

# Fitness for Duty in the Nuclear Industry: Update of the Technical Issues 1996

---

---

Manuscript Completed: April 1996  
Date Published: May 1996

Edited by  
N. Durbin and T. Grant, BSRC

## Contributors

A. Bittner, N. Durbin, I. Field, C. Forslund, T. Grant, J. Hauth, J. Macaulay, C. Moore, C. Orians, J. Toquam,  
M. Silbermagel, R. Wilson, BSRC  
B. Baxter<sup>1</sup>, S. Kimme<sup>2</sup>, A. Zebelman<sup>3</sup>

Battelle Seattle Research Center  
4000 NE 41st Street  
Seattle, Washington 98105

Pacific Northwest National Laboratory  
Richland, Washington 99352

L. Bush, NRC Project Manager

Prepared for  
Division of Reactor Program Management  
Office of Nuclear Reactor Regulation  
U.S. Nuclear Regulatory Commission  
Washington, DC 20555-0001  
NRC Job Code I2007

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

---

1. Alcohol and Drug Abuse Institute, University of Washington, Seattle, WA 98105
2. Center for the Disability Policy and Research, University of Washington, Seattle, WA 98105
3. Dynacare Laboratory of Pathology of Seattle, Inc., Seattle, WA 98104

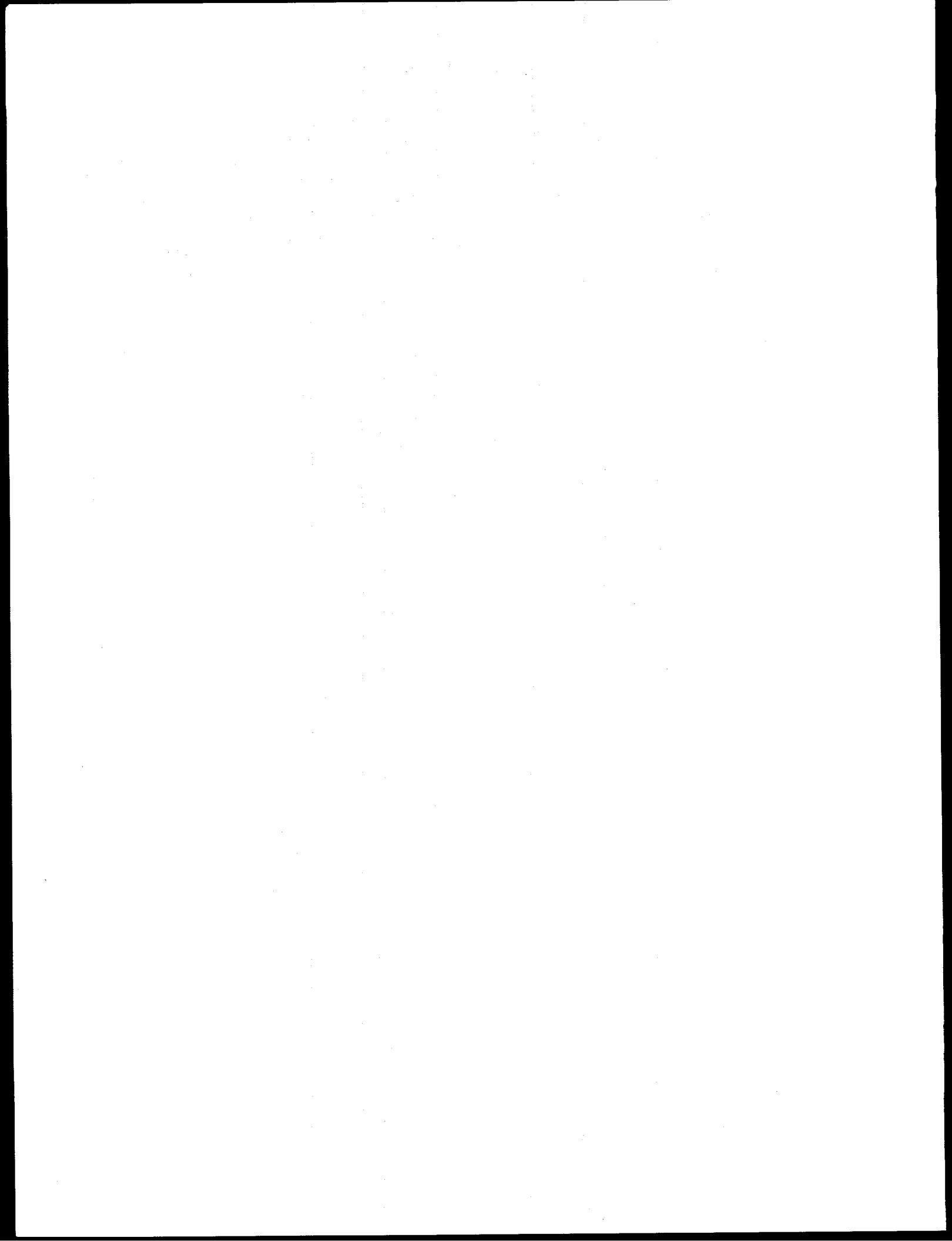
**DISCLAIMER**

**Portions of this document may be illegible  
in electronic image products. Images are  
produced from the best available original  
document.**

## ABSTRACT

The purpose of this report is to provide an update of information on the technical issues surrounding the creation, implementation, and maintenance of fitness-for-duty (FFD) policies and programs. It has been prepared as a resource for Nuclear Regulatory Commission (NRC) and nuclear power plant personnel who deal with FFD programs. It contains a general overview and update on the technical issues that the NRC considered prior to the publication of its original FFD rule and the revisions to

that rule (presented in earlier NUREG/CRs). It also includes chapters that address issues about which there is growing concern and/or about which there have been substantial changes since NUREG/CR-5784 was published. Although this report is intended to support the NRC's rule making on fitness for duty, the conclusions of the authors of this report are their own and do not necessarily represent the opinions of the NRC.



# TABLE OF CONTENTS

## CHAPTER 1: INTRODUCTION

1.1: Background .....	1-1
1.2: Overview of the Report .....	1-1
Endnotes .....	1-4

## CHAPTER 2: GENERAL REVIEW OF THE TECHNICAL ISSUES CONSIDERED IN THE DEVELOPMENT AND IMPLEMENTATION OF THE NUCLEAR REGULATORY COMMISSION'S FITNESS-FOR-DUTY REGULATION

2.1: Background and Organization .....	2-1
2.2: Drug Use Prevalence .....	2-1
2.3: Legal Drugs .....	2-2
2.4: Legal and Regulatory Issues .....	2-3
2.5: Fitness-for-Duty Program Issues .....	2-4
2.5.1: Comprehensive Fitness-for-Duty Programs .....	2-4
2.5.2: Random Testing .....	2-5
2.5.3: Deterrence .....	2-5
2.5.4: Medical Review Officer Qualifications and Duties .....	2-6
2.5.5: Chain of Custody .....	2-7
2.5.6: Subversion .....	2-9
2.5.7: Employee Assistance Programs .....	2-10
2.5.7.1: EAP Scope of Services .....	2-10
2.5.7.2: Confidentiality .....	2-11
2.5.7.3: Evaluation Efforts .....	2-11
2.5.7.4: Managed Care .....	2-12
2.5.8: Return to Work .....	2-13
2.6: Drug Testing Issues .....	2-14
2.6.1: Department of Health and Human Services Changes .....	2-14
2.6.2: On-Site Screening Issues .....	2-14
2.6.3: Opiate Testing Requirements .....	2-15
2.6.3.1: 6-AM Requirements .....	2-14
2.6.3.2: Opiate Screening Level .....	2-16
2.6.4: Reliability of the Testing Process .....	2-16
2.6.5: Drug Retention Over Time .....	2-18
2.6.6: Impairment .....	2-18
2.6.7: Deterioration of Metabolites in Urine Specimens .....	2-19
2.6.8: Extrapolation Issues .....	2-19
2.6.9: Alternative Testing Methods .....	2-20
2.6.9.1: Alternatives to Urinalysis Drug Testing .....	2-20
Saliva Analysis .....	2-21
Hair Analysis .....	2-21
Sweat Analysis .....	2-23
Blood Analysis .....	2-23
New Developments in Urinalysis .....	2-24

## Table of Contents

2.6.9.2: Alternatives to Breath Alcohol Testing .....	2-25
Saliva Analysis .....	2-25
Blood Analysis .....	2-25
New Developments in Breath Tests .....	2-25
2.6.9.3: Performance-Based Assessment Tests .....	2-26
2.6.10: Shy Bladder .....	2-27
2.6.11: Alcohol Testing Issues .....	2-27
2.6.12: Occupational Risks in Drug Testing .....	2-28
2.6.12.1: Exposure Control .....	2-28
2.6.12.2: Methods of Compliance .....	2-28
2.6.12.3: Hepatitis B Vaccination and Post-Exposure Evaluation and Follow-Up .....	2-29
2.6.12.4: Communication of Hazards to Employees .....	2-29
2.6.12.5: Recordkeeping .....	2-29
2.7: A Review of Findings from NRC Inspections .....	2-29
Endnotes .....	2-31
References .....	2-32

## CHAPTER 3: PREVALENCE OF DRUG USE IN THE UNITED STATES

3.1: Introduction .....	3-1
3.2: Prevalence of Illegal Drug Use and Alcohol Abuse Among Americans .....	3-1
3.2.1: The National Household Survey on Drug Abuse (NHSDA) .....	3-2
3.2.1.1: Illegal Drug Use: Discussion of NHSDA Findings .....	3-3
3.2.1.2: Alcohol Use: Discussion of NHSDA Findings .....	3-4
3.2.1.3: Differences Across Population Categories: Discussion of NHSDA Findings .....	3-5
3.2.2: The Monitoring the Future (MTF) Survey of College Students and Young Adults .....	3-5
Summary and Discussion of MTF Findings (Compared with NHSDA Findings) .....	3-6
3.2.3: The Drug Abuse Warning Network (DAWN) Data System of Drug-Related Emergency Room Episodes .....	3-6
Summary and Discussion of DAWN Findings .....	3-8
3.2.4: The Pulse Check Profile of National Trends in Drug Abuse .....	3-8
Summary and Discussion of Pulse Check Findings .....	3-9
3.2.5: Conclusions Regarding the Prevalence of Drug Use Among Americans .....	3-9
3.2.5.1: Examination of Interstudy Differences in Findings .....	3-9
3.2.5.2: Conclusions Regarding Study Findings .....	3-10
3.3: Prevalence of Illegal Drug Use and Alcohol Abuse in the American Workforce .....	3-12
3.3.1: Workforce Illegal Drug Use and Heavy Alcohol Use Prevalence Across Employment Status .....	3-12
3.3.2: Workforce Illegal Drug Use and Heavy Alcohol Use Prevalence Across Income .....	3-12
3.3.3: Workforce Illegal Drug Use and Heavy Alcohol Use Prevalence Across Industry .....	3-12
3.3.4: Workforce Alcohol Problems Across Occupation .....	3-13

3.4: Prevalence of Illegal Drug Use and Alcohol Abuse On the Job in the American Workplace .....	3-15
3.4.1: Workplace Drug Use Prevalence Across Occupation .....	3-16
3.4.2: Workplace Drug Use Prevalence in the Nuclear Industry .....	3-16
3.5: Prevalence of Drug Use Detected in Workplace Drug Testing .....	3-17
3.5.1: Test Results Under NRC's FFD Rule .....	3-18
3.5.2: Drug Testing Index Results .....	3-19
3.5.3: NIDA Test Results .....	3-21
3.5.4: AMA Test Results .....	3-21
3.5.5: Summary of Workplace Drug Test Results .....	3-21
3.6: Chapter Conclusions and Implications for Drug Testing in the Nuclear Power Industry .....	3-22
3.6.1: Chapter Conclusions .....	3-22
3.6.2: Implications for Drug Testing in the Nuclear Power Industry .....	3-23
3.6.2.1: Panel of Drugs .....	3-23
Recommendation .....	3-23
Supporting Evidence .....	3-23
3.6.2.2: Balanced Approach to Illegal Drug Use and Abuse of Legal Drugs .....	3-23
Recommendation .....	3-23
Supporting Evidence .....	3-24
3.6.2.3: Program Flexibility with Regard to Workforce Differences .....	3-24
Recommendation .....	3-24
Supporting Evidence .....	3-24
3.6.2.4: Currency of Data on Drug Use Prevalence .....	3-24
Recommendation .....	3-24
Supporting Evidence .....	3-24
Endnotes .....	3-26
References .....	3-28
<b>CHAPTER 4: LEGAL DRUGS</b>	
4.1: Introduction .....	4-1
4.2: The Use, Misuse, and Abuse of Legal (Prescription and OTC) Drugs .....	4-1
4.3: Potential Impairing Effects of Legal Drugs .....	4-2
4.4: Dangerous Legal Drug Combinations .....	4-3
4.5: Legal Drugs of the "90s" .....	4-3
4.5.1: Prozac® and Other New Antidepressants .....	4-4
4.5.2: Steroids .....	4-4
4.6: Conclusion .....	4-7
Endnotes .....	4-8
References .....	4-9

## Table of Contents

### CHAPTER 5: THE DETERRENT EFFECT OF FITNESS-FOR-DUTY PROGRAMS

5.1: Introduction . . . . .	5-1
5.2: Deterrence and Drug Use . . . . .	5-1
5.3: The Deterrent Effect of FFD Programs . . . . .	5-3
5.3.1: Random Testing . . . . .	5-4
5.3.1.1: Random Testing Rates and Detection of Drug Use . . . . .	5-5
5.3.1.2: Selection Methods . . . . .	5-7
5.3.1.3: Other Selection Practices . . . . .	5-9
5.3.1.4: Summary . . . . .	5-9
5.3.2: Pre-Employment Testing . . . . .	5-9
5.3.3: For-Cause Testing . . . . .	5-9
5.3.4: Periodic Testing . . . . .	5-10
5.3.5: Sanctions . . . . .	5-10
5.3.6: Integrity of the Testing Process and Program Administration . . . . .	5-10
5.3.7: Education, Training, and Awareness of Risks . . . . .	5-10
5.3.8: Employee Assistance Programs . . . . .	5-11
5.3.9: Organizational Influences on Deterrence . . . . .	5-11
5.4: Conclusions . . . . .	5-12
Endnotes . . . . .	5-14
References . . . . .	5-15

### CHAPTER 6: SUBVERSION AND COUNTERSUBVERSION

6.1: Introduction . . . . .	6-1
6.2: Subversion Techniques . . . . .	6-1
6.2.1: Specimen Substitution . . . . .	6-2
6.2.2: Specimen Dilution . . . . .	6-3
6.2.3: Specimen Adulteration . . . . .	6-4
6.3: Evidence of and Experience with Subversion . . . . .	6-4
6.3.1: <i>In Vitro</i> Specimen Substitution . . . . .	6-5
6.3.2: Specimen Hydration and Dilution . . . . .	6-5
6.3.3: Other Evidence of Subversion . . . . .	6-6
6.3.4: Summary . . . . .	6-7
6.4: Countersubversion Strategies . . . . .	6-7
6.4.1: Overview of Countersubversion Measures . . . . .	6-7
6.4.2: Preventive Countermeasures . . . . .	6-7
6.4.3: Detection Measures . . . . .	6-9
6.4.3.1: Observation Techniques . . . . .	6-9
6.4.3.2: Nonanalytical Examination of Specimens . . . . .	6-9

6.4.3.3: Analytical Examination of Specimens .....	6-9
Temperature .....	6-9
pH .....	6-10
Specific Gravity .....	6-10
Creatinine Testing .....	6-11
Tests to Detect UrinAid® .....	6-11
Ionic Strength .....	6-12
Positive Control Assay .....	6-12
6.4.3.4: Lower Cutoff Levels .....	6-12
6.4.4: Discussion of Analytic Countersubversion Measures .....	6-12
6.5: Subversion Issues .....	6-13
6.5.1: Measuring Prevalence .....	6-13
6.5.2: Workers' Rights .....	6-14
6.6: Summary and Conclusions .....	6-14
Endnotes .....	6-16
References .....	6-17

## CHAPTER 7: FITNESS FOR DUTY: RETURN TO WORK

7.1: Introduction .....	7-1
7.2: Categories of Users .....	7-2
7.3: Treatment .....	7-2
7.3.1: Treatment Approaches .....	7-2
7.3.1.1: Medical or Physical Models .....	7-3
7.3.1.2: Psychological Models .....	7-3
7.3.1.3: Social Models .....	7-3
7.3.2: Treatment Goals .....	7-4
7.3.3: Treatment Selection Process .....	7-4
7.3.3.1: Diagnosis .....	7-5
7.3.3.2: Limitations of Patient Matching .....	7-6
7.4: Treatment Outcomes .....	7-7
7.4.1: Definitions of Rehabilitation, Recovery, and Relapse .....	7-7
7.4.2: Outcome Measures .....	7-8
7.4.3: Treatment Outcome Studies .....	7-8
7.4.4: Limitations of Outcome Studies .....	7-9
7.5: Assessment of Fitness to Return to Work .....	7-10
7.5.1: Declaration of Fitness .....	7-10
7.5.2: Continued Monitoring of Fitness .....	7-10

## Table of Contents

7.6: Conclusion .....	7-11
Endnotes .....	7-13
References .....	7-14

### CHAPTER 8: DRUG RETENTION OVER TIME

8.1: Introduction .....	8-1
8.2: General Drug Metabolism Issues .....	8-1
8.3: Marijuana (Cannabinoids) .....	8-2
8.3.1: The Occasional User .....	8-3
8.3.2: The Moderate User .....	8-3
8.3.3: The Heavy (Chronic) User .....	8-3
8.3.4: Potential for Successive Positive Drug Test Results for Marijuana .....	8-4
8.4: Cocaine .....	8-4
8.5: Alcohol .....	8-5
8.6: Opiates .....	8-6
8.7: Amphetamines .....	8-8
8.8: Phencyclidine (PCP) .....	8-9
8.9: Summary and Conclusions .....	8-9
Endnotes .....	8-11
References .....	8-12

### CHAPTER 9: EFFECT OF STORAGE TEMPERATURE AND EXTREME pH LEVELS ON MARIJUANA AND COCAINE METABOLITE CONCENTRATIONS

9.1: Introduction .....	9-1
9.2: Specimen Stability Pilot Study .....	9-2
9.2.1: Approach .....	9-2
9.2.2: Results .....	9-2
9.2.2.1: Immunoassay Results .....	9-2
9.2.2.2: GC/MS Results .....	9-4
9.3: Continuation of Specimen Stability Pilot Study .....	9-4
9.3.1: Approach .....	9-4
9.3.2: Results .....	9-5
9.4: Discussion .....	9-5
9.5: Conclusion .....	9-6
Endnotes .....	9-7
References .....	9-8

## CHAPTER 10: PERFORMANCE-BASED ASSESSMENT TESTS

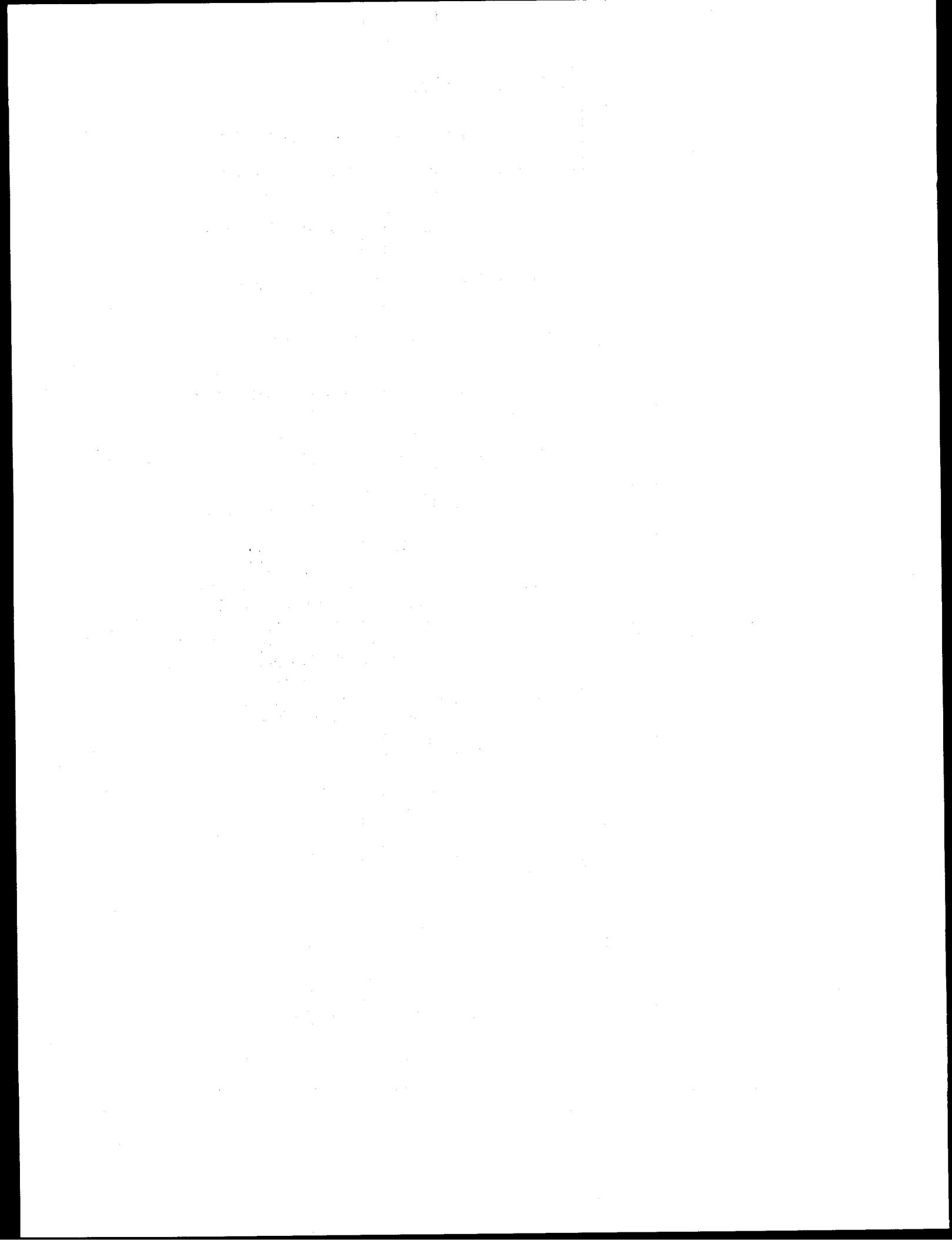
10.1: Scope and Optimal Characteristics of Performance-Based Tests .....	10-1
10.1.1: Scope of Performance-Based Tests .....	10-1
10.1.2: Optimal Characteristics of Performance-Based Tests .....	10-2
10.1.2.1: Reliability .....	10-3
10.1.2.2: Sensitivity .....	10-4
10.1.2.3: Validity .....	10-4
10.1.2.4: Accuracy .....	10-5
10.2: Performance-Based Testing Literature .....	10-5
10.2.1: Review of Performance-Based Testing Literature .....	10-5
10.2.1.1: Cognitive Tests .....	10-6
Attention/Divided Attention .....	10-6
Time Estimation .....	10-6
Perceptual Speed .....	10-6
Memory .....	10-9
Perception .....	10-9
Spatial Ability .....	10-10
Reasoning .....	10-10
Fluency .....	10-10
Number Facility .....	10-10
Reaction Time .....	10-10
10.2.1.2: Psychomotor Tests .....	10-10
10.2.1.3: Physiological Tests .....	10-11
Vision and Visual Acuity .....	10-11
Ocular Movement .....	10-11
10.2.1.4: Summary of Performance-Based Testing Literature .....	10-11
10.2.2: Performance-Based Tests Available for Implementation .....	10-11
10.2.2.1: SEDI Corporation .....	10-12
10.2.2.2: Essex Corporation .....	10-12
10.2.2.3: Nova Technology Incorporated .....	10-12
10.2.2.4: Performance Factors Incorporated .....	10-12
10.2.2.5: Eye Dynamics, Inc. .....	10-12
10.2.2.6: Summary of Currently Available Performance-Based Tests .....	10-13
10.3: Performance-Based Test Accuracy .....	10-13
10.3.1: Test Accuracy in Quantitative Terms .....	10-13
10.3.2: Selecting a Test for Accuracy Calculations .....	10-14
Computing Test Accuracy ( $d'$ ) .....	10-14
10.3.3: Accuracy Evaluation of the Selected Test .....	10-14
10.4: Summary and Implications for the Nuclear Industry .....	10-15
10.4.1: Overall Summary of the Status of Performance-Based Tests .....	10-15

## Table of Contents

10.4.2: Implementing Performance-Based Tests in Industry .....	10-16
10.4.2.1: Linking Performance-Based Tests to Job Performance Requirements .....	10-16
10.4.2.2: Procedures for Implementing a Performance-Based Test Battery .....	10-17
10.4.2.3: Implications of a Test Failure .....	10-17
10.4.2.4: Conclusions .....	10-17
Endnotes .....	10-18
References .....	10-20
APPENDIX A: DEPARTMENT OF HEALTH AND HUMAN SERVICES CHANGES .....	A-1
APPENDIX B: NON-INSTRUMENTED SCREENING TEST DEVICES .....	B-1
APPENDIX C: REQUIREMENTS OF THE OCCUPATIONAL EXPOSURE TO BLOODBORNE PATHOGENS RULE .....	C-1
APPENDIX D: VIOLATIONS, NON-CITED VIOLATIONS, AND UNRESOLVED ITEMS FROM THE NRC INSPECTION REPORTS .....	D-1
APPENDIX E: METHODOLOGICAL STRENGTHS AND LIMITATIONS OF SELECTED MAJOR DRUG USE SURVEYS .....	E-1
APPENDIX F: SUBVERSION ISSUES RELEVANT TO THE PROPOSED REVISIONS TO 10 CFR PART 26 .....	F-1
APPENDIX G: SUMMARIES FOR PERFORMANCE-BASED TESTS IDENTIFIED IN THE PUBLISHED LITERATURE .....	G-1

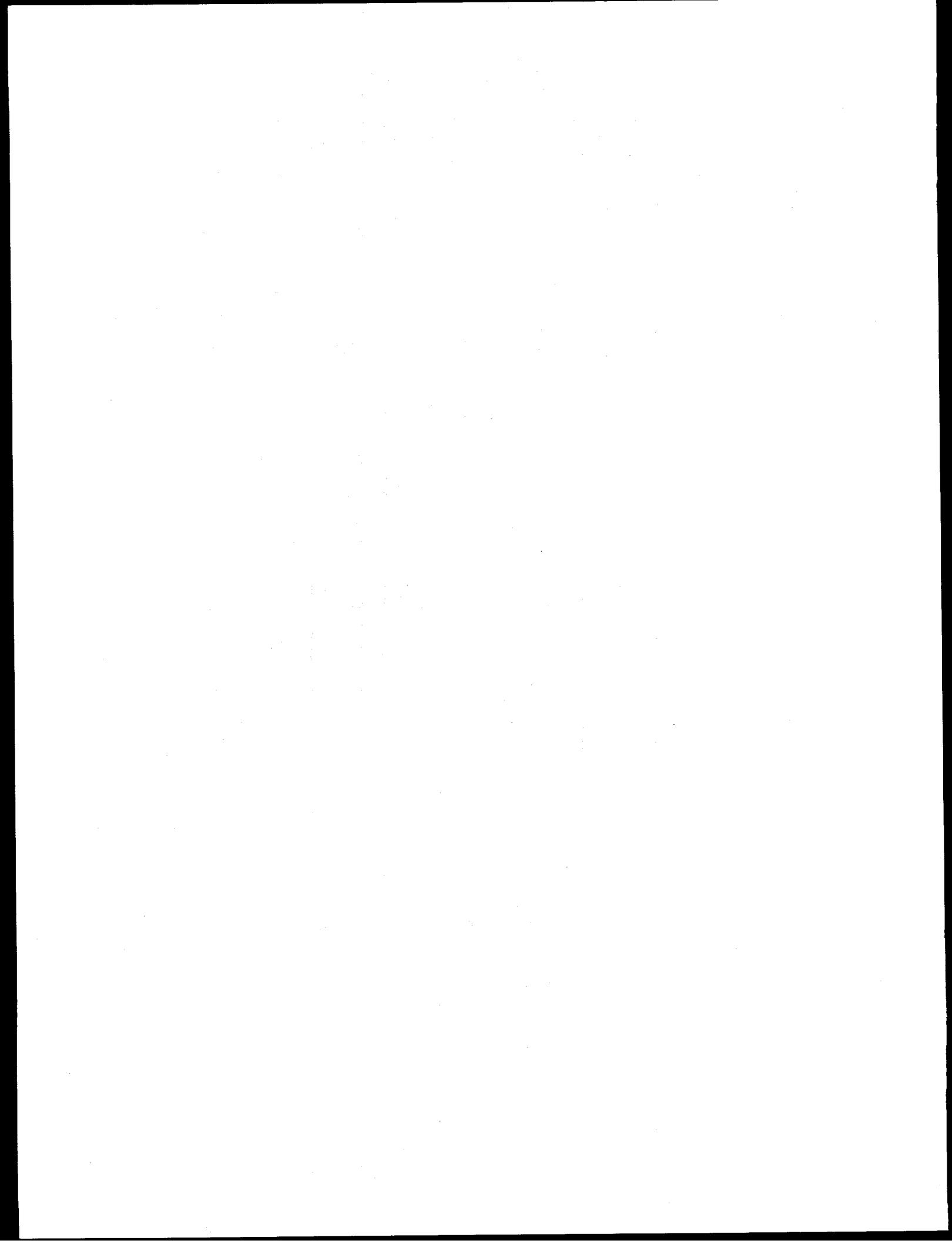
## LIST OF FIGURES

Figure 3.1: Trends in Americans' Self-Reported Past-Month Drug Use, 1985-1993 . . . . .	3-3
Figure 5.1: Simple and Modified Random Testing at 100% Annual Testing Rate . . . . .	5-8



## LIST OF TABLES

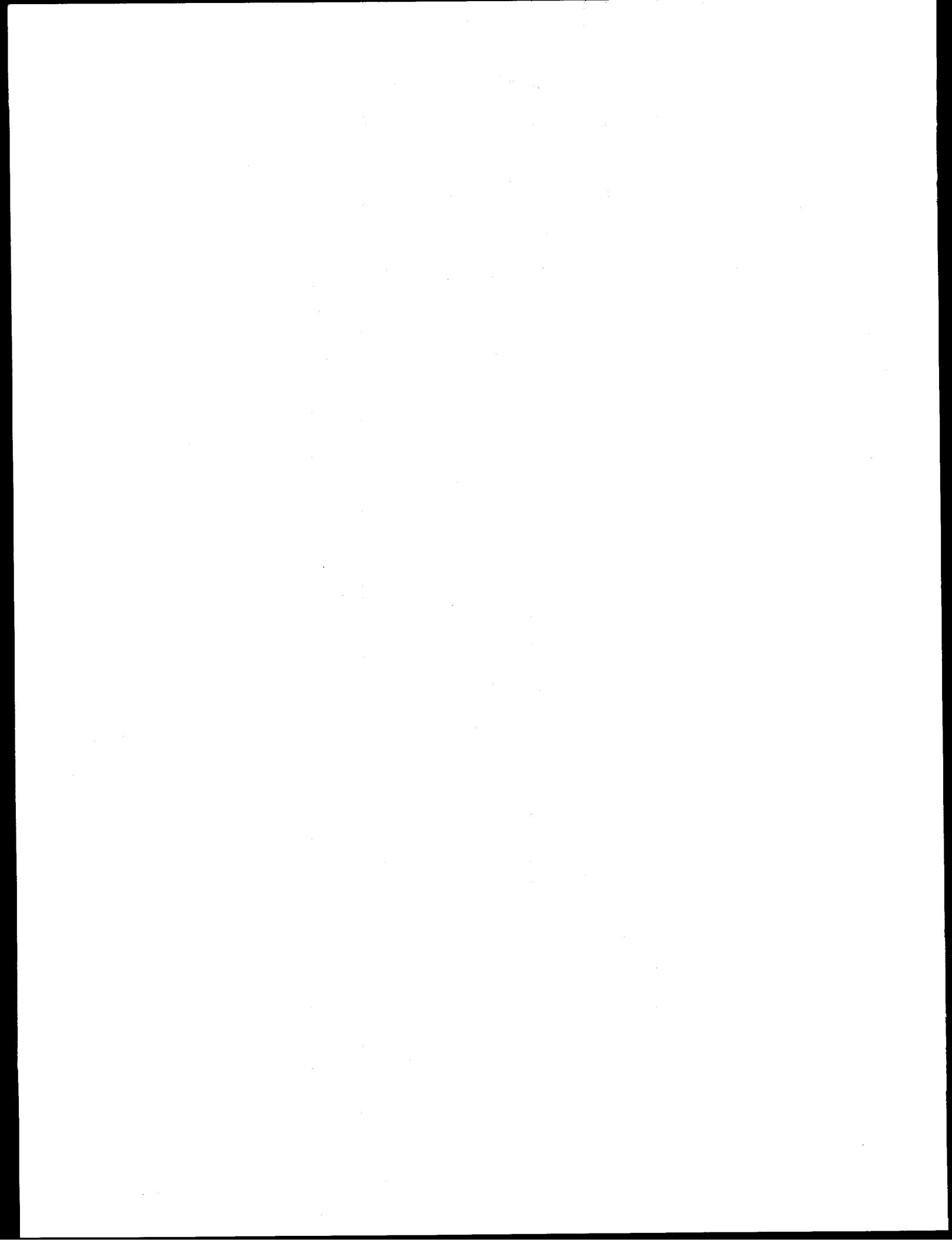
Table 2.1: Summary of Unsatisfactory Testing Results (January 3, 1990 through December 31, 1992) .....	2-17
Table 3.1: Annual (Past Year) Prevalence Rates, for Selected Drug Categories, Based on Self-Reported Results .....	3-6
Table 3.2: Past-Year Use of Illicit Drugs and Alcohol among the General Population and College Students in Particular .....	3-7
Table 3.3: Prevalence Rates of Self-Reported Drug Use, for Selected Drug Categories, by Employment Status and Age: 1988, 1990, 1992 .....	3-13
Table 3.4: Prevalence Rates of Self-Reported Drug Use Among Full-Time Workers, for Selected Drug Categories, by Income and Gender: 1988, 1990 .....	3-14
Table 3.5: Prevalence Rates of Self-Reported Drug Use Among Full-Time Male Workers, Aged 18-34, for Selected Drug Categories, by Selected Industries: 1988, 1990 .....	3-14
Table 3.6: Trends in Alcohol and Other Drug Use at Work, Past 12 Months, Young Adults Aged 19-28 Years .....	3-16
Table 3.7: Use of Alcohol and Other Drugs at Work in Past 12 Months, Young Adults Aged 19-28 by Occupational Category .....	3-17
Table 3.8: Positive Test Rates: Comparison of Nuclear Power Industry Results .....	3-19
Table 3.9: Positive Test Rates by Testing Category: Comparison of Nuclear Power Industry Results and SmithKline Beecham Drug Testing Index Results for Safety-Sensitive Transportation Workers (SSTWs) .....	3-20
Table 3.10: Distribution of Positive Test Results by Drug Category: Comparison of Nuclear Power Industry Results and SmithKline Beecham Drug Testing Index Results for Safety-Sensitive Transportation Workers (SSTWs) .....	3-20
Table 5.1: Expected Detection Rates for Random Testing at 25% Test Rate and 50% Test Rate .....	5-6
Table 5.2: FRA Industry-Wide Random Drug Testing Results .....	5-7
Table 5.3: Probability of Being Selected According to Different Annual Random Testing Rates .....	5-7
Table 6.1: Examples of the Six Types of Subversion .....	6-2
Table 6.2: Subversion Techniques and Countersubversion Measures .....	6-8
Table 8.1: Summary of Approximate Drug Detection Time in Days .....	8-10
Table 9.1: Experimental Results: EMIT® Immunoassay .....	9-3
Table 9.2: Experimental Results: Gas Chromatography/Mass Spectroscopy (GC/MS) .....	9-4
Table 9.3: Benzoyllecgonine (BE) Concentration Following Storage at 37°C for Various Times .....	9-5
Table 10.1: Summary of Performance-Based Test Information .....	10-7
Table 10.2: Evaluation of Two-Hand Tapping Test Accuracy: Proportion Detected of BAC Equivalent Impairments for Two Test Lengths and Two False Alarm Rates .....	10-15



## ACKNOWLEDGEMENTS

We would like to acknowledge the help and cooperation of many individuals who supported the preparation of this report. We would like to especially thank Mr. Loren Bush, project technical monitor, for his patience, guidance, and cooperation during the production of this report. In addition, we would like to thank the many other Nuclear Regulatory Commission staff members who supported this effort. Ms. Nancy Sutherland, research librarian, and the staff of the Alcohol and Drug Abuse Institute in Seattle, Washington generously shared resources with us and provided valuable assistance. In addition, the library staff at the Battelle Seattle Research Center—Ms. Karen Aoyama, Ms. Jan Schueller, and Ms.

Miyoko Wolf—provided critical and timely support. Mr. Jon Brichoux, Ms. Kate Lynch, Ms. Susan Hauth, and Mrs. Barbara Hollosi provided much needed and appreciated support and assistance throughout the project and in producing this report. We would also like to thank Ms. Toni Slavich and Mr. Mitchel Cunningham for their careful and thoughtful review of the document. Finally, we would like to thank the many individuals who either in person or on the phone provided their time and valuable information. Although the individuals acknowledged here made this report possible, they are not responsible for the quality or the content of this report.



## ACRONYMS<sup>1</sup>

6-AM	6-Acetylmorphine
6-MAM	6-Monoacetylmorphine
AA	Alcoholics Anonymous
AAMRO	American Association of Medical Review Officers
ADA	Americans with Disabilities Act
AIDS	Acquired Immune Deficiency Syndrome
AMA	American Management Association
AMA	American Medical Association
APA	American Psychiatric Association
APTS	Automated Portable Test System
AUDADIS	Alcohol Use Disorders and Associated Disabilities Interview Schedule
BAC	Blood Alcohol Concentration
BE	Benzoyllecgonine
BZ	Benzodiazepine
CDC	Centers for Disease Control and Prevention
CFR	Code of Federal Register
CLIA	Clinical Laboratory Improvements Amendments
CNS	Central Nervous System
CPL	Conforming Procedures List
CRIT	approach of the Criteria Committee of the National Council on Alcoholism
CY	Calendar Year
DARP	Drug Abuse Reporting Program
DAST	Drug Abuse Screening Test
DATOS	Drug Abuse Treatment Outcome Study
DAWN	Drug Abuse Warning Network
DHHS	Department of Health and Human Services
DIS	Diagnostic Interview Schedule
DOD	Department of Defense
DOT	Department of Transportation
DRE	Drug Recognition Expert
DSM	Diagnostic and Statistical Manual of Mental Disorders
DSM-III	Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition
DSM-III-R	DSM-III, Revised
DUI	Driving Under the Influence
EAP	Employee Assistance Program
EDI	Eye Dynamics, Inc.
EEOC	Equal Employment Opportunity Commission
EHAP	Employee and Health Assistance Program
EIA	Enzyme Immunoassay
EME	Ecgonine Methyl Ester
EMIT®	Enzyme Multiplied Immunoassay Test
ER	Emergency Room
FAA	Federal Aviation Administration
FBI	Federal Bureau of Investigation
FDA	Food and Drug Administration
FFD	Fitness for Duty
FR	Federal Register
FRA	Federal Railroad Administration
FRN	Federal Register Notice
GC/MS	Gas Chromatography/Mass Spectrometry
GSA	General Services Administration
HBV	Hepatitis B Virus

<sup>1</sup> Due to the nature of this report, acronyms are defined in each chapter.

## Acronyms

HIV	Human Immunodeficiency Virus
HPP	Health Promotion Program
HR33	Drug Testing Quality Act
INS	Immigration and Naturalization Service
LOD	Level of Detection
MAO	Monoamine Oxidase
MAST	Michigan Alcoholism Screening Test
MMPI	Minnesota Multiphasic Personality Inventory
MMWR	Morbidity and Mortality Weekly Report
MODCRIT	modified CRIT
MRO	Medical Review Officer
MTF	Monitoring the Future
NCAA	National Collegiate Athletics Association
NCCLS	National Committee for Clinical Laboratory Standards
NFL	National Football League
NHIS	National Health Interview Survey
NHSDA	National Household Survey on Drug Abuse
NHTSA	National Highway Traffic Safety Administration
NIDA	National Institute on Drug Abuse
NIJ	National Institute of Justice
NIST	National Institute of Standards and Technology
NLCP	National Laboratory Certification Program
NLRB	National Labor Relations Board
NORML	National Organization for the Reform of Marijuana Laws
NRC	Nuclear Regulatory Commission
NTI	Nova Technology Incorporated
NUMARC	Nuclear Management and Resources Council
ONDCP	Office of National Drug Control Policy
OSHA	Occupational Safety and Health Administration
OSS	Operations Shift Supervisor
OTC	Over the Counter
PCP	Phencyclidine
PETER	Performance Evaluation Tests for Environmental Research
PFI	Performance Factors Incorporated
PNNL	Pacific Northwest National Laboratory
PPA	Phenylpropanolamine
QA	Quality Assurance
RIA	Radioimmunoassay
RSPA	Research and Special Programs Administration
SADS	Schedule for Affective Disorders and Schizophrenia
SAMHSA	Substance Abuse and Mental Health Services Administration
SCID	Structured Clinical Interview for DSM-III-R
SEDI	Simulated Evaluation of Drug Impairment
SG	Specific Gravity
SOFT	Society of Forensic Toxicologists
SSRI	Selective Serotonin Reuptake Inhibiting
SSTW	Safety-Sensitive Transportation Workers
TCA	Tricyclic Antidepressant
THC	D-9-tetrahydrocannabinol
TI	Temporary Instruction
TLC	Thin Layer Chromatography
TOPS	Treatment Outcome Prospective Study
UTS-PAB	Unified Tri-Services Performance Assessment Battery

# CHAPTER 1: INTRODUCTION

## Nancy Durbin and Tom Grant<sup>1</sup>

Programs to assure staff fitness for duty have become increasingly required in safety sensitive industries such as the commercial nuclear power industry. These fitness-for-duty (FFD) programs can involve drug and alcohol testing (including pre-employment, for-cause, random, return-to-service, and follow-up testing); employee assistance programs (EAPs); education of workers and managers about FFD problems (including fatigue, stress, and alcohol and drug abuse); training of workers and managers in the recognition of FFD problems; sanctions (including removal from work either temporarily or permanently); and policies and programs for returning employees to work after recovery from an FFD problem. People who develop and implement FFD policies and programs need information on research and experience related to these areas. Unfortunately, there is little research directly applicable to specific FFD policy and program issues, and the research that is available is often written for the research community rather than for practitioners. Therefore, the U.S. Nuclear Regulatory Commission (NRC) has supported the preparation of this report to provide a compilation and interpretation of current research and experience relevant to the implementation of its regulation of commercial nuclear power industry FFD programs.

The purpose of this report is to provide an update of information on the technical issues surrounding the creation, implementation, and maintenance of FFD policies and programs. This report has been prepared as a resource for NRC and nuclear power plant personnel who deal with FFD programs. It contains a general overview and update on the technical issues that the NRC considered prior to the publication of its original FFD rule and the revisions to that rule. It also includes chapters that address issues about which there is growing concern.

This report is intended to support the NRC's rule making on fitness for duty. However, the conclusions of the authors of this report are their own and do not necessarily represent the opinions of the NRC.

### 1.1: Background

Since 1989, the NRC has developed and implemented a program to address concerns about potential impacts on public health and safety of FFD problems among personnel with unescorted access to protected areas in commercial nuclear power plants. As part of this program development, the NRC published a rule that requires its power reactor license holders to implement

FFD programs applicable to such personnel. This rule became effective upon publication (*Federal Register*, 54 FR 24468, June 7, 1989; 10 CFR Part 26, Fitness-for-Duty Programs), with full implementation required by January 3, 1990. The rule establishes broad requirements for the control of FFD problems stemming from illegal drug use, alcohol abuse, abuse of legal drugs, or from any other mental or physical problems that could impair performance or that in other ways create concerns about the reliability and trustworthiness of employees having unescorted access to protected areas of nuclear power plants. Due to the nature of this regulation, the NRC contracted with the Pacific Northwest National Laboratory (PNNL)<sup>2</sup> to examine the technical background associated with the FFD rule. This information was published in NUREG/CR-5227; NUREG/CR-5227 Supplement 1; NUREG/CR-5758, Volumes 1-5; and NUREG/CR-5784.

As part of its original rulemaking, the Commission required its staff to continue to review and evaluate licensee programs, to assess the effectiveness of the rule, and to recommend appropriate rule revisions. The Commission also tasked the staff to monitor technical developments in case alternatives to urinalysis and random testing became viable. In November, 1991, the Commission strongly encouraged its staff to continue considerations of implementation experiences to identify areas of the FFD rule where amendments may be needed. In keeping with these directives, the staff has continuously reviewed licensee FFD programs' performance, policy, procedures, legal issues, and technical developments. As part of its ongoing activities to improve its regulations, the NRC has made various revisions to 10 CFR Part 26<sup>3</sup> and is currently considering additional revisions that would have a substantial impact on the rule. These proposed revisions will be published in the *Federal Register* for public comment. This report provides technical background on a wide range of issues surrounding the implementation of the NRC's FFD program by the Commission's licensees, focusing on issues of particular relevance to those proposed revisions.

### 1.2: Overview of the Report

This report contains a general overview of all of the issues that were considered by the NRC prior to the publication of 10 CFR Part 26; subsequently discussed in NUREG/CR 5227, NUREG/CR 5227-Supplement 1, and NUREG/CR 5784; or which have become of increasing concern since this latter NUREG/CR was published in

## Introduction

September, 1991. Below is a brief synopsis of the chapters that follow this introductory chapter.

Chapter 2, "General Review of the Technical Issues Considered in the Development and Implementation of the Nuclear Regulatory Commission's Fitness-for-Duty Regulation," reviews the issues surrounding the NRC's FFD program considered in previous NUREG/CRs and discusses any new findings or concerns that have emerged during recent years. When appropriate, it refers to the more in-depth treatment of the issue(s) in previous volumes and/or in subsequent chapters. Highlights of this chapter include a review of the recent legal and regulatory activities that may have an impact on implementation of the NRC's FFD rule, the impact of managed care programs on EAPs, and recent changes to Department of Health and Human Services (DHHS) drug testing guidelines.

Chapter 3, "Prevalence of Drug Use in the United States," summarizes the current survey findings regarding drug and alcohol abuse in the U.S., the types and prevalence of drug testing programs in U.S. industry, and the outcomes of testing programs. A key finding reported in this chapter states that the overall self-reported rate of illegal drug use and alcohol use has declined in the U.S. It also notes that marijuana, cocaine, and alcohol are the drugs that are most commonly reported in self-report surveys and are most commonly detected in commercial nuclear power industry FFD programs.

Chapter 4, "Legal Drugs," reviews the effects of legal (prescription and over-the-counter [OTC]) drugs. It also provides a discussion of two types of legal drugs that have been the focus of regulatory and media attention recently: the new antidepressants (e.g., Prozac<sup>®</sup>) and anabolic steroids.

Chapter 5, "The Deterrent Effects of Fitness for Duty Programs," discusses the potential deterrent effects of FFD programs, reviewing the expected effect of each program element on deterrence. The chapter provides the reader with a discussion of the issues surrounding deterrence and the information that is currently available regarding the impact of individual and combined elements of FFD programs on deterrence. Overall, the chapter concludes that deterrent strategies that ensure that the risks of the use of illegal drugs and the abuse of alcohol and legal drugs are well understood, support an organizational culture that is intolerant of drug and alcohol abuse, and provide support and assistance to those seeking help are most effective.

Chapter 6, "Subversion and Countersubversion," covers common means used to subvert the testing process and methods that can thwart or discover subversion attempts. The chapter discusses *in vivo* and *in vitro* techniques for substituting, diluting and adulterating samples; evidence of subversion; and counter-subversion strategies. Countersubversion measures include preventative measures such as security and control of specimens and detection measures such as checking the temperature and pH of a specimen. The chapter also includes a discussion of balancing the need to prevent or detect subversion with the protection of workers' rights.

Chapter 7, "Fitness for Duty: Return to Work," addresses key questions faced by managers of FFD programs that return workers to the job after those workers have tested positive for illegal use of drugs or abuse of alcohol and have undergone treatment. This chapter reports on research regarding recidivism rates and other factors related to return-to-work issues, discusses the relevance of the current research findings in addressing return-to-work questions in a work setting such as the nuclear industry, and provides a summary of general guidelines that might be derived from this literature. Key recommendations of the chapter include the involvement of professionals throughout the process and the centrality of professional judgment in making return-to-work decisions.

Chapter 8, "Drug Retention Over Time," summarizes the literature regarding how long drug metabolites remain in the body and are detectable in urine samples or breath samples. This chapter highlights the difficulties of detecting occasional abuse of drugs, including alcohol.

Chapter 9, "Effects of Storage Temperature and Extreme pH Levels on Marijuana and Cocaine Metabolite Concentrations," reports the findings of a pilot experiment examining the effects of time and temperature on the quality of specimens, and, hence, on the ability of testing techniques to successfully detect drug metabolites in specimens that have been stored for a period of time or at high temperatures. This pilot study was prompted by concerns that excessive heat during shipment or while awaiting testing would cause the drug metabolites in the specimen to degrade, resulting in a false negative test. Although none of the experimental conditions significantly affected the metabolite concentration of the marijuana metabolite, the more extreme experimental conditions (high pH level and high temperature) resulted in a significant decrease in the cocaine metabolite benzoylecgonine (BE) in two specimens. The results suggest that further research regarding the potential deterioration of specimens during storage and shipment is

warranted and that, when testing specimens for the presence of cocaine, care should be taken to ensure that urine specimens are not subjected to excessive heat prior to testing.

Chapter 10, "Performance-Based Assessment Tests," addresses the current state of performance-based testing techniques. It includes discussions of the scope of performance-based tests (what they can and cannot address in an FFD program), the optimal characteristics of performance-based tests, the literature regarding these tests, and characteristics of tests currently available from

vendors. One of the key contributions of this chapter is a summation of the current literature regarding the characteristics of tests designed to measure performance on thirteen classes of processes and abilities that are important for functioning in the workplace. The chapter proposes a technique for assessing the accuracy of a particular performance-based test in measuring alcohol induced impairment. Finally, the chapter presents conclusions regarding the current usefulness of performance-based tests as an element of fitness-for-duty programs.

## Introduction

## Endnotes

1. Nancy Durbin, Ph.D., Battelle Seattle Research Center  
Tom Grant, J.D., Battelle Seattle Research Center
2. Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract DE-AC06-76RLO 1830.
3. 56 FR 41922, August 6, 1991; 57 FR 55443, November 25, 1992; 58 FR 31467, June 3, 1993; and 59 FR 502, January 5, 1994.

# CHAPTER 2: GENERAL REVIEW OF THE TECHNICAL ISSUES CONSIDERED IN THE DEVELOPMENT AND IMPLEMENTATION OF THE NUCLEAR REGULATORY COMMISSION'S FITNESS-FOR-DUTY REGULATION

Charlene Forslund, Nancy Durbin, Idene Field,  
Tom Grant, Michelle Silbernagel, and Ron Wilson<sup>1</sup>

## 2.1: Background and Organization

This chapter reviews the topics contained in NUREG/CR-5227 (Barnes et al., 1988), NUREG/CR-5227, Supplement 1 (Moore et al., 1989), and NUREG/CR-5784 (Durbin et al., 1991). The purpose of this chapter is to briefly summarize the conclusions of the previous NUREG/CRs, to note recent relevant literature, and to discuss any new findings that support previous conclusions or that may suggest different conclusions.

Section 2.2 reviews the prevalence of drug use in American society in general and in the nuclear industry in particular (a topic covered in detail in Chapter 3 of this report). This section also discusses changes in drug use for the five drugs represented in the Substance Abuse and Mental Health Services Administration (SAMHSA) panel (marijuana, cocaine, opiates, amphetamines, and phencyclidine) and alcohol, and presents recent findings about the abuse of other illegal and legal drugs in American society. Section 2.3, summarizing Chapter 4 of this report, reviews potential concerns stemming from current trends in legal drug use. Section 2.4 reviews legal issues concerning drug testing, particularly as those issues relate to job duties with safety implications and to the Americans with Disabilities Act (ADA) that became effective in 1992. Section 2.5 reviews Nuclear Regulatory Commission (NRC) fitness-for-duty (FFD) program issues, such as the NRC's broad-brush approach, random testing rates, medical review officer (MRO) qualifications, employee assistance programs (EAPs), subversion (discussed more fully in Chapter 6), and return-to-work issues (discussed more fully in Chapter 7). Section 2.6 reviews drug testing issues within the NRC's FFD rule. Topics discussed in this section include on-site screening issues, revised Department of Health and Human Services (DHHS) guidelines, reliability of test results, and some alternative testing techniques. Finally, Section 2.7 reviews results from NRC inspections of licensees' FFD programs.

## 2.2: Drug Use Prevalence

Chapter 3 in this report presents a summary of drug use prevalence in the United States, based on the results of such diverse sources as self-report surveys, emergency room data, and workplace drug testing reports. FFD staff will benefit by understanding these national, industry-specific, and emerging drug use trends and by using such knowledge to refine their programs. A brief summary of Chapter 3 is provided below.

While Americans' reported use of both illegal drugs and alcohol has decreased since 1985, drug use still poses considerable FFD problems in the workplace. Findings suggest that drug testing in the nuclear power industry should focus on five specific types of drugs: alcohol, marijuana, cocaine, amphetamines, and pharmaceuticals.

Findings presented in Chapter 3 suggest that alcohol abuse in the workplace poses the most formidable threat to FFD goals. Despite the popular attention paid to illegal drugs, alcohol clearly remains the most commonly used and abused mind-altering drug among the general population, including among employed workers. Survey findings show that while casual alcohol use is declining, heavy alcohol use has remained stable. Drug test results showed alcohol to be the third most common drug detected in nuclear power workers. This is particularly troubling, given that a positive test result for alcohol may indicate on-the-job alcohol abuse.

Marijuana clearly remains the most popular illegal drug (in fact, most illegal drug users reported using marijuana only). Marijuana was the most common drug detected in tests of nuclear power workers.

Cocaine, though declining in use, is still the third most prevalent illegal drug among the general population after marijuana and illegally-used pharmaceutical drugs. Cocaine is the second most common drug identified by testing of nuclear power workers. Cocaine is consistently the most common illegal drug in drug abuse-related emergency room visits. Of Americans aged 18 to 25, a considerable portion (7.7%) reported using cocaine in 1991.

## General Review

While amphetamine abuse is relatively rare among the general population and within the nuclear power industry as a whole, it was the second most common drug detected in the NRC's former Region V (California, Oregon, Washington, Nevada, Arizona, Hawaii, and Alaska), producing over 22 percent of the positive test results among workers covered by the NRC's FFD rule in this region in 1993.

Finally, Chapter 3 finds that the increased use of pharmaceuticals (especially among women), when combined with common complacency regarding their use, poses an FFD concern. Survey data report pharmaceutical abuse as the second most prevalent form of illegal drug use in the United States.

### 2.3: Legal Drugs

Chapter 4 in this report presents a summary of current issues and trends in the use of legal drugs, both prescription and over-the-counter (OTC) (excluding alcohol), as they may affect FFD programs. The potential for safety concerns from the use, misuse, or abuse of legal drugs is clearly present, but because of their legality, these drugs present unique challenges to FFD programs. One challenge is to overcome the tendency to underestimate the potential risks involved. A second challenge is the need to address the special testing issues posed by the lack of accepted testing standards and the presence of various legal mandates that protect employee privacy concerns. Chapter 4 discusses in more detail the types of misuse and abuse that can occur, potential impairing effects of legal drugs (alone or in combination), and trends in the use of legal drugs of the "90s," particularly the new anti-depressants and anabolic steroids. A brief summary is provided below.

The risks posed by legal drugs are of particular concern because:

- prescription and OTC drugs are commonly used,
- the misuse and abuse of these drugs often goes undetected (sometimes even by the user), and
- too little is known about the impairing effects of legal drug use.

Even when properly used, some of these drugs can pose FFD risks either because of their intended effects or their unintended side effects. When misused or abused, the consequences can be even more severe.

Furthermore, changes in the drug approval process by the Food and Drug Administration (FDA) and increases in stress-producing lifestyles are expected to increase the use of OTC drugs.

Previous reports prepared for the NRC identified barbiturates and benzodiazepines as particular safety risks (Moore et al., 1989; Durbin et al., 1991). Those drugs are frequently misused, with serious implications for the workplace. OTC medications with the potential for misuse or abuse include cough and cold medications, analgesics, cathartics (laxatives), menstrual preparations, asthma medications, and camphor.

Recently, there has been a growing concern over two types of legal drugs: the new antidepressants and steroids. A new group of selective serotonin reuptake inhibiting (SSRI) antidepressants have recently become available in the U.S. and are being prescribed with increasing frequency. The new SSRIs appear to have fewer of the side effects that were common among traditional antidepressants, but there has been a growing concern, as yet unsubstantiated, regarding a possible association between some SSRIs and an increased risk of suicide and violence. Another potential concern is that the beneficial effects of these new drugs in treating depression have not been fully studied. Yet it is known that in the absence of effective treatment, depression does cause clear work-related impairments. Further study is therefore warranted on the potential FFD concerns relating to these drugs.

Steroids, in contrast, are not new drugs, but their use has been increasing among the general population. (As yet, there is little evidence of wide-spread use in the workplace.) While anabolic steroids have legal, beneficial medical uses, these substances are controlled as a result of their potential negative effects. Illicit use has tended to involve doses far higher than those that are recommended, with potentially more serious consequences as a result of these higher doses. Potential impairment effects from the use of steroids include both physiological and psychological damage. Because steroid users often view their drug use as being short term, research and employee awareness programs that stress the immediate side effects of anabolic steroids, and the possible long-term consequences of short-term use, may be most effective in curbing steroid use and abuse.

## 2.4: Legal and Regulatory Issues

Several pieces of legislation, enacted or proposed during the last few years, have a potential bearing on fitness-for-duty programs. These include the Americans with Disabilities Act of 1990 (ADA), the Clinical Laboratory Improvements Amendments of 1988 (CLIA), and the Drug Testing Quality Act (HR33) that was introduced in the first session of the 103rd Congress in 1993. Each of these is briefly summarized here.

The Americans with Disabilities Act of 1990 was signed into law on July 26, 1990. The purpose of the ADA is to provide a clear and comprehensive national mandate for the elimination of discrimination against people with disabilities and to provide enforceable standards for addressing such discrimination.

The ADA's fundamental concept is contained in its section 102(a) which prohibits employers from discriminating in regard to all terms, conditions, and privileges of employment against people who the Act includes within the definition of "qualified individual with a disability." The term "disability" is defined to mean (1) a physical or mental impairment that substantially limits one or more of a person's major life activities, or (2) a record of such an impairment. Section 104(a) specifically provides that a "qualified individual with a disability" does not include an employee or applicant who is a current user of illegal drugs although it may include a person who has had a prior drug or alcohol addiction but who has been successfully rehabilitated, provided that the severity of the past addiction was sufficient to result in a "substantial limitation of one or more major life activities." Discrimination on the basis of such a past impairment is prohibited under the ADA.

It appears that the ADA will have no effect on the functioning of licensees' and contractors' fitness-for-duty programs as they apply to current employees. As described above, a current user of drugs or alcohol does not qualify as an individual with a disability and therefore is not protected under the Act. Furthermore, the ADA contains a provision with a specific reference to NRC-mandated fitness-for-duty programs. Section 104(c) states that employers may require employees in sensitive positions, as defined by the NRC's regulations regarding alcohol and drug use, to comply with the standards established by those regulations. The section also provides that employers (1) may prohibit the use of alcohol or illegal drugs at the workplace by all employees, and (2) may require that employees not be under the influence of alcohol or illegal drugs at the

workplace. This appears to allow pre-access, random, for-cause, and follow-up drug testing.

The ADA may, however, affect the conditions under which licensees can deny unescorted access to individuals who have tested positive for drugs or use of alcohol in the past, depending on whether they are considered disabled under the terms of the Act. Section 26.27(a) of 10 CFR requires the licensee to obtain a statement from a suitable applicant for unescorted access regarding past denial of such status and to conduct a suitable inquiry into the applicant's past drug test results. While the ADA prohibits an employer from making any inquiries as to the existence or nature of an applicant's disability prior to an offer of employment, the U.S. Equal Employment Opportunity Commission (EEOC) has ruled that an employer will not be found in violation of the ADA if their actions were taken in response to other federal regulations. Thus the ADA does not supersede NRC regulations. Therefore, it is unlikely that interpretations of the ADA will require changes in the conduct of these background checks. However, to the extent that a licensee has discretion on how they treat those they identify as being disabled, including those with prior drug or alcohol addiction problems, they may be required to make a reasonable accommodation to an otherwise qualified individual.

The Clinical Laboratory Improvements Amendments of 1988 (CLIA) may affect the approximately one third of all NRC licensees that have their own facilities for conducting testing of specimens for the presence of drugs. CLIA's intent was to require DHHS to regulate all laboratories that handle materials derived from the human body for the purpose of providing information for the diagnosis, prevention, or treatment of any disease or impairment of human beings, or the assessment of their health. The most significant requirements that the Act imposes on covered laboratories include:

1. all positive tests must be confirmed with a second test;
2. work must be supervised by an M.D. or D.O.;
3. technicians performing such tests must meet certain personnel requirements;
4. laboratories must pay a fee to become certified;
5. DHHS may perform unannounced inspections;

## General Review

6. laboratories will need to conduct proficiency testing to determine if their procedures are accurate; and
7. other quality control requirements must be followed.

Workplace drug testing was initially included in the rules, but on January 19, 1993 the Secretary for DHHS suspended the application of CLIA to such testing. The Secretary cited concern about balancing the need to ensure accuracy of testing with the need to encourage drug-free workplace programs (58 FR 5220, 1/19/93). Because of the cost to comply with CLIA, it was believed that its implementation would effectively put a halt to all on-site employment drug testing programs and otherwise have a chilling effect on implementation of such programs (*Drugs in the Workplace*, 1993). A final decision is pending. The Secretary for DHHS has said that the agency will be looking to the National Institute on Drug Abuse for guidance and will await findings from a task force looking into possible standards for assuring accuracy and reliability for on-site drug testing (58 FR 5220).

In related legislation, the Drug Testing Quality Act (HR33) was introduced on January 5, 1993 to provide standards for the certification of laboratories performing urine drug testing and to require all drug testing programs to use only certified laboratories. The bill addresses how testing must be done but does not address whether or when testing should be performed. More specifically, this legislation addresses specimen collection procedures to address chain of custody and other concerns, qualifications for labs analyzing the specimens, procedures to ensure accuracy, integrity, and reliability in the results, including the review of positive tests by a qualified professional, guidelines for confidentiality, and the opportunity for persons affected to bring a civil action for damages resulting from errors in the process.

No Congressional action was taken on HR33. The committee to which the legislation was referred, the Health and Environment Subcommittee of the House Energy and Commerce Committee, was aware of the CLIA regulations and informed the Secretary for DHHS that it would like to see CLIA applied to workplace drug testing. The committee was strong in its support for applying legislation to workplace drug testing facilities. In fact, Congressman Dingell, in his introduction of the proposed legislation, used an example from a nuclear utility plant to show why such legislation was needed. If CLIA is applied, it is possible that legislation similar to HR33 could again be introduced. If either piece of legislation is applied to workplace drug testing, NRC

licensees that rely on their own laboratories to analyze specimens may be affected. Therefore, further developments in these areas should be closely monitored.

## 2.5: Fitness-for-Duty Program Issues

The NRC's FFD regulation requires that licensee FFD programs address a number of ongoing concerns that are useful to re-evaluate in light of new information. This section reviews previous NUREG/CR conclusions about these programmatic concerns and updates the conclusions in light of new information. This section includes discussions of comprehensive FFD programs, random testing, deterrence, MRO qualifications and duties, chain-of-custody requirements, subversion techniques, EAPs, and return-to-work issues.

### 2.5.1: Comprehensive Fitness-for-Duty Programs

Executive Order 12564, signed by President Reagan in 1986, ordered the development and adoption of drug testing programs consisting of

- a comprehensive written policy,
- supervisory training,
- employee education,
- an EAP, and
- means to identify users of illegal drugs, including drug testing on a controlled and carefully monitored basis (51 FR 32889).

This is a multidimensional method to assure drug-free workplaces. It recognizes the need to inform employees of drug control policies in a clear manner, to train employees and supervisors in risks of drug use and in recognition of impairment, to assist those with drug-related problems, and to enforce the policy by testing for compliance. This program model reflects a belief that is well-accepted in other areas of worksite health promotion: a comprehensive program of education and activities that address different dimensions of the target behavior will be more effective in promoting behavior change than a program that focuses on one dimension only (Abrams et al., 1986; Glasgow and Terborg, 1988).

The rationale for a comprehensive program draws on several theoretical frameworks of learning and change

(Bandura, 1977; Ajzen and Fishbein, 1980; Prochaska, 1991). Their common component is the recognition that individuals adopt new behaviors (in this case, compliance with an FFD policy) at different speeds, based on a variety of decision-making strategies. To reach and convince all employees in a workforce, a range of information and sanctions may be necessary. Some drug-using employees may wish to comply with the policy, but may need help in addressing their use patterns. Others may be referred by management to forestall incipient problems. One additional important aspect of a comprehensive program: it recognizes that behaviors are influenced by conditions and norms in the workplace, and changes the environment to support the behavior and attitude changes required by the policy (Abrams et al., 1986).

For a more detailed discussion of the effect of comprehensive programs, see Barnes et al. (1988) and Chapter 5, Deterrence, of this report.

### 2.5.2: Random Testing

Random testing is considered a key component of many FFD programs. During the four years following the publication of the NRC's FFD rule on June 7, 1989 (54 FR 24468) and implementation by nuclear power reactor licensees on January 3, 1990, the NRC required licensees to maintain a 100 percent random testing rate. This was required until January 1, 1994, when that rate was reduced to 50 percent.

During the public comment period prior to implementation of the rule, the Nuclear Management and Resources Council (NUMARC) and most licensees proposed that the NRC adopt a 100 percent random test rate. These commenters also recommended that this rate be re-evaluated on the basis of utility experience and be reduced to 25 percent, if warranted (54 FR 24468 at 24472). As a result, the NRC indicated that it would consider reducing testing rates after several years if it obtained information that experience in the industry with the existing rate had been positive (54 FR 24468 at 24474).

In 1991, the NRC directed staff to report on work done to date on the deterrent effect of different testing rates and make recommendations about the applicability of the work to the nuclear industry. The staff found no research that directly addressed the issue of whether reducing the random testing rate affects the deterrent effect of drug testing (see also Chapter 5, Deterrence, in this report).

The NRC considered several alternatives in determining the appropriate random drug testing rate for the nuclear power industry. The NRC considered conducting a study that would reduce the random testing rate of some licensees to 50 percent (experimental sites) and analyze that data against the data of licensees who would continue a 100 percent testing rate (control sites). The experiment would have had to run for several years to allow for delayed effects caused by adjusted testing rates and to obtain a sufficient number of test results. The design of the study and the analysis of the results would have taken an additional year. Due to the relatively long period of time required to collect and analyze the data for such a study, and because variables from site to site could mask any statistical differences between data from two test groups in the small absolute number of expected positive tests, the NRC decided not to conduct such a study (58 FR 15810 at 15811).

Taking into account several factors, including the uncertainties involved in determining the appropriate random drug testing rate for the nuclear power industry, the current low rate of positive tests in that industry, and the unlikely eventuality that the rate of positive random tests for nuclear industry workers will increase, the NRC has concluded that lowering the random testing rate from 100 percent to 50 percent would cause little, if any, decrease in the deterrent effect of random testing (59 FR 502). Therefore, the NRC has modified Section 26.24(a)(2) to permit licensees to randomly test their employees and contractors at an annual rate equal to at least 50 percent. This does not preclude licensees from testing the employee workforce, or portions thereof, at a higher rate (59 FR 502).

### 2.5.3: Deterrence

Chapter 5 reviews the deterrent effect of FFD programs. As discussed in that chapter, a host of factors may affect drug use deterrence in the workforce. These factors include employee attitudes concerning the health, safety, and employment risks of drug use; peer and family influences; and the singular and combined effectiveness of unique FFD program elements such as testing, sanctions for violations, training, education, employee assistance, and organizational culture toward drug use. In general, an effective deterrence strategy should ensure that the risks of drug use are well known, support an organizational culture intolerant of drug use, and provide employee assistance to those seeking to cease their drug use.

## General Review

Deterrence is affected by multiple factors: individual risk perceptions, FFD program elements such as the types of testing being used and employee assistance available, and peer, family, and organizational attitudes toward drug use. These factors support a multifaceted approach to maximize deterrence in FFD programs:

- Effective detection capabilities coupled with strong, consistent, and well-understood sanctions and efficient program administration.
- Education or training on and awareness of the negative consequences of drug use and the effectiveness of the program in detecting drug use.
- An organizational culture that discourages drug use, supported by an EAP that is perceived as being effective, confidential, and non-punitive.

There are limitations to a deterrence strategy. Some occasional users will continue to use drugs regardless of risk. Chronic users cannot be deterred *per se*. Consequently, it is imperative to detect these individuals through an effective testing program.

Several of the studies reviewed in Chapter 5 have concluded that declining drug use detection rates indicate a deterrent effect from random testing, but these inferences are speculative. Detection rate data, often used to document a deterrent effect, provide little meaningful information on deterrence. In fact, there is no empirical data to justify the random testing rates currently used in FFD programs. The lack of empirical research on the deterrent effect of FFD programs in general and random testing in particular is unfortunate, given widespread claims on the importance of random testing's singular deterrent effect. The effect of changes in random testing rates on deterrence in FFD programs is unknown.

Although there is some indication that other FFD program elements serve to deter drug use, little is known about their independent and synergistic effects in many workplace settings where these programs are carried out. It is clear from the review in Chapter 5 that additional information on attitudes toward drug testing programs is needed in order to ascertain the role of risk perception behaviors and other influences in the context of FFD programs. This would permit the design of programs that maximize deterrence of drug use. The deterrent value of FFD programs will remain questionable until these issues are addressed.

FFD program managers and policymakers recognize these issues. Accordingly, judgment and common sense must be the basis for establishing and modifying random testing rates and selection methods. The need for an effective detection and deterrence capability must be balanced against practices that are perceived as overly Draconian, punitive, or unjustified. Likewise, the importance of all FFD program elements working in concert should be recognized and overemphasis on any one element should be avoided. For a more thorough discussion of these deterrence issues, see Chapter 5 in this report.

### 2.5.4: Medical Review Officer Qualifications and Duties

Durbin et al. (1991) reported on the duties and qualifications of MROs, as summarized in NUREG/CR-5784. The NRC's FFD rule, 10 CFR Part 26, requires that an MRO be

A licensed physician . . . who has knowledge of substance abuse disorders and has appropriate medical training to interpret and evaluate an individual's positive test result together with his or her medical history and any other relevant biomedical information.

The primary role of the MRO is to review positive drug test results reported by the laboratory. In this review, the MRO performs a verification and validation function. The verification function consists of verifying that the FFD test results are accurate. A positive urinalysis test result is verified when the MRO determines that the test unequivocally establishes the presence of specific drugs or drug metabolites in the urine at a concentration equal to or above the established cutoff level(s). In addition, the MRO interviews the individual to determine whether or not an alternative legitimate medical explanation for a test result exists (e.g., the test results might be attributable to legitimate uses of medications). Another important function of the MRO is to determine whether or not drug use that can be attributed to legitimate medical uses of the drug(s) in question is also drug use that is likely to lead to on-duty impairment. As the last person to review a test result before actions are taken towards a worker, the MRO serves a critical role in performing the final check for such errors and to validate the meaning of positive drug tests.

DHHS has made several administrative changes that apply to MROs (59 FR 29908). Of the three proposed changes concerning the responsibilities of the MRO, two

proposed changes create new restrictions for MROs, while one change makes the MRO's duties somewhat easier (for more details, see Appendix A). SAMHSA is currently revisiting MRO duties in accordance with DHHS administrative changes; consequently, an up-to-date MRO handbook is expected to be forthcoming from SAMHSA in the near future. Readers are encouraged to review that handbook upon its publication.

### 2.5.5: Chain of Custody

In the context of drug urinalysis, "chain of custody" refers to a set of procedures and activities designed to ensure the authenticity of each urine specimen and to confirm that the condition of the specimen is substantially unchanged from when it was collected from the person being tested. In the commercial nuclear power industry, as in many other industries, chain-of-custody procedures are used to track the handling and storage of urine specimens in drug testing programs. These procedures enable program staff to identify all people who handle specimens from initial collection through handling, shipment, storage, and chemical testing to final disposition at the DHHS-certified drug testing laboratory or, in some cases in the nuclear power industry, at the NRC licensee's drug testing facility.

Chain-of-custody procedures protect the interests of those subject to testing and of their employers responsible for administering the testing. An effective chain of custody protects those being tested by guaranteeing that a specimen being analyzed does in fact belong to the particular person identified on the label affixed to the individual specimen container. Both the person being tested and that person's employer are protected when the chain of custody ensures that the specimen has not been tampered with at any time during the entire testing process. The employer also benefits when the chain of custody provides a strong foundation for the authenticity and unchanged condition of the specimen if the employer is called upon to defend an employment action based upon a drug test result in an arbitration or court proceeding.

Chain-of-custody procedures are applied in three phases in NRC FFD drug and alcohol testing programs. First, NRC licensees use chain-of-custody forms to document the collection, handling, and transportation of urine and blood specimens from one authorized person or place to another as required by sections 2.4(d) and 2.7(a)(2) of the Appendix to 10 CFR Part 26. These forms accompany specimens from the collection site to the licensee's testing facility and/or to the DHHS-certified testing laboratory.

The second phase involves the means of tracking and safeguarding specimen packages sent by courier and express mail companies or the U.S. Postal Service when licensee personnel transfer specimens to DHHS-certified testing laboratories. While not actually considered to be chain-of-custody procedures per se, the internal procedures used by courier and express mail companies and the Postal Service have been shown to be reliable for tracking and maintaining the integrity of parcels in their custody (Evans, 1990). The DHHS-certified testing laboratories' own internal chain-of-custody procedures make up the third phase. Section 2.4(b)(2) of the DHHS Mandatory Guidelines for Federal Workplace Drug Testing Programs requires testing laboratories to use their own chain-of-custody forms to document the integrity of aliquots drawn from specimens while initial and confirmatory testing is conducted (59 FR 29921). The original specimen and the licensee's chain-of-custody forms that had accompanied the specimen are to remain in secure storage during those testing processes.

Durbin et al. (1991) emphasized the importance of reliable chain-of-custody procedures for maintaining the integrity of urine specimens from a number of viewpoints. Since publication of that document, both DHHS and the U.S. Department of Transportation (DOT) have adopted revisions to their respective procedures applicable to drug and alcohol testing. A brief review of these two agencies' revisions pertaining to chain-of-custody requirements can provide useful insight for possible changes in chain-of-custody procedures in NRC licensee FFD programs.

The DHHS issued revisions to its Mandatory Guidelines on June 9, 1994 (59 FR 29908). These revisions include two modifications to chain-of-custody procedures to be used by federal agency drug and alcohol testing programs subject to those guidelines. Like the original guidelines adopted in 1988, the modified DHHS Mandatory Guidelines continue to require federal agencies' collection site personnel to properly execute chain-of-custody standardized forms upon receipt of urine specimens. Thereafter, handling and transportation of specimens from one authorized person or place to another must always be accomplished through chain-of-custody procedures. The original guidelines required collection site personnel to attach chain-of-custody documentation to the exterior of securely sealed containers when shipping specimens to drug testing laboratories. Those original guidelines were silent as to whether courier company or Postal Service personnel who take custody of the specimen packages for transportation had to sign the chain-of-custody documentation during transit.

## General Review

The revised section 2.2(i) of the DHHS Mandatory Guidelines has changed this procedure by requiring chain-of-custody documentation pertaining to particular specimens to be enclosed within each container being used to transport those specimens to the drug testing laboratory (59 FR 29920). Specimen containers are to be securely sealed to eliminate the possibility of tampering during shipment. That section furthermore explicitly states that couriers, express carriers, and Postal Service personnel need not sign the testing agency's chain-of-custody form during transportation of the specimen package. DHHS explained the reason for this modification by stating in section 2.2(i) that: "Specimens are sealed in packages and any tampering with a sealed specimen would be noticed by the laboratory and documented on the specimen chain of custody. In addition, as a practical matter, couriers, express couriers, and Postal Service personnel do not have access to the specimen chain-of-custody form since the form is inside the sealed package."

The U.S. Department of Transportation (DOT) adopted amendments to its Procedures for Transportation Workplace Drug and Alcohol Testing Programs, 49 CFR Part 40, on August 19, 1994 (59 FR 42996). These amendments include revisions to DOT's chain-of-custody procedures similar to those adopted by DHHS. Like DHHS, the DOT requires transportation industry collection site personnel to properly execute chain-of-custody documentation upon receipt of specimens. Thereafter, handling and transportation of specimens from one person to another or from one place to another must always be accomplished using chain-of-custody procedures (49 CFR 40.25(c)). Specimens must be shipped in containers designed to minimize damage and be securely sealed to eliminate the possibility of undetected tampering with the specimens and/or chain-of-custody documents. Those documents are to be shipped inside the containers rather than being attached to the exterior of the containers (49 CFR 40.25(c), (h) and (k)). Because the specimens and documentation are sealed in tamper-evident shipping containers during shipment to the laboratory, courier, express carrier, or Postal Service personnel need not document chain of custody for the shipping container during shipment.

The revised DOT procedures also explicitly state that there need not be a chain-of-custody entry when a specimen which is sealed in such a shipping container is put into or taken out of secure storage at the collection site prior to pickup by courier, express, or postal personnel. Each of the three revised sections emphasizes that the chain of custody has not been broken because

courier, express carrier, or Postal Service personnel have not documented chain of custody for the shipping container during shipment or because the chain of custody does not contain entries related to putting the specimen into or removing it from secure temporary storage at the collection site.

The DOT revised its form for use in establishing drug specimen chain of custody and modified its form used to document alcohol test results. DOT revised its drug specimen chain-of-custody form, known as the Federal Drug Testing Custody and Control Form, in cooperation with DHHS. This form is to be used in testing of federal employees in federal agency drug and alcohol testing programs subject to DHHS guidelines as well as for testing under DOT rules. While the NRC's FFD rule does not mandate it, NRC licensees can use these forms to maintain chain of custody and document alcohol test results in their own FFD programs. (These revised forms were published at 59 FR 43003-43019 on August 19, 1994 and now constitute Appendices A and B, respectively, to 49 CFR Part 40.)

Some of these revisions to DHHS and DOT chain-of-custody procedures were made in response to a December 16, 1993 arbitrator's decision ordering a Birmingham, Alabama bakery to reinstate a truck driver who had tested positive for drugs. The drug testing had been conducted under the DOT regulations in effect at that time. In that decision, the arbitrator was called upon to decide the validity of the chain of custody associated with the positive drug test that subsequently led to the job termination. The arbitrator framed the issue as whether the urine test had met the criteria set forth in the then unamended DOT regulations governing drug and alcohol testing in 49 CFR Part 40. The arbitrator put particular emphasis on the DOT requirement that "every individual in the chain shall be identified" on the chain-of-custody form (49 CFR 40.26(k)) and that "Laboratories shall use chain-of-custody procedures ... during storage ...." (49 CFR 40.29(a)(2)). The arbitrator interpreted these regulations to require that *all* people who handle the specimen from time of collection to final disposition be identified on the chain-of-custody form. According to the arbitrator this included the person who placed the specimen in temporary storage at the collection site before sending it on to the testing laboratory and all courier personnel who handled the specimen during transportation to the laboratory. The chain-of-custody form accompanying the specimen contained no reference to storage nor did it identify any courier employee. Because these people were not identified on the chain-of-custody form, the arbitrator found that the employer had

not followed DOT's chain-of-custody procedures. He, therefore, ordered the employee's job to be reinstated. The arbitrator's decision was later upheld by a U.S. District Court (Interstate Brands, 1994a).

In December 1994, the U.S. Court of Appeals, Eleventh Circuit reversed the District Court's decision and voided the arbitrator's decision (Interstate Brands, 1994b). The Court agreed with the DOT's interpretation of its own regulation that couriers, postal employees, and other intervening personnel involved in the transportation of urine samples need not sign chain-of-custody forms. People involved in transporting sealed shipping boxes containing specimens are not considered to be in the chain of custody and, therefore, need not sign such documentation. The Court acknowledged that tamper-evident seals on the specimen bottle and shipping box are sufficient guarantors of specimen integrity during transportation. Likewise, the Court found DOT's lack of a requirement that a person putting a specimen into secure temporary storage note that action on the specimen's chain-of-custody form to be consistent with long-standing legal precedent.

The DOT requirements, as assented to by the Court of Appeals, are consistent with normal and acceptable chain-of-custody practice in federal workplace drug testing programs and in forensic science practice generally. It is normal practice in drug testing programs and forensic science to not require that all transportation personnel who have custody of specimens sign chain-of-custody forms accompanying those specimens. The internal shipping documentation and tracking systems of courier companies and the U.S. Postal Service have been found over many years to be reliable for the purpose of identifying those who handle and transport urine and blood specimens. It would appear, therefore, that NRC licensee programs need not require each courier or postal employee who handles specimen shipping containers to be identified.

Another reason for not requiring such identification is that it would mean forgoing the benefits of having chain-of-custody documentation enclosed in the sealed shipping containers. As the DOT noted when adopting its August 1994 revisions to its regulations, "... the Department recommends enclosing chain-of-custody documentation in the shipping container, as opposed to attaching it to the exterior of the container, since this minimizes the likelihood of loss or damage to the documents." (59 FR 42997).

Requiring a notation on the chain-of-custody form when a specimen is put into temporary storage at a testing facility would also seem to go beyond the realistic needs of tracking the whereabouts of specimens in the context of workplace drug testing. Courts reviewing the admissibility of evidence protected by chain-of-custody procedures have long held that such procedures need not exclude absolutely all possibility of tampering with the evidence (see e.g., *Kansas v. Dodd*). For an object to be admissible in a law suit, it suffices that the chain of custody establishes that it is more probable than not that the object is the one connected with the case (*Louisiana v. Kinchen*). This is particularly true in workplace drug testing which takes place in a non-criminal, administrative context. While stringent evidence tracing requirements may arguably be appropriate in criminal prosecutions, some courts have suggested that such requirements can be relaxed in civil cases (*Woolley v. Hafner's Wagon Wheel*).

The arbitrator's decision, and the DHHS, DOT, and Federal Circuit Court responses to that decision, indicate the importance of having a clear and complete statement of chain-of-custody procedures in federal regulations governing drug and alcohol testing programs. By making the revisions discussed above, DOT and DHHS have provided such a statement for their regulated organizations. In so doing these agencies have also conformed their regulations to widely accepted chain-of-custody practices in workplace substance abuse programs and the forensic sciences. The revised DHHS and DOT requirements and the Federal Court decision should provide NRC licensee FFD program administrators with useful guidance for maintaining strong internal chain-of-custody procedures that comply with the NRC's Part 26 requirements.

### 2.5.6: Subversion

Testing employees to determine whether they use illegal drugs and alcohol is central to many FFD programs. Program effectiveness hinges on valid, accurate results and fairness. Effectiveness is lost if some drug users evade detection and compromise the integrity of the FFD program. When subversion of FFD testing programs occurs, the program fails in its mission to identify and remove drug abuse and its consequent effects. In addition, allowing subversion implies to workers that the program is not taken seriously by management and undermines support for the program.

In addition to damaging the effectiveness of a single program, subverting the testing process can also

## General Review

compromise the credibility of industry-wide program performance. FFD program performance data, and data from other drug testing programs such as the military (Bray et al., 1992) and transportation industry (Department of Transportation, 1993), have shown continuing declines in rates of positive results from drug testing from the 1980s until 1992, when the declining trend in drug detection leveled off. Proponents of drug testing suggest that declining drug detection rates over the last several years indicate that program policies and procedures are effective. However, declining drug detection rates may be due in part to drug users successfully eluding detection. It is impossible to know how many drug users in these testing populations have successfully subverted the testing process, thereby escaping detection. A more detailed discussion of subversion and countersubversion is provided in Chapter 6 of this report.

### 2.5.7: Employee Assistance Programs

Employees Assistance Programs (EAPs) are a required element of the NRC's FFD program. Based on a comprehensive review of EAPs in NUREG/CR-5227, Barnes et al. (1988) concluded that, although EAPs could not be the sole mechanism for assuring that employees are fit for duty, EAPs play a vital role in assuring fitness for duty in the nuclear power industry. The review provided in NUREG/CR-5227 includes a discussion of the elements of an ideal EAP, the scope of EAP services, confidentiality of EAP services, funding options for EAPs, and evaluation of effectiveness of EAPs.

In the years since publication of Barnes et al. (1988), several EAP issues addressed in that publication continue to be debated. These issues include the scope of EAP services, the need for confidentiality for employees who use EAP services, and the need for evaluation of EAP services. The most important "new" issue facing the EAP industry today is the effect managed care initiatives will have on EAP services. Updates of each of these issues are provided in the sections below.

#### 2.5.7.1: EAP Scope of Services

The debate about whether EAPs should offer treatment and referral services beyond problems with alcohol or drug abuse has broadened since Barnes et al. first summarized this issue in 1988. Hirshberg (1993) suggests that with so much change and uncertainty occurring in the workplace today, EAPs should expand into areas formerly the domain of human resource professionals. Such areas include educating organizations

and their employees to recognize and manage change processes; assessing and addressing the impacts of downsizing and mergers on workers and organizations; increasing services for everyday work/family issues, such as child care and elder care; offering information on topics such as diversity; assisting employees with career counseling; administering wellness or health promotion programs; and helping organizations manage health care benefits for mental health, drug, and alcohol services. Although Hirshberg seems to be advocating extreme broadening of EAP services, other researchers see a need for EAPs, at the very least, to either integrate separate efforts for, or assume responsibility for, workplace wellness programs (also called health promotion programs, or HPPs) (Cook and Youngblood, 1990; Washington State Drug Free Business, 1992a; Schechter, 1992; Ames, 1993).

Traditionally, wellness programs have been much more prevention-based than EAPs. And although some of the goals of wellness programs overlap with those of EAPs, the goals of wellness programs are more general:

- to promote the physical and emotional health of employees and their families;
- to resolve disturbances in the physical and emotional health of employees at the earliest stage of problem development; and
- to prevent the return of physical and emotional health problems once the employee has been treated (Schechter, 1992, p. 38).

Specifically, wellness programs most often focus on helping employees with stress management, time management, and weight control. These programs also provide screenings for high blood pressure, high cholesterol, and other health risks related to lifestyle choices (e.g., cigarette smoking and lack of exercise) (Washington State Drug Free Business, 1992a; Schechter, 1992). Cook and Youngblood (1990) advocate combining the substance abuse prevention services of EAPs with other health promotion efforts offered by workplace wellness programs. According to Cook and Youngblood, if EAPs and workplace wellness programs were to combine their efforts, EAP staff would contribute valuable knowledge about the dynamics of substance abuse—expertise that wellness program staff often do not have. Schechter (1992) states that letting EAPs evolve into "Employee and Health Assistance Programs (EHAP) would help reach employees who do not normally seek assistance until later stages of their difficulties" (p. 38).

Support for Schecter's statement can be found in evidence from EAP studies. For instance, in one study, EAP counselors "assessed more emotional, alcohol, or other drug, and physical or sexual abuse problems than clients had expressed as their main problem. Conversely, staff assessed fewer occupational, financial, educational, and job stress problems than were presented by clients. The largest discrepancies were observed for clients deemed to have alcohol or drug problems, of whom 56 percent had first indicated a different problem" (Tompkins, 1991, p. 89). Tompkins concludes that substance-abusing workers may enter into treatment earlier if they can establish contact with EAP staff initially through a "safer" complaint.

Given findings from EAP industry trends and studies, there are reasons to review and evaluate the scope of services currently offered by EAPs. EAPs that offer not only assessment and treatment or referral services for drug and alcohol problems but also broader-based health promotion services, such as those traditionally provided by a workplace wellness program, can increase the overall emphasis on employees having a healthy lifestyle in general. At the very least, EAPs that undertake some of the workplace wellness activities, such as screenings for blood pressure and cholesterol or classes in smoking cessation, make EAP staff members more visible and accessible to workers, thus perhaps increasing the chances that a worker will seek out the help of the EAP for personal problems, before those problems start affecting a worker's job performance.

Regardless of precisely what array of services EAPs offer, a more important ongoing debate concerns the likelihood that employees will use EAP services and the extent to which EAP services solve employee problems and improve the health and safety of workers. A few external evaluations of EAPs have provided some insight into what aspects of EAPs increase the likelihood that workers will use their services. For example, some researchers have found that EAPs that are more in tune with the "organizational culture" in which they are operating or with the values and needs of particular types of employees in the workforce, have an increased likelihood that employees will use their services when problems arise (Ames, 1993; Drake, 1993).

#### 2.5.7.2: Confidentiality

Studies have shown that several factors contribute to the likelihood that employees will use EAP services. At the top of this list of factors was confidentiality. According to Hall et al. (1991), "Employees were more likely to use

EAP services if they perceived that (a) confidentiality existed; (b) the EAP was helpful, affordable, convenient, and of assistance to individuals in keeping their jobs; and (c) their supervisors viewed the EAP as helpful" (p. 72). Other researchers report the importance of confidentiality in delivering effective EAP services and the need for organizations to state in writing to EAP staff and employees alike the circumstances under which information given to the EAP need not be held confidential (Seymour and Smith, 1990; Osterloh and Becker, 1990; Bickerton, 1990; Major, 1990; Petracek, 1992, *The Seattle Times*, July 15, 1994). Although many EAPs must comply with the DHHS final rule regarding confidentiality of alcohol and drug abuse patient records (52 FR 21796), according to Seymour and Smith (1990), federal confidentiality protections do not apply in the following circumstances:

- reporting under state law of incidents of suspected child abuse and neglect;
- communications within an EAP program or between a program and an entity having direct administrative control over an EAP program;
- communications between an EAP program and a qualified service organization; and
- disclosures to law enforcement officers concerning a patient's commission of or threat to commit a crime at the EAP site or against personnel of the EAP (p. 404).

In summary, Eickerton (1990) states that "the EAP that manages to maintain most patient/client information below the level of disclosure has the greatest chance of being the most successful program" (p. 44).

The NRC's FFD rule encourages licensees to protect the confidentiality of EAP services provided to employees unless the employees' behaviors or conditions pose a threat to themselves or others. The NRC has proposed to strengthen this language to require confidentiality of EAP services (58 FR 15810).

#### 2.5.7.3: Evaluation Efforts

Ironically, the assurance of confidentiality in delivering EAP services seems to be one reason for the lack of further research data on the effectiveness of EAP services. As Barnes et al. reported in 1988, EAPs seldom self-evaluate, and organizations seldom require such evaluations. Since that time, more concern has been

## General Review

expressed about the need for EAP evaluations, but there is still little empirical evidence that such evaluations are taking place. According to Tompkins (1991), "There is a general lack of detailed recordkeeping among EAPs due to the primacy of confidentiality and the absence of any industry compliance standards" (p. 88). Researchers continue to attribute this lack of evaluation to the difficulty of choosing appropriate outcome measures by which to evaluate EAP effectiveness, given the complexity and diversity of the employee problems EAPs are tasked with solving (Tompkins, 1991). A recent report by the Committee on Drug Use in the Workplace (Normand, et al., 1994) concluded that it is misguided to ask whether the (generic) EAP is effective. Different EAP designs are necessary "to accommodate differences in organizational structures, employee demography, methods of dealing with good or poor performance, benefit packages, supervisor training and authority, and the organizational experience with different types of human resource innovations" (p. 257). Therefore, we should instead seek to understand how EAPs contribute to a range of different outcomes in different settings. The focus should be on particular EAP techniques and services rather than on the presence or absence of an EAP.

Definitive studies of this type have yet to be conducted. However, a few studies completed to date suggest that (1) constructive confrontation and referral by supervisors may improve outcomes, (2) long-term follow-up support by EAP staff may improve outcomes, and (3) initial referral of alcoholic employees to intensive inpatient treatment followed by Alcoholics Anonymous (AA) attendance produces better outcomes than referral to AA without other treatment (Normand et al., 1994, p. 257).

The report concluded that more high-quality critical case studies of EAPs need to be conducted before theoretical understandings can be developed to permit generalized statements about what is likely to work in particular settings (Normand et al., 1994). If EAPs are increasingly integrated with managed care initiatives in the future, they will likely be required to devote more attention to documenting and reporting the effectiveness of their services in the workplace.

### 2.5.7.4: Managed Care

As the federal government starts to take a serious look at nationwide health care reform initiatives, cost-containment of health care benefits and services through managed care is the foremost issue in the EAP industry. Many professionals in the EAP industry see managed care as a

threat to providing employees with adequate, quality care for their substance abuse and mental health problems. Traditionally, EAPs have focused on improving access to and increasing employee use of EAP services; managed care is centered around avoiding "unnecessary" costs by scrutinizing use patterns and reshaping reimbursement and the delivery of services (Tompkins, 1991). Nonetheless, some employers see EAPs as good avenues for controlling costs through managed care because "EAPs combine face-to-face intake of potential clients with the capacity to select and negotiate with providers and to monitor outcomes" (Tompkins, 1991, p. 86). Some researchers and EAP professionals argue that if EAPs are required to conform to managed care standards, there will be a shift from employees seeing EAPs as confidential and helpful resources to seeing EAPs as management enforcers or obstacles to treatment, thereby diminishing the effectiveness of EAPs in the workplace (Tompkins, 1991; Neuhaus and Caplan, 1992; Bennett, 1993).

Other researchers and EAP professionals are more hopeful. These advocates of managed care see a positive role for EAPs in cutting the costs of substance abuse treatment. They argue that EAPs can cut costs through being more proactive in using early intervention (when probability of success is greater) to help employees before their problems become treatable only on a long-term, inpatient basis (when there is a lower probability of successful treatment); the strength of EAP staff lie in their ability to provide accurate assessments and appropriate referrals, thereby finding the most cost-effective way to treat employee problems (Washington State Drug Free Business, 1992b). These researchers and EAP professionals also stress that employers can manage cost and provide quality care to employees if they use EAPs as case managers for mental health/chemical dependency benefits and if they set "appropriate care" as the goal for serving troubled employees. One professional again stressed the need for EAPs to be more proactive in providing the following services:

- unlimited short-term counseling at no pocket cost to employees;
- patient-oriented, not protocol-centered, utilization review and case management; and
- a benefit plan designed to encourage outpatient treatment, the use of the EAP as the gatekeeper, and a limit on per diem charges (Washington State Drug Free Business, 1992b, p. 5).

Another professional provides more general, less tangible suggestions that organizations can use to achieve cost-effective, quality care for employees. He states that the integration of EAP services and managed care can be achieved by following these steps:

- use the EAP or other independent professional caregiver as a gatekeeper;
- provide face-to-face assessment;
- have quality measured by other than the provider; and
- measure outcomes (Stetzer, 1993, p. 39).

Regardless of how employers choose to use EAPs in a managed care setting, employers and managed care EAPs should be aware of possible liabilities they may encounter in providing services. Gerner and Kerns (1994) warn that there are four potential areas of liability for managed care: 1) selection of providers, 2) utilization review, 3) denial of claims, and 4) payment of financial incentives. To avoid liability involving the selection of a provider, such as might occur if an employee blames not only a provider but also the provider's EAP and the contracting employer for inappropriate or harmful treatment received, the employer should clearly state in every way possible that there is no relationship between the employer and the EAP or the treatment provider (*Drugs in the Workplace*, April 1994; Gerner and Kerns, 1994). Similarly, if EAPs are involved with a utilization review company, the EAP should carefully check into the utilization review company's policies for claim denial and protocols for admission, continued stay, emergency care, and discharge. These protocols should be based on standards of care rather than dollar amounts. In this way, EAPs can avoid liabilities associated with utilization review and denial of claims (*Drugs in the Workplace*, April 1994; Gerner and Kerns, 1994). Finally, although the courts to date have sided with managed care providers and legislative bodies that have approved the use of financial incentives, risk-sharing, and utilization review to contain health care costs, EAPs that are part of a managed care program should be aware that claims for liability based on financial incentives not to treat or on referral restrictions have been and continue to be brought against managed care entities (*Drugs in the Workplace*, April 1994; Gerner and Kerns, 1994).

To date, researchers and EAP professionals seem to agree that managed care will be integrated into the delivery of EAP services in the near future, but there is still much

debate about how that integration will be implemented and the effect that integration will have on employees receiving adequate, appropriate resources to overcome problems that may affect their performance at work. An important issue is the extent to which cost-containment through managed care initiatives may affect the ability of EAPs to provide adequate, appropriate services to nuclear power plant employees under the FFD rule.

### 2.5.8: Return to Work

Following a confirmed violation of the worksite policy on substance abuse, several organizational responses are possible. Some employers choose to terminate the offending employee. Others may remove the offending employee from the job and then provide the opportunity for reinstatement. Employers providing the opportunity for reinstatement are faced with major decisions regarding treatment, rehabilitation, and return to work of the affected employee. There is little in the way of professional consensus to guide those who must make these decisions.

Chapter 7 in this report reviews the literature on types of drug and alcohol users, treatment methods, and treatment outcomes, as well as the scant literature on return-to-work issues, and discusses the limitations of current research to address these issues. Chapter 7 also discusses the important elements of an assessment of fitness to return to work. The results of this review suggest the following:

- Satisfactory compliance with the chosen treatment program provides the minimum assurance of a worker's fitness to return to work. While the selection of the best treatment program for the employee is based on professional judgment rather than professional consensus, the importance of completing the agreed upon treatment is well established in the outcome literature. Those patients who complete primary treatment and who follow the aftercare plan are consistently those who do best in long-term follow-up studies. Compliance with an agreed upon treatment program and aftercare plan is an important element of fitness to return to duty and might be regarded as a measure of employee reliability.
- Most drug and alcohol treatment outcome studies focus on the highly dependent drug or alcohol abuser, typically the "street addicts." Therefore, the relevance of these findings to the typical worker and to the decision to permit workplace re-entry following treatment is limited.

## General Review

- Patient factors, including motivation, chronicity and severity of drug or alcohol abuse, age and gender, and personal and social resources, appear to be better predictors of successful outcomes than treatment program factors. Because employed workers are more likely to have patient characteristics associated with successful treatment than patients who have been unable to hold a job due to their drug or alcohol use (e.g., employed workers are expected to have greater personal and social resources and to be highly motivated to abstain in order to maintain employment), treatment success for employed workers who have been identified through an FFD program is expected to be high. One expert familiar with the nuclear industry has found that a high-paying job does provide a strong motivation to comply with treatment.
- Appropriate treatment for employed workers who test positive for illegal drugs is not well studied, although, as noted above, research on factors that predict successful treatment outcomes suggests that these individuals would have better outcomes than are typically reported in the treatment literature. Developments in patient and treatment matching warrant attention over the next few years as researchers and practitioners seek a better understanding of the relationship between patients, treatment methods, and successful outcomes.
- In light of the difficulty in defining suitable treatment and appropriate treatment outcomes for the violator of the worksite policy on substance abuse, many employers are instead choosing to focus efforts on developing procedures that must be followed in treatment, rehabilitation, and the decision to permit re-entry into the workplace. The procedures specify the decision-makers and address some of the criteria, but ultimately leave the determination of fitness to professional judgment.
- Treatment and aftercare plans should be developed with explicit goals that correspond to workplace expectations: abstinence from illegal drugs and abstinence or controlled use of alcohol and legal drugs in a manner that does not interfere with productivity and safety.
- The EAP is in a position to provide the ongoing monitoring of fitness upon return to work that is necessary to assure compliance with treatment and aftercare goals and to provide worksite support and encouragement for long-term recovery.
- Discussions of employee reliability and trustworthiness do not appear in the treatment literature because they are not concepts that fit within the prevailing illness model of drug and alcohol dependency.

The discussion of return to work issues in Chapter 7 is intended to assist the reader in understanding the knowledge base upon which treatment and return to work decisions are made. It is intended to assist the reader in understanding the currently available literature in this arena and to reinforce the need for professional judgment in selecting treatment and determining fitness to return to work.

## 2.6: Drug Testing Issues

This section reviews drug testing issues within the NRC's fitness-for-duty rule. Topics discussed in this section include new Department of Health and Human Services guidelines, on-site screening issues, reliability of test results, and different testing techniques.

### 2.6.1: Department of Health and Human Services Changes

SAMHSA, under DHHS, published changes to its mandatory guidelines for federal workplace drug testing programs in the June 9, 1994 edition of the *Federal Register* (59 FR 29908). The changes became effective September 1, 1994. These changes create some new restrictions, but overall these changes lessen the effects of existing regulations, making drug testing regulations less costly and easier to enforce for federal agencies and DHHS-certified laboratories. In addition to several general definition changes and several clarifications regarding the role of the Secretary of DHHS, DHHS changes primarily affect three components of drug testing programs: specimen collection, MROs, and DHHS-certified laboratories. Appendix A presents a brief summary of the changes regarding DHHS mandatory guidelines for federal workplace drug testing programs.

### 2.6.2: On-Site Screening Issues

Although most drug testing programs, including the NRC's FFD program, require that testing be completed at a DHHS-certified laboratory, some companies have chosen to perform immunoassay screening tests at the work site. Performing screening tests using immunoassay techniques on site can reduce costs by reducing the number of specimens that are sent to a DHHS-certified

laboratory for testing. In addition, such testing can reduce the time until determination of a negative result. On-site screening of urine specimens is an established practice at many worksites, including nuclear power plants. Although proponents of on-site screening of specimens believe that the screening is a desirable component of a drug testing program, others have raised concerns about the procedures used in such a program, whether they are adequate to assure accurate results and the protection of employee rights. This section summarizes a report by Doug Rollins, M.D., Ph.D., on the results of a recent study that addressed these concerns.

The National Institute on Drug Abuse (NIDA)<sup>2</sup> formed a committee to gather data on a limited number of on-site testing facilities. Dr. Rollins led this committee. Eleven sites were surveyed, including military, criminal justice system, construction, power generation, transportation, and off-shore drilling sites. Information was gathered from site visits and questionnaires (Rollins, 1993). Rollins (1993) reports that the committee concluded that "all sites were attempting to perform on-site testing in a reasonable manner" and that "no facility was performing drug testing in an overtly inaccurate or false manner" (Rollins, 1993, p. 1). However, several of the facilities had serious flaws in their procedures, some of which could lead to errors in testing or handling of specimens. Examples of flaws included poorly trained personnel, weak security, and incomplete quality control and quality assurance. The committee developed three recommendations, reproduced below, based on the findings of the study.

#### **Recommendations of the Committee for the Evaluation of On-Site Drug Testing**

1. Guidelines should be developed that establish operational consistency and analytical accuracy among on-site urine drug testing facilities. Such guidelines should include, but need not be limited to:

- a. criteria for education, training, demonstration, and documentation of skills of on-site laboratory personnel performing the appropriate test procedures.
- b. criteria for the standardization of chain-of-custody and control procedures for specimens and aliquots of specimens.

- c. criteria for the assurance of the security and integrity of specimens and the security of data.
- d. criteria for the development of a complete and updated standard operating procedure manual.
- e. criteria for the standardization of the drug tests, including cutoff concentrations, blanks, standards and controls to be tested at or near the cutoff concentrations, testing methods to be used, and procedures for the validation of methods.
- f. criteria for the establishment of quality control and quality assurance programs to cover every step in the drug testing process.

- 2. Laboratories involved with the criminal justice system *only* should not necessarily be included in consideration of the development of the above guidelines, as their purpose for testing is not similar to workplace testing. However, the consequences of drug testing in the criminal justice system are no less serious than in the workplace, and inaccurate or unsubstantiated results should not be permitted. Thus, a set of criteria similar to the above guidelines must also be developed for drug testing within the criminal justice system. This seems, however, to be beyond the scope and mission of this Committee.
- 3. A system should be developed for monitoring on-site drug testing facilities to assure compliance with the above guidelines.

#### **2.6.3: Opiate Testing Requirements**

Two issues have arisen regarding testing requirements for opiates:

- Requirements for 6-AM testing discussed in 2.6.3.1, and
- Opiate screening test level requirements, discussed in 2.6.3.2

##### **2.6.3.1: 6-AM Requirements**

The fitness-for-duty rule currently requires NRC licensees to test for opiates and conduct a 6-acetylmorphine (6-AM, formerly described as 6-monoacetylmorphine or 6-MAM)

## General Review

test for heroin. This section provides an update on the usefulness of the 6-AM test for heroin.

All opiates are metabolized by the liver and excreted in urine as the metabolite morphine glucuronide. Morphine is excreted in urine after consumption of codeine, heroin, poppy seeds, opium, or morphine. The source of morphine cannot be determined through urinalysis (Paul et al., 1988). Although there are a number of common opioids that can be prescribed by physicians, there are no legal circumstances under which heroin can be prescribed. Therefore, a specific test that could differentiate heroin from the other opioids is desirable.

Numerous studies including Mulè and Casella (1988) and Cone and Welch (1989) have demonstrated that the by-product 6-AM will only be found in the urine after heroin consumption. This compound is not found in the urine after morphine, codeine, or poppy seed consumption. Although 6-AM can identify a heroin user, the 6-AM test may not be practical or cost-effective in the nuclear power industry due to the very short duration of the presence of 6-AM in the urine after heroin consumption and the very low incidence of heroin use in the industry. Fehn and Megges (1985) found that the half-life of 6-AM in whole blood is only 38 minutes and not dose related. Cone and Welch (1989), in a controlled study, found that levels of 6-AM in the urine could not be detected by urinalysis as few as 8 hours after test subjects injected heroin. They state that "Overall, it appears that there is a very limited time window for detection of 6-AM after heroin use" (p. 449). However, as pointed out in the *MRO Alert*, (October/November, 1995), as detection testing methods become more sensitive at lower levels, drugs (6-AM) can be detected for longer periods of time, regardless of the half-life time of elimination (Quadrangle Research, LLC, *MRO Alert*. Vol. VI, No. 8, p. 3).

### 2.6.3.2: Opiate Screening Level

The DHHS, through SAMHSA, has proposed to change the urine opiate drug screening cutoff level used presently in opiate drug detection tests from 300 ng/mL to 2000 ng/mL. The proposed rule change to the Mandatory Guidelines for Federal Workplace Drug Testing Programs was published in the *Federal Register*, Vol. 60, No. 221 on November 16, 1995. The current 300 ng/mL cutoff level was adopted when the Federal Workplace Drug Testing Program was established, in accordance with testing levels for opiates used by the Department of Defense (DOD) for testing service members. These levels were selected in an effort to provide the greatest opportunity to identify anyone who may have used

heroin. However, industry experience indicates that at these levels, many individuals who have taken prescriptive medication such as codeine or morphine or who have consumed normal dietary amounts of poppy seeds have also tested positive. Recently, DHHS evaluated 7,294 specimens (out of 1.1 million urine specimens) that tested positive for opiates in five SAMHSA certified laboratories. Three different MRO groups evaluated the specimen results from January of 1992 through March of 1993. Of these, 87% were verified as negative by the MRO based on consumption of prescriptive and OTC drugs, or poppy seeds in food ingested, or the absence of clinical evidence of heroin use. Of the positive specimens, 81% (5,931/7,294) had less than 2,000 ng/mL of codeine or morphine. Furthermore, 84% of the 7,292 positives were tested for 6-AM. Only 16 specimens tested positive and 14 of these 16 had morphine in concentrations greater than 2,000 ng/mL (Quadrangle Research, LLC, 1995).

These results may indicate that the current standard of 300 ng/mL is not accurately identifying opiate drug abusers and that raising the screening cutoff level to 2000 ng/mL in conjunction with the required 6-AM testing will provide adequate protection from and deterrence of illicit opiate use. The DOD adopted similar increases in the testing cutoff levels as of April 1, 1994 in response to similar concerns. The proposed rule change by DHHS should result in a reduced cost to federal agencies, laboratories, MROs, and participating companies since fewer specimens will be screened positive, resulting in a reduction in the number of specimens sent to confirmatory testing (Quadrangle Research, LLC, 1995). In spite of the potential cost savings and the reduced likelihood of inaccurately identifying positive opiate drug users, there is some concern that the higher cutoff level may result in false negative test results.

### 2.6.4: Reliability of the Testing Process

Because 10 CFR Part 26 requires NRC licensees to use DHHS-certified laboratories to perform the analytical testing for the licensees' FFD programs, SAMHSA, within the DHHS, periodically analyzes data pertaining to the nature of unsatisfactory testing results that have been reported to the NRC. This analysis is part of the NRC's ongoing FFD performance monitoring program, which was designed to identify unforeseen problems in the drug testing process of NRC FFD programs and to initiate corrective actions when problems are identified. An important goal of the monitoring program is to protect nuclear industry workers under NRC jurisdiction from unfair adverse effects of false testing results. This section

summarizes the findings from a SAMHSA analysis of unsatisfactory testing results that have occurred in the nuclear industry. This analysis was compiled and submitted to the NRC by Dr. Michael R. Baylor and Dr. Donna M. Bush of the Division of Workplace Programs within SAMHSA. A complete copy of this analysis may be found in Appendix D: Analysis of Testing Process Errors, of NUREG/CR-5758, Vol. 3, *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports CY 1992* (Fleming et al., 1993).

The SAMHSA report describes the unsatisfactory testing results that occurred in the nuclear industry between January 3, 1990 and December 31, 1992. During this 36-month period, the 52 NRC licensees tested approximately 807,400 urine specimens as part of their FFD programs. Of those urine specimens tested, 44 licensees reported that there were unsatisfactory testing results on 199 test specimens. These included 188 double-blind performance specimens and 11 specimens that were provided by licensee or contractor personnel. However, all of the unsatisfactory test results were resolved and no NRC employees or contractors suffered damaging consequences from the unsatisfactory test results. Table 2.1 summarizes the unsatisfactory testing results by type and year. Data from calendar years 1993 and 1994 indicate that there were 25 unsatisfactory results in 1993 and 23 in 1994. Almost all unsatisfactory results were based on blind performance test samples.

Overall, the number of unsatisfactory tests declined 86 percent from January 3, 1990 to December 31, 1992. According to the SAMHSA report, the decrease from 99 unsatisfactory testing results in 1990 to 14 unsatisfactory testing results in 1992 "appears to be significant"

(Fleming et al., 1993, p. D-3). The types of factors contributing to unsatisfactory testing results were categorized by the report authors into four general areas, which are defined as follows: *False-negative test results* refer to a specimen that is reported to be negative even though the actual concentration of drug in the specimen is above the stated cutoff level. *False-positive test results* are specimens that do not contain any drugs that either test positive for drugs (analytical false positive) or that are reported to be positive for drugs (administrative false positive). *Other-improper manufacture of blind performance specimens* involve general problems in the drug testing process that are linked to the improper manufacture, formulation, or packaging of the blind quality control specimens. And *other-improper processing of specimens* pertains to the improper processing of urine specimens on site prior to their shipment to the laboratory for testing, or inappropriate handling or actions by the MRO.

The significant majority of the unsatisfactory testing results (58.3%), involved either the improper manufacture of blind performance specimens (95 specimens) or the improper processing or handling of specimens (21 specimens). Problems with these results included specimens with inconsistent immunoassay performance due to concentrations of drugs that were too close to the cutoff levels and specimens that were improperly labeled as negative when in fact they contained drugs above specified cutoff levels. Still other specimens were used after their shelf-life "sell-by" dates had expired, allowing the drugs in the specimens to degrade.

The second most common factor producing unsatisfactory testing results was false negative laboratory results. Seventy-eight specimens (or 39.2%) were deemed false

**Table 2.1: Summary of Unsatisfactory Testing Results (January 3, 1990 through December 31, 1992)**

Year	Personnel Specimens				Blind Performance Specimens				Total	
	False Negative	False Positive	Other		False Negative	False Positive	Other			
			Manufacturing	Processing			Manufacturing	Processing		
1990	3	1	0	1*	39	1	41	13	99	
1991	1	1	0	2	27	2	50	3	86	
1992	0	0	0	2	8	0	4	0	14	
Total	4	2	0	5	74	3	95	16	199	

\*No immunoassay testing directed by MRO. True positive THC, administrative false positive benzodiazepene.  
Source: Fleming et al., 1993, p. D-7.

## General Review

negative laboratory results. Seventy-four of these results involved quality control tests with analytical problems (primarily tests that screened below cutoff levels) or administrative errors. Administrative errors included errors in entering data into computers, errors in clerical transcriptions, the use of higher cutoff levels, and test specimens being misidentified. Of the four personnel tests with false negative results, all involved administrative errors. Three of these were errors in clerical transcription, while one result involved two specimens being accidentally switched.

Although a total of five administrative false positives were reported for 1990 through 1991, no false positives were reported during 1992. Subsequent to the SAMHSA report, one licensee reported to the NRC in October, 1994 one false positive. A sample was analyzed by a special testing protocol of a suspect specimen (the specimen had a low creatinine level) that extended beyond DHHS-approved methods and practices and NRC FFD testing requirements. The results, which were at first thought to be positive, were determined to be falsely positive. The false positive result was quickly resolved by the licensee and the testing laboratory, and the licensee test subject suffered no adverse impact as a result of the false positive (L. Bush, Nuclear Regulatory Commission, personal communication, October 15, 1993). Another licensee reported to the NRC in December, 1994 a false-positive result that involved a unique combination of prescription drugs, OTC drugs, and habitual consumption of many cans of soda daily. The individual suffered no adverse impact from this false-positive result.

The SAMHSA report stated that the data compiled for the 36-month review "clearly indicate that the performance monitoring program has been adept not only in identifying the numerous unforeseen problems that have occurred in the drug testing process of NRC licensees' FFD programs, but also in initiating corrective actions" (Fleming et al., 1993, p. D-5). The overall decrease in unsatisfactory test results has been attributed to licensee initiatives, such as bar code labeling of specimens, additional review steps, and procedural modifications to keep specific problems associated with unsatisfactory testing results from recurring. The SAMHSA report also suggests that "those errors associated with the inappropriate manufacture, formulation, or packaging of quality control materials may need to be addressed by regulatory guidance" (Fleming et al., 1993, p. D-6). Overall, however, SAMHSA found that the small number of unsatisfactory testing results compared to the large number of specimens tested during the 36-month period

"attests to a high level of performance" in the NRC's monitoring program (Fleming et al., 1993, p. D-6).

### 2.6.5: Drug Retention Over Time

Chapter 8 of this report discusses the period during which testing can detect drugs in urine. Normal metabolism breaks down drugs and eliminates them from the body relatively quickly. Most drugs, including the substances for which the NRC's FFD rule requires testing, are eliminated from the body in three days or less. Marijuana is an exception; it may be retained at detectable levels for relatively long periods. Frequent use can cause marijuana to be stored in fat cells and released slowly over a period of weeks or even months following use. In contrast, alcohol is retained for the shortest period, being metabolized at a rate of about one drink per hour (DuPont, 1991). The other four drugs covered under the SAMHSA and NRC guidelines (cocaine, opiates, amphetamines, phencyclidine) fall somewhere between alcohol and marijuana in how long they or their metabolites may be retained in the body at detectable levels.

There are numerous conditions that affect how long a drug or metabolite may be retained in the body and, thus, detected. Such conditions include the chemical nature of the drug, i.e., how amenable it is to being metabolized by the body; the relative health of the user; and the efficiency of the liver and kidneys in breaking drugs down into their metabolites and eliminating the metabolites from the body. Frequent, chronic drug use can overcome normal physiological cleansing, and some drugs or their metabolites can be retained for considerable periods.

The discussion of retention of drugs over time in Chapter 8 concludes that all but one of the drugs covered under the SAMHSA guidelines are relatively quickly metabolized and eliminated by normal body physiology. Marijuana, the only exception, remains in the system at detectable levels for extended periods of time in the small subset of users that use marijuana heavily over a long period. Thus, the drug retention pharmacokinetics of these drugs limits the effectiveness of drug testing as a tool for identifying the occasional or moderate drug abuser.

### 2.6.6: Impairment

The NRC's FFD rule is intended to prevent worker impairment. To understand the relationship between impairment and drug test results, it is useful to distinguish

between *acute*, *carry-over*, and *chronic* effects (DuPont, 1991). Acute effects are those that are usually short acting. For example, the "high" caused by alcohol and other stimulants happens relatively quickly—within a few minutes to an hour.

Carry-over effects of drug use include effects such as hangovers caused by alcohol and other drugs. Depending on the amount and frequency of drug use, these carry-over effects may be mild, moderate, or severe. The performance of workers can be seriously impaired by any of these effects. Carry-over effects usually last from one to two days.

Chronic effects result from long-term and/or heavy use. It is widely accepted that some drugs leave the chronic user with a "protracted abstinence syndrome." This long-term effect can prevent the once heavy user from returning to "normal" for months or years. The discomfort a former addict may experience could lead to the resumption of drug use at any time (Schuckit, 1991). Some of the signs of this syndrome are decreased appetite, depression, increases or decreases in blood pressure or body temperature, and insomnia. The effects of chronic drug use are serious due to the mental and physical damage caused by prolonged drug use. Chronic effects last from weeks to years.

Acute, carry-over, and chronic effects of drug use can result in visible and non-visible impairment (DuPont, 1991). Acute, visible impairment effects include stumbling gait and slurred speech. Non-visible effects include failing memory, depression, and poor judgment. Both types of drug use effects can have serious implications for the safety of workers and the public.

For FFD testing programs conducted by NRC licensees, it is important to remember that although an individual can test positive while impaired, an individual may also test positive for drug or alcohol use when he or she is no longer acutely impaired. However, with carry-over and especially chronic effects, the reverse may be true; an individual may be impaired even when drugs or alcohol are no longer detectable in the body. In other words, an individual may be suffering from carry-over or chronic impairment effects but not test positive. Drug or alcohol testing results alone are not adequate to evaluate impairment. For example, a negative test result does not demonstrate that a worker is not impaired.

### 2.6.7: Deterioration of Metabolites in Urine Specimens

There is some concern that if a specimen is subjected to excessive heat during shipment or while awaiting testing, drug or drug metabolites in the urine specimen may degrade, resulting in a false negative test result. Chapter 9 in this report discusses a study that was conducted to determine whether storage of urine samples for different periods of time and at different temperatures could alter the concentration of marijuana or cocaine metabolites in specimens "spiked" with a known quantity of these two drug metabolites. Experimental results indicate that marijuana metabolite concentration is not significantly reduced in urine at various pH levels when stored for 72 hours at temperatures up to 37°C. However, storing a specimen with pH level of 9 at 37°C for 72 hours resulted in a reduction in the concentration of the cocaine metabolite benzoylecggonine (BE) in two specimens. A follow-up experiment focused on potential loss of BE and found that there was a significant reduction of BE in specimens with a pH level of 9 after 36 hours of storage at 37°C. The study reported in Chapter 9 concludes that when testing urine for the presence of cocaine, care should be taken to ensure that urine specimens are not subjected to excessive heat prior to testing.

### 2.6.8: Extrapolation Issues

As presented in Moore et al. (1989) and Durbin et al. (1991), procedures for back extrapolation, or back calculation, of blood alcohol concentration (BAC) levels have been developed. These procedures allow estimates of the probable BAC level at some previous point in time (for example, in the case of NRC licensees, the beginning of a worker's shift), based on observed BAC levels at a later point. These calculations may include adjustments for sex, age, percent body fat, and body water content, all of which are known to influence the rate at which alcohol is absorbed into the body and metabolized. Moore et al. (1989) concluded that although back extrapolations for BAC must be used with caution, back extrapolations can provide highly reliable evidence of a policy violation. Moore et al. also concluded that the reliability of back extrapolations appears to be substantial enough that it is difficult to justify not performing them—especially if a policy requiring a pre-duty abstinence period exists. Based on the work of Field and Basteyns (1988), however, Durbin et al. (1991) concluded that back extrapolation of blood alcohol levels at very low BAC measurements, such as 0.01 percent, are of limited accuracy. Nonetheless, they suggest that action may be

## General Review

advisable on BAC levels of 0.02 and 0.03 percent in certain cases, particularly for-cause testing cases.

A number of recent studies continue to raise concerns about the reliability of BACs predicted from back extrapolations. The first concern involves alcohol being used in combination with other drugs. In such a case, the metabolism of one substance may interfere with the metabolism of the other substance. In most cases where this occurs, alcohol metabolism is not affected and the metabolism of the other drug is slowed. However, there are exceptions to this pattern, and medications that interfere with alcohol metabolism have been identified. In such cases, a given level of alcohol consumption will result in an atypically high BAC. Case studies in which this effect has been observed are reported by Naranjo and Bremmer (1993) and in an article in the *Journal of the American Medical Association* (Modell, Taylor and Lee, 1993).

The findings of a relatively small scale study by Watkins and Adler (1993) indicate that the presence of food in the stomach may slow rates of alcohol metabolism. In this study, the effects of stomach contents on absorption were observed to be relatively minor, while the presence of food in the stomach appeared to affect alcohol elimination rates. As would be expected, the presence of food in the stomach reduced peak BACs (on average, peak BACs were 21% lower under the full-stomach condition). Elimination rates under both conditions were within the range commonly used by forensic scientists to calculate prior BAC levels.

Two additional studies indicate that the assumption that BAC rates both rise and fall at constant rates throughout the time course of alcohol absorption and elimination may be inaccurate and may lead to inaccurate back calculated BACs. Smith et al. (1993) found that for many subjects, BACs dropped relatively quickly for a short period of time subsequent to drinking. The findings indicate that back extrapolations based upon BACs measured within one hour of drinking may overestimate BACs. Smith et al. (1993) also found that the time course of alcohol disappearance from the blood differed by gender and previous drinking behaviors. For one of these groups—women defined as alcohol misusers—rates of alcohol elimination were 30 percent faster than in the control group. This finding indicates that the assumption of “typical” pharmacokinetics could lead to underestimating BACs in cases where the subject is female and a heavy user of alcohol. The alcohol pharmacokinetics observed by Al-Lanqawi et al. (1992) also found that the magnitude of error in back

extrapolations can be expected to increase as one attempts to calculate further back in time.

The concerns described above are important in traffic safety situations where a BAC at a particular point in time must be clearly established. In these cases, where the traffic violation may have occurred shortly after drinking, it is important to consider variation in absorption rates. Even under these difficult situations, however, the practice has withstood legal challenges (Jackson et al., 1991).

Back extrapolation is less problematic under the NRC's FFD regulation because of the required pre-duty abstinence period. In the nuclear industry, it can be assumed that all alcohol absorption has taken place during this pre-duty abstinence period. The potential for overestimating BAC levels within the first hour reported by Smith et al. is likewise not relevant to this context. Therefore, the conclusions presented in Durbin et al. (1991) are still valid. Use of back extrapolation does appear to be a valid tool in the nuclear industry for determining violations of the FFD policy on alcohol consumption.

### 2.6.9: Alternative Testing Methods

Urine testing is currently the most widely established and accepted testing method for the detection of illegal drug use; for the detection of alcohol, evidential grade breath alcohol analysis is the primary testing method. A variety of alternative drug and alcohol testing methods are also available. Chemical tests that rely on blood, saliva, and hair specimens were reviewed previously (Barnes et al., 1988; Moore et al., 1989; and Durbin et al., 1991). Alternatives to chemical testing (performance- and behavioral-based testing) were also reviewed previously and are discussed in detail in Chapter 10 of this report. The strengths and weaknesses of these and other alternative drug and alcohol testing methodologies are summarized in this section.

#### 2.6.9.1: Alternatives to Urinalysis Drug Testing

Alternatives to urine testing either rely on the analysis of different specimens for the detection of illegal drug use, or they use new methods of analyzing urine that differ from traditional immunoassays. The methods discussed in this section include the analysis of saliva, hair, sweat, and blood and new developments in urinalysis.

### Saliva Analysis

Methods for using saliva to detect illegal drug use are not well established (saliva testing methods for alcohol are more established, see Section 2.6.9.2). Many questions and potential problems need to be addressed before using saliva to test for illegal drug use is feasible in the workplace.

Some of the potential benefits of performing saliva tests apply to both alcohol and illegal drug use detection. Saliva is a readily available substance. The collection of saliva probably does not carry the stigma of intrusiveness and anxiety to the extent that urine or blood collection does. Saliva sample collection, handling, shipping, testing, and storing requires less time, expense, and training to perform than other methods, and testing kits are readily portable (Miller, 1993).

Drugs such as marijuana, cocaine, phencyclidine, opiates, barbiturates, amphetamines, and diazepines can be detected in saliva using various methods, including immunoassay, gas chromatography/mass spectrometry (GC/MS), and thin layer chromatography—TLC (Schramm et al., 1992). It appears that saliva contains higher concentrations of many drugs than blood or urine and provides a more immediate picture of what drugs are active in a person's system (Schramm et al., 1992; *Drugs in the Workplace*, 1993). Therefore, saliva has the potential for acting as a better indicator of impairment at the time of testing.

However, numerous doubts have been raised about the accuracy of drug testing using saliva. For example, food consumption prior to testing has been found to affect the ability to detect marijuana in saliva. In addition, the length of the detection period for drugs in saliva is still unclear. Most research has focused on testing within a few hours or days from the time of ingestion. Saliva's usefulness may be limited by a short detection period (Schramm et al., 1993; Schramm et al., 1992). Additional studies are required to determine the relationship between the strength and time of drug dosage and subsequent concentrations in saliva (Miller, 1993).

Another concern regarding saliva drug testing is related to the effect of saliva stimulation on testing results. In a study that examined the effects of stimulated saliva production (participants were given sour candy) compared to unstimulated saliva production (Kato, et al., 1993), concentrations of cocaine and its metabolites were reduced considerably with the stimulation of saliva. The lowering of the concentrations in saliva in the stimulated

flow condition was attributed to the increase in saliva pH associated with increased saliva flow rate. This finding has implications for collection techniques used in saliva testing.

In sum, although one study tentatively concluded that saliva analysis had potential as a screening mechanism (Schramm et al., 1993), the use of saliva for drug testing in a mandatory program requires further study.

### Hair Analysis

The use of hair analysis to detect drug use has increased in recent years. It is now being used by the Federal Bureau of Investigation (Miller, et al., 1995), in forensic cases (Moeller, et al., 1995), and as a tool for assessing prenatal exposure of newborns (Klein and Koren, 1995 and Katikaneni, et al., 1995). In addition, it has been used in the workplace for employee testing (Gest, 1990). Casino operators in Nevada have been using hair analysis for drugs in pre-employment drug testing (*Drug Detection Report*, 1993b).

Hair has a number of advantages for use in workplace drug testing programs. First, hair testing is more effective in identifying a history of drug abuse. While urine and blood retain drugs and their metabolites for only a short time (see section 2.6.5 of this chapter and Chapter 8 of this report), hair retains drugs for long periods. Second, hair is a better source of information on drug abuse by those who use drugs only occasionally. The period during which an occasional user is at risk of testing positive is much greater with hair analysis (two or more months) than with urine analysis (usually only one or two days).

Despite the advantages of hair analysis, there remain questions regarding its use in federally mandated programs.

Since this methodology is still relatively new, it has faced comparatively few legal tests. Two recent events, however, provide early indications of mixed acceptance of hair analysis as a tool to assist employers in making personnel decisions. In one case, the Nevada Employment Security Department Board of Appeals upheld a previous decision that random hair testing was a valid means to determine that unemployment benefits could be denied to a worker who tests positive and is fired (*Drug Detection Report*, 1993c). In Maryland, however, the state Attorney General's office temporarily suspended a firm from testing the hair of its employees for drugs until a decision could be made as to whether

## General Review

the testing method was acceptable and fell within the state law definition of drug testing by certified labs (*Drug Detection Report*, 1993a).

In addition to questions regarding the acceptance of hair analysis in court cases, some of the disadvantages and unresolved questions surrounding this methodology include:

- Hair analysis is more expensive than urinalysis—approximately \$40-\$65 per test compared to \$25-\$30 for urinalysis (Lesar, 1992).
- Hair does not provide an indication of recent drug use because drugs appear to take a period of days to be excreted into the hair shaft (Gropper, 1995). Hair analysis is thus inappropriate for use in for-cause or follow-up testing.
- Many in the scientific community continue to perceive hair analysis for employee and pre-employment drug testing as premature due to the lack of clinical tests to establish the reliability and validity of analytical methodologies (see Cone and Welch, 1995 for a discussion of the current status of this issue). Establishing generally accepted hair analysis techniques and determining accuracy, precision, sensitivity, and specificity of hair analysis will probably be necessary before this testing is mandated in federal programs. Different methods have been successfully used for hair analysis, but few clinical studies have been performed and more are required to verify their accuracy and reliability (Cone, 1990; Society of Forensic Toxicologists, 1992; Stephenson, 1993; Mieczkowski et al., 1993). Some of the techniques that are currently used by a major marketer of hair analysis have not been independently studied due to the proprietary nature of the technology (Mieczkowski, 1992; Psychomedics, 1992). The Society of Forensic Toxicologists (SOFT) performed a technical review of hair testing and determined that confirmation using technology similar to GC/MS should be required for any positive screen (Society of Forensic Toxicologists, 1992).
- The potential for false positives due to environmental contamination has not been adequately researched, but some evidence suggests that environmental contaminants to hair can result in positive test results (Cone and Wang, 1995). University of California researchers, funded by the National Institute of Justice (NIJ) and SAMHSA, determined that hair is susceptible to such contamination (Stephenson, 1993).

The extent to which external contaminants could seep into or onto hair shafts and affect test results is not clear (Mieczkowski, 1992; *Drugs in the Workplace*, 1993). Standardized procedures for washing hair specimens to remove contaminants have not been established. The National Institute of Standards and Technology (NIST) recently evaluated hair testing and washing techniques and found they could remove most but not all environmental contaminants (Stephenson, 1993; Cone and Wang, 1995). An established means to either remove all contaminants or distinguish between drug traces that enter hair externally and internally would appear to be necessary if external contaminants are found to affect test results.

- The minimum dose required to produce a positive test result is not known, nor have cutoff levels been established (Mieczkowski, 1992; Mieczkowski et al., 1991). Some research has addressed this question. For example, a University of California study determined a general relationship between cocaine dose and quantities measured in hair (Stephenson, 1993). However, variability between individuals was still great. This raises the question of the extent to which individual differences can affect hair analysis.
- Research on the effects of individual differences in race, sex, and hair texture and type on test results indicates that hair type has a significant impact on test results as do chemical treatments such as hair dyes and permanent wave solutions (Kidwell and Blank, 1995).
- The mechanisms of drug excretion into hair are not fully understood (see Kidwell and Blank, 1995, for a review). The mechanism by which drugs are transferred from the body to hair is another area that requires additional study. It has been widely believed that the transfer occurs directly from the blood to the hair. However, researchers from the University of California, expecting to find in hair the cocaine metabolites that are broken down quickly in the bloodstream, found instead the "parent" cocaine drug. This has raised the question of whether some other mechanism, like sweat, may be the medium for drug excretion into hair (Stephenson, 1993). These issues have led many in the scientific community to conclude that hair analysis cannot be supported at this time.

Overall, hair analysis has potential advantages for identifying drug abuse, particularly in pre-employment

testing and in identifying occasional users, and the development of analytic techniques for hair analysis is proceeding rapidly (see Cone, Welch, and Babbecki, 1995, for current research on the standards and technology for hair testing). However, at this time there are many remaining questions about the accuracy and fairness of hair testing. These questions are particularly salient to federally mandated drug testing programs where protection of individual rights is critical.

### Sweat Analysis

A transdermal patch that is designed to permit detection of drug use through analysis of collected sweat is one of the latest developments in drug testing technology. The FDA has recently granted approval to PharmChem Laboratories for its transdermal patch, PharmChek. PharmChek can detect the five drugs covered under the SAMHSA guidelines (marijuana, cocaine, opiates, phencyclidine, and amphetamines). However, at present, FDA approval has only been granted for cocaine, opiates, and amphetamines. Marijuana and PCP are still waiting for authorization. The FDA approval allows PharmChem to sell the device for workplace drug testing to private industry, and also for criminal justice, rehabilitation, and military testing programs (*MRO Newsletter*, December, 1995). The sweat patch consists of a two-inch by three-inch transparent film, with an absorbent pad in the center, that is adhered to the body for a period of approximately two weeks (PharmChem, 1993). The absorbent pad collects passive uncontrolled sweat loss and its components, including drugs. When the patch is removed, the pad is shaken in a liquid medium to remove any drugs and the liquid is subsequently screened with an immunoassay. As with other screening methods, GC/MS can be used to confirm positive test results (Pharmchem, 1993).

The patches appear to be most useful for drug testing programs in criminal justice and drug treatment settings, but could potentially be used for employee testing (PharmChem, 1993). There are several potential benefits of this new technology. The sweat patch will collect cumulative amounts of drugs and their metabolites over the period of wear, so it has the potential to provide a longer period of testing. For example, a sweat patch has been demonstrated to record a cocaine dose for up to seven days following a dose, whereas urine tests were no longer positive after 72 hours (Burns and Baselt, 1993). The developer claims the cocaine detection patch can detect low levels of cocaine as soon as two to five hours after use, depending on the dose. The patch is also easy to apply and is simpler to handle and transport than are

urine specimens. The patch may also offer benefits in avoiding subversion of the testing process. According to the developer (PharmChem, 1993), the patch cannot be removed without damaging the adhesive, clearly indicating that it was tampered with.

There are a variety of issues to resolve before the sweat patch could be considered for use in a workplace drug testing program. Minimum drug detection levels and cutoff levels have not been established. The potential for environmental contamination leading to false positives has not been thoroughly examined for most of the drugs on the panel. Individual variability in drug detection levels following similar doses has been found (PharmChem, 1993), pointing out the need to examine the effect of variables such as age, sex, race, metabolism, sweat rates, and activity levels on sweat patch drug collection. Little independent research has been performed to date on the sweat patch to attempt to resolve these issues.

Even if these issues are resolved in the future, the most significant drawback of the sweat patch—a perceived high level of intrusiveness—may seriously limit its application in employee drug testing programs. It is likely that being asked to wear a collection device for a week or two at randomly selected periods (or continuously) would be highly objectionable to many people. This factor alone would appear to seriously limit the potential for widespread application of this technology.

### Blood Analysis

The analysis of blood for drug use detection has been discussed in previous reports (Barnes et al., 1988; Moore et al., 1989; and Durbin et al., 1991). Some of the potential advantages of blood testing over urinalysis that were identified in these reports include the potential for establishing a greater correlation between drug levels in the blood and impairment, the small likelihood of subversion, and the perception by some that it may be less intrusive than urinalysis.

Previously discussed disadvantages of blood testing include the lack of knowledge concerning the relationship between dose and effect needed to establish a firm link between blood test results and impairment, the potential that some individuals would view blood collection as more intrusive than urine or breath testing, the short detection window due to rapid metabolism of drugs and alcohol, and public perceptions of potential health risks posed by blood collection.

## General Review

Resolution of these issues does not appear to be any closer now than when these earlier reports were published. No new research was identified that addresses these issues.

### New Developments in Urinalysis

Urinalysis is currently widely used. Techniques for urinalysis are specified under mandatory DHHS guidelines (59 FR 29908) and are reviewed in previous reports (Barnes et al., 1988; Moore et al., 1989; Durbin et al., 1991). New non-instrumented testing technologies have been developed since those reports were published.

These new technologies are designed to provide a fast, simple, accurate methodology that is relatively inexpensive to perform. The tests are portable and have potential as screening tests for use on site or in remote areas, particularly when access to formal laboratory testing equipment is limited by time or location (Schwartz et al., 1993). The tests have been shown to be relatively reliable and are being used in hospital emergency rooms (Schwartz et al., 1991). These tests are being considered for use under the DOT drug testing program. However, little if any research has been performed on these procedures, especially for drugs other than marijuana and cocaine. The DHHS has been given a mandate by Congress to investigate the accuracy and reliability of these devices. However, as of January, 1996, sufficient funds for such an effort had not been earmarked (Dr. Donna Bush, SAMHSA, personal communication, January 31, 1996).

Under the current rule, licensees are not precluded from using these rapid, non-instrumented testing devices. (However, drug testing procedures within the nuclear industry require that all screened positives (whether using instrumented or non-instrumented devices) be forwarded for confirmation to a certified laboratory that uses highly trained personnel and proven testing equipment.) There are indications that these devices are being used and/or will be used by NRC licensees for on-site screening in their FFD programs. There are several issues that make use of these devices a potential problem. These issues include:

Quality assurance procedures may be difficult. For example, procedures are needed for users (e.g., licensees performing on-site testing) to validate the accuracy and reliability of these devices around a cutoff level.

- Although manufacturers state that the accuracy and specificity of these tests (e.g., the expectations of

false positives and false negatives) are comparable to other immunoassays (e.g., EMIT®, RIA), their accuracy and specificity have not been thoroughly established by independent agencies.

- The potential for subversion using these devices has not been addressed. For example, the impact of an adulterated or dilute specimen on the test results is not known.

One rapid on-site test that currently appears promising is AbuScreen ONTRAK®, an immunoassay test for individual drugs based on latex agglutination. A positive or negative test result is provided within three minutes after a relatively simple procedure, and is indicated by the level of latex particle agglutination on the card. Several tests can be performed simultaneously. Accuracy checks done by the manufacturer, including one comparison to GC/MS results for the cocaine and marijuana tests, concluded the test was highly accurate. Although only studies of marijuana and cocaine test cards were identified, additional kits have been developed for opiate, amphetamine, barbiturate metabolites, and phencyclidine (PCP). Some training is required to ensure proper results (Schwartz et al., 1993; *Drug Detection Report*, 1993d). This testing device was recently selected by the Florida Department of Corrections for use in testing of probationers (*Drug Detection Report*, 1993d).

Another non-instrumented testing device, EZ-SCREEN®, is a qualitative enzyme immunoassay test that, while being fairly accurate for cocaine and marijuana tests, requires following a set of labor-intensive, critically-timed steps. A small number of false negatives for cocaine were produced in one study (Schwartz et al., 1993). Another testing device, microLINE® Screen, has exhibited a tendency to produce false positives, false negatives, or no definite results (*Drug Detection Report*, 1993d). Factors such as complexity and inaccuracy diminish the potential benefits of these testing devices.

Another test developed for rapid, on-site screening is the KDI Quik Test®. The KDI Quik Test® is a broad-spectrum paper chromatography screening test for cocaine, opiates, amphetamines, and phencyclidine. A separate test can be performed for marijuana. According to Schwartz et al. (1993), paper chromatography tests have not been found to be extremely accurate. Their study of the KDI test produced many false positives and found the methodology to be labor-intensive and time-consuming, failing to warrant use in any setting.

Appendix B of this report discusses additional potential problems with these devices and provides additional information on seven non-instrumented testing devices.

#### 2.6.9.2: Alternatives to Breath Alcohol Testing

Alcohol testing is primarily performed using breath alcohol testing devices. Nonetheless, a variety of alternative alcohol testing methods undergo continual study and evaluation, and are discussed in this section.

##### Saliva Analysis

A great deal of saliva testing research has been directed toward its applicability in testing for alcohol, and it is in alcohol testing that the development of saliva testing technology is most advanced. In addition to the potential benefits discussed above for illegal drugs (less intrusive, easy to collect and handle, less expensive, easy to use, portable), certain saliva testing methods may also produce results that are relatively sensitive and correlate well with BACs derived from blood analysis.

Several methods have been developed for saliva alcohol screening. Two methods were reviewed by Durbin et al. in 1991. One method uses a test strip with an alcohol-oxidase methodology that reacts when wetted with saliva to produce a color change corresponding with BAC. The test is inexpensive, non-invasive, highly portable, and easy to administer. Bates et al. (1991) concluded, however, based on a research consensus, that this type of semi-quantitative saliva test, which requires comparison to a color chart, was not precise enough to provide accurate estimates of BAC. While the method shows some promise as a screening mechanism to determine the presence or absence of alcohol, this method has produced inconsistent results, particularly with blood alcohol levels below .05 percent (Schwartz et al., 1993).

Another saliva alcohol screening methodology has been developed that requires the placement of a detection reagent and a drop of saliva on a card. A particular color appears if alcohol is present to react with the reagent's chromium trioxide at concentrations greater than 20 mg/dL. The manufacturer (Toxi-Lab) claims a 98 percent correlation with GC/MS results. No studies have been identified that support or disprove this claim. The manufacturer claims that the test can provide similar results with urine instead of saliva. However, the results are qualitative and provide a simple "pass or fail" result. The test also produces positive results within 10 minutes after use of 14 percent alcohol mouthwash (Schramm et al., 1993).

The most advanced saliva alcohol screening method requires the use of a swab to collect saliva and an electronic device that determines BAC. Considerable research on the Enzymatics QED A-150® (Enzymatics, 1990; Christopher and Zecardi, 1990) has demonstrated its usefulness as an alcohol screening device, and Durbin et al. (1991) reported the viability of this device. The system has received final FDA clearance as a quantitative saliva alcohol test (Miller, 1993). A DOT study (National Highway Traffic Safety Administration, 1992) evaluated the device's potential in a highway safety setting and confirmed its potential after finding that resulting estimated BACs were highly correlated (0.98 or greater) with a wide range of known alcohol concentrations. Varying ambient conditions were established during the DOT study with no considerable impact on test results. The tests consistently underestimated the target BACs at all alcohol concentrations and temperatures. Like the Toxi-Lab saliva test, the QED A-150® is susceptible to registering a positive up to 10 minutes after use of 14 percent alcohol mouthwash (Schwartz et al., 1993). The DOT has added the QED Saliva Alcohol Test to its Conforming Procedures List (CPL). This product, manufactured by STC Diagnostics, can now be used for alcohol screening performed under the Omnibus Employee Testing Act of 1991 (*MRO Alert*, 1995).

##### Blood Analysis

The analysis of blood to test for alcohol has been discussed in previous reports (Barnes et al., 1988; Moore et al., 1989; and Durbin et al., 1991). After evaluating several alcohol testing methodologies against the performance criteria of specificity, sensitivity, and precision, the Federal Aviation Administration (FAA) found blood testing to be the most desirable of fixed methods (McLean et al., 1991). Nevertheless, there are a variety of disadvantages of testing blood. Barnes et al. (1988), Moore et al. (1989), and Durbin et al. (1991) previously discussed some of these potential problems, including concern that blood analysis may be viewed as more intrusive than urine or breath testing, the short detection window due to rapid metabolism of alcohol, and public perceptions of potential health risks posed by blood collection. Resolution of these issues does not appear to be any closer, as evidenced by the lack of new research identified for this update.

##### New Developments in Breath Tests

A variety of breath alcohol testing devices have recently been developed or improved. One of these is the

## General Review

electronic fuel cell instrument, which has recently been classified as meeting evidential-grade breath alcohol analysis device standards by the National Highway Traffic Safety Administration (49 FR 48855; 58 FR 48705). The advantage this device has over many current breath alcohol measuring devices is cost and portability. Portable units typically cost approximately \$500 to \$800 and have low per-test cost. Fixed units can also be purchased (McLean et al., 1991). The electronic fuel cell instrument may provide additional flexibility to licensees in the breath alcohol testing equipment they have available for screening and confirmatory purposes.

Other developments include a tin oxide ceramic sensor and dichromate crystals. These methods are cruder and more qualitative in nature. The tin oxide ceramic sensor has demonstrated poor specificity, reacting to hydrocarbons and alcohols other than ethanol. It is also not as sensitive as the methods currently used (McLean et al., 1991). Glass tubes packed with impregnated dichromate crystals provide a semiquantitative indication of breath alcohol (Schwartz et al., 1993). When breath is exhaled into the tube, the crystals change color in response to the presence of alcohol. Although some of these devices are calibrated to indicate low breath alcohol levels, they are not extremely sensitive or precise and appear to work best at high alcohol levels (Schwartz et al., 1993; McLean et al., 1991; Flores, 1994).

### 2.6.9.3: Performance-Based Assessment Tests

Performance-based assessment tests can provide direct measures of an individual's ability to perform specific tasks required on the job. These tests have been proposed to detect worker performance impairment, a key concern in safety-sensitive industries.

Previous reviews of the technical issues surrounding the NRC's FFD program examined performance-based tests and test batteries and concluded that the empirical data did *not* support using these tests in an FFD program (Moore et al., 1989; Durbin et al., 1991). Current research, however, supports the re-examination of these tests. For example, during the past few years, continued research using performance-based testing technology has advanced this area by examining performance-based testing methods and outcomes. In addition, the number of vendors marketing these tests and test batteries for purposes of FFD testing has increased.

Government agencies and research programs are currently evaluating this body of literature and vendor materials to determine the feasibility of performance-based tests for

safety-sensitive jobs (e.g., Federal Aviation Administration and National Research Council; see Gilliland and Shelegal, 1993, and Bittner et al., 1993). There is also evidence that these tests are being used in some industries to screen out impaired workers on a daily basis. Results from the American Management Association survey of drug testing policies in 1993 indicated that about one percent of the companies surveyed used performance-based testing, but these tests were often used in conjunction with other FFD testing procedures (American Management Association, 1993; see also Chapter 3, Drug Use Prevalence, in this report). Finally, a recent informal survey of vendors indicates that many process-control industries and other organizations are currently using performance-based testing to ensure worker fitness for duty.

Chapter 10 in this report provides a detailed examination of performance-based tests that may potentially have a role in the commercial nuclear power industry's FFD testing programs. The chapter identifies and describes a variety of specific tests that have been developed or have a potential for application to FFD programs, and evaluates these tests in terms of their ability to identify impaired employees (i.e., those who cannot adequately perform their required job duties). The chapter concludes that, at the present time, performance-based tests are available for implementation in industry. These tests are reliable and sensitive to drug use, environmental exposure to chemicals, and problems associated with sleep loss and stress. Furthermore, an analysis of a specific selected test presented in the chapter indicates that such tests can be highly accurate in detecting significant impairment such as that caused by a BAC of 0.15. One problem associated with performance-based tests is the lack of quantitative criterion-related validity information for any of the performance-based tests summarized in the literature. Further studies linking performance-based test results to impairment are needed.

Information supplied by vendors suggests that some performance-based tests are currently being used in industry. At this time, specific information about implementation procedures and procedures for handling failures is not available. Before performance-based tests become widely used, it is important to collect information about the procedures that other industries are using or have tried and to collect data about the merits of performance-based tests versus urinalysis in terms of employee acceptability, avoidance of incidents and accidents in the workplace, and organizational outcomes. For further information regarding performance-based assessment tests, see Chapter 10 in this report.

## 2.6.10: Shy Bladder

An issue of concern in workplace drug testing is the "shy bladder" issue. A shy bladder refers to the difficulty that some individuals encounter in trying to provide a sufficient urine specimen within a designated time period. The American Association of Medical Review Officers (AAMRO) formed a technical committee to develop guidance on this issue for the DOT and the MRO community.

Under previous rules, the DOT permitted donors up to 8 hours to provide a urine specimen. This procedure was subject to abuse. Some employees would take the day or afternoon off, stating that they were unable to provide a specimen. A related problem that concerned DOT was over-hydration resulting in a compromised specimen. To correct these subversion problems, DOT limited the collection period to 2 hours and limited the amount of liquids that an individual could consume to 24 ounces. When this change was officially announced, it raised a great deal of concern. The concern centers around the possibility of genuine medical explanations for failure to provide an adequate urine specimen.

The AAMRO Technical Committee on Shy Bladder released a preliminary report on the subject (*MRO Alert*, 1994). The majority of the committee members responding indicated that there appeared to be a number of conceivable circumstances where an individual trying to provide urine in good faith may not be able to produce the required 45 mL in two hours. Furthermore, the majority of members felt that the risk of internal dilution was minimal at intakes of up to 48 ounces. The preliminary report thus recommended that consideration be given to lengthening the period to provide the urine to more than two hours and permitting more than 24 ounces of fluid.

In response to the numerous complaints by employers, employees, and MROs, as well as the AAMRO report's findings, the DOT has proposed to change the rule so that employers and job applicants would be allowed up to 4 hours to produce the required specimen. The change would also allow the person to consume up to 40 ounces of liquid during that period. The proposed rule was published in the July 25, 1995 *Federal Register* (*Drugs in the Workplace*, 1995).

In addition to the "four hour/forty ounce" proposal, the DOT is proposing to adopt some of the language from its alcohol testing rules into the drug testing rule in an effort to provide MROs with more guidance when trying to

determine whether the inability to produce urine is genuine. The language to be incorporated deals with the inability to provide enough breath for the breath testing device, and would be altered to apply to urine. In essence, the proposed guidance would specify that if the individual has not provided a sufficient specimen within four hours from the time the original insufficient urine specimen was presented to the collection site person, the collection site person shall discontinue the collection and notify the employer. The individual would then be directed to obtain, as soon as practical, an evaluation from an acceptable, licensed physician regarding the individual's medical ability to provide an adequate amount of urine. If a medical condition is determined to be the cause, failure to provide an adequate amount of urine will not be considered a refusal to take the test. Otherwise, failure will be deemed a refusal.

The report also provides some support for reinstating the ability to give the donor a second chance to provide a specimen on short notice, within 1-2 days of the examination. If the donor failed to provide a specimen, then he or she could submit to a medical examination by a licensed physician.

## 2.6.11: Alcohol Testing Issues

In a previous review of research, Moore et al. (1989) addressed the question of balancing the objectives of the NRC's FFD rule, 10 CFR Part 26, with numerous other factors relevant to workplace alcohol programs. Among the recommendations made in that publication were prohibitions against on-duty use of alcohol and support for the use of breath testing for the detection of unacceptable off-duty alcohol use (using a 0.04% BAC cutoff level). Moore et al. also recommended requiring that workers covered by the FFD rule abstain from all consumption of alcohol for a period of at least five hours prior to a scheduled tour of duty. In addition to these general conclusions, Moore et al. examined the question of whether or not any actions are appropriate when workplace breath tests indicate recent alcohol use, but use that did not result in a BAC at or above the 0.04 percent BAC cutoff level. Finally, that publication also addressed in detail technical considerations pertaining to testing for unacceptable alcohol use such as specimen selection and the devices and techniques to be used in testing.

Since the publication of Moore et al. (1989), no evidence has been identified that provides a significant challenge to the analysis presented as a basis for the recommendation of a five-hour pre-duty alcohol abstinence period. In addition, no evidence has been identified since the

## General Review

publication of Moore et al. that significantly challenges any of the findings that led to the recommendation of the 0.04 percent BAC cutoff level. In general, ongoing research substantiates the findings presented in Moore et al. that provided the basis for the recommendation of this cutoff level. In recent years, the reliability of breath testing technology to support testing at this cutoff level has improved, and saliva tests are becoming a viable form of alternative testing for alcohol (see Section 2.6.9, Alternatives to Breath Alcohol Tests, in this chapter). In addition, the use of a 0.04 percent BAC cutoff level appears to be emerging as a standard for workplace alcohol programs. This standard has been implemented or proposed by the Department of Transportation (58 FR 48705; 59 FR 7372) and the Department of Energy (56 FR 10075).

New research has added to the discussion of back extrapolation of BAC. Section 2.6.8, Extrapolation Issues, summarizes this new evidence. Use of back extrapolation still appears to be a valid tool in the nuclear industry.

### 2.6.12: Occupational Risks in Drug Testing

Since the last published update on FFD technical issues in the nuclear power industry (see Durbin et al., 1991), the Occupational Safety and Health Administration (OSHA) has published guidelines relevant to protecting workers from occupational exposure to bloodborne pathogens. From 1981 through 1992, the Center for Disease Control (CDC) has confirmed that 32 health care workers have been occupationally infected with the human immunodeficiency virus (HIV), the virus that causes acquired immune deficiency syndrome, or AIDS. The CDC concedes that the "number of persons with occupationally acquired HIV infection is probably greater than the totals presented [in its study] because not all health care workers are evaluated for HIV infection following exposures and not all persons with occupationally acquired infection are reported" (*Journal of the American Medical Association*, 1992, p. 3294). Although the risk of contracting HIV in the workplace does not appear to be great (Turk, 1993), data from OSHA have indicated that health care workers who work in a health care, clinical, or HIV-laboratory setting face a significant enough health risk as a result of occupational exposure to blood and other potentially infectious materials because these materials may contain bloodborne pathogens, including the hepatitis B virus and the HIV virus (54 FR 23042). To mitigate these potential health care risks, OSHA established the Occupational Exposure to Bloodborne Pathogens rule (29 CFR Part 1910.1030),

which is applicable to all employees who may encounter occupational exposure to blood or other potentially infectious materials during the normal course of their daily job duties. The OSHA standard does apply to clinical personnel involved in the collection and handling of blood specimens collected pursuant to the NRC fitness-for-duty rule.

Compliance with the OSHA bloodborne pathogen standard does not appear to impose undue hardship on affected facilities. In a 1990 survey of facility types in which occupational exposure to human blood or other potentially infectious material is likely to occur, OSHA found that about 90 percent of affected facilities already had infectious bloodborne disease policies in place. The majority of those facilities required all blood to be handled as infectious (55 FR 10250). Likewise, reviewers in the *MRO Newsletter* concluded that the OSHA standard enforces only the minimum standard required by law (*MRO Newsletter*, 1993).

The requirements of the Occupational Exposure to Bloodborne Pathogens rule fall into five categories: exposure control; methods of compliance; hepatitis B vaccination and post-exposure evaluation and follow-up; communication of hazards to employees; and record keeping. The sections below summarize some of the requirements in these categories. Appendix C provides more specifics concerning the requirements of the Occupational Exposure to Bloodborne Pathogens rule.

#### 2.6.12.1: Exposure Control

Employers are required to establish a written Exposure Control Plan if they have personnel who have a reasonable chance of being exposed to bloodborne pathogens, including the hepatitis B virus (HBV) and HIV, during the normal performance of their job duties. The Exposure Control Plan should be accessible to employees and should include an exposure determination. The Exposure Control Plan should also include schedules and methods of implementation for compliance; hepatitis B vaccination and post-exposure evaluation and follow-up; communication of hazards to employees; and recordkeeping.

#### 2.6.12.2: Methods of Compliance

In general, employees should observe universal precautions when working with blood and other potentially infectious materials; that is, they should treat all human blood and certain human body fluids as if they were known to be infectious for HIV, HBV, and other

bloodborne pathogens. In addition, employers are required to provide the following:

- Engineering and work practice controls, such as handwashing facilities and protocols for handling contaminated equipment and work areas, to eliminate or minimize employee exposure.
- Personal protective equipment, including but not limited to gloves, gowns, face shields, eye protection, and mouthpieces, if employees are still subject to occupational exposure after implementation of the engineering and work practice controls.
- Assurance, at all times, that the worksite is maintained in a clean and sanitary condition. Employees should have access to appropriate equipment and protocols for cleaning and sanitizing the worksite.

#### **2.6.12.3: Hepatitis B Vaccination and Post-Exposure Evaluation and Follow-Up**

In general, employers are required to make available, at no cost to the employee, the hepatitis B vaccine and vaccination series to all employees who may have occupational exposure to blood or other potentially infectious materials. Employers must also provide, at no cost to the employee, post-exposure evaluation and follow-up to all employees who have had an exposure incident.

#### **2.6.12.4: Communication of Hazards to Employees**

Provisions for communication of hazards to employees include ensuring that appropriate warning and biohazard labels and signs are chosen and affixed to containers with potentially infectious materials. These provisions also address the situations in which such labels and signs should be used.

Other provisions for communicating hazards to employees pertain to information and training. Information and training provisions include ensuring that employers provide, at no cost to the employee, appropriate and accessible training to all employees with occupational exposure to blood or other potentially infectious materials. This training must be provided during working hours when an employee is initially assigned to a task where occupational exposure may take place, or within 90 days after the effective date of the OSHA standard, and at least annually thereafter. The person(s) conducting the training must be knowledgeable of the text and requirements of

the OSHA standard; general bloodborne disease epidemiology, symptoms, and path of transmission; and information on protective equipment and protocols. Employers must provide additional training when changes such as modification of tasks or procedures or institution of new tasks or procedures affect the employee's occupational exposure. The additional training may be limited to addressing the new exposures created.

#### **2.6.12.5: Recordkeeping**

Provisions for recordkeeping pertain to medical records, training records, availability of records, and transfer of records. Employers should establish an accurate, confidential record for each employee with occupational exposure to blood or other potentially infectious materials, and should maintain this medical record for the duration of an employee's employment plus 30 years. Employers should maintain training records containing information such as the dates of the training sessions, the contents or a summary of the training sessions, the names and qualifications of the trainers, and the names and job titles of all people attending the training sessions, for three years. Employers should ensure that all medical and training records are maintained in accordance with the OSHA standard. Finally, as necessary, the employer should comply with the requirements of 29 CFR 1910.20 (h) involving the transfer of records, including provisions for transferring all medical and training records to the director of OSHA if the employer ceases to do business and there is no successor employer to receive and retain the records.

### **2.7: A Review of Findings from NRC Inspections**

A previous report, *Fitness for Duty in the Nuclear Power Industry* (Durbin et al., 1991), discussed the results of 24 initial inspections and 3 reactive inspections conducted by NRC regional inspectors between January 3, 1990 and March 1, 1991. The purpose of these inspections was to evaluate implementation of licensee fitness-for-duty programs, to provide information needed to assess the effectiveness of the FFD rule, and to determine whether further FFD rulemaking is needed. These inspections were conducted in compliance with Temporary Instruction (TI) #2515/106, which requires that each site undergo an initial inspection to evaluate its compliance with the FFD rule. Through the spring of 1995, the NRC has conducted more than two dozen additional inspections under TI #2515/106, a few special reactive inspections, and a number of safeguard inspections where FFD issues

## General Review

were addressed. This section discusses the results of all the inspections conducted to date, including those reported earlier. In conducting these inspections, inspectors were encouraged to provide constructive feedback to licensees about their FFD programs. This feedback included discussions of program strengths as well as potential problems with the program.

Overall, NRC inspectors found that licensees have functioning FFD programs devoted to providing reasonable assurance that nuclear power plant personnel are not impaired as they perform their duties. The programs also aim at achieving a drug-free workplace. Although there is some variance in how individual licensees implement the rule and the effectiveness of rule implementation, the program performance reports submitted by licensees and the NRC's FFD inspection efforts disclosed no major problems on an industry-wide level with licensees' FFD programs. Inspectors noted many program strengths in a majority of the inspections. The most frequently mentioned program strengths included strong management support for the FFD

program, effective quality assurance (QA) audit programs, and professionalism exhibited by well-trained FFD personnel. Other program strengths included adequate procedures and testing facilities for the FFD program, effective employee and supervisor training, and well-implemented EAPs that sometimes extended to licensee employee family members and contractor personnel.

In addition to feedback on program strengths, inspectors noted problems in several areas. Violations and non-cited violations were identified in seven FFD program areas: written policy and procedures, supervisor training, contractors and vendors, chemical testing, management actions and sanctions, recordkeeping requirements, and reporting requirements. Unresolved items were identified in nine areas: written policy and procedures, policy communication and awareness training, supervisor training, chemical testing, EAPs, appeals process, protection of information, recordkeeping requirements, and reporting requirements. Details concerning these violations, non-cited violations, and unresolved items are reported in Appendix D of this report.

## Endnotes

1. Charlene Forslund, M.S., Battelle Seattle Research Center  
Nancy Durbin, Ph.D., Battelle Seattle Research Center  
Idene Field, Battelle Seattle Research Center  
Tom Grant, J.D., Battelle Seattle Research Center  
Michelle Silbernagel, M.S., Battelle Seattle Research Center  
Ron Wilson, Battelle Seattle Research Center
2. Since the formation of this committee, NIDA has become the Substance Abuse and Mental Health Services Administration (SAMHSA).

## References

10 CFR Part 26. (January 1, 1990). U.S. Nuclear Regulatory Commission. "Fitness-for-duty programs." *U.S. Code of Federal Regulations*.

29 CFR Part 1910.1030. (July 1, 1992). Occupational Safety and Health Administration. "Bloodborne pathogens." *U.S. Code of Federal Regulations*.

49 CFR Part 40. (August 19, 1994). U.S. Department of Transportation. "Procedures for transportation workplace drug and alcohol testing programs." *U.S. Code of Federal Regulations*.

49 FR 48855. (December 14, 1984). National Highway Traffic Safety Administration. "Highway safety programs; model specifications for evidential breath testing devices: Publication of a conforming products list." *Federal Register*.

51 FR 32889. (September 17, 1986). Executive Office of the President. "Executive Order 12564: Drug-Free Federal Workplace." *Federal Register*.

52 FR 21796. (June 9, 1987). Department of Health and Human Services. "Confidentiality of alcohol and drug patient records." *Federal Register*.

54 FR 23042. (May 30, 1989). Occupational Safety and Health Administration. "Occupational exposure to bloodborne pathogens." *Federal Register*.

54 FR 24468. (June 7, 1989). Nuclear Regulatory Commission. "Fitness for Duty Programs; Final Rule and Statement of Policy." *Federal Register*.

55 FR 10250. (March 20, 1990). Occupational Safety and Health Administration. "Occupational exposure to bloodborne pathogens." *Federal Register*.

56 FR 10075. (March 8, 1991). Department of Energy. "Personnel security assurance program: Substance abuse testing procedures." *Federal Register*.

58 FR 15810. (March 24, 1993). Nuclear Regulatory Commission. "Modifications to Fitness-for-Duty Program Requirements." *Federal Register*.

58 FR 48705. (September 17, 1993). National Highway Traffic Safety Administration. "Highway safety programs;

model specifications for devices to measure breath alcohol." *Federal Register*.

58 FR 5220. (January 19, 1993). Department of Health and Human Services. "Medicare, Medicaid and CLIA Programs; Regulations Implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA) and Clinical Laboratory Improvement Act Program Fee Collection." *Federal Register*.

59 FR 29908-29931. (June 9, 1994). U.S. Department of Health and Human Services. "Mandatory guidelines for federal workplace drug testing programs." *Federal Register*.

59 FR 502. (January 5, 1994). Nuclear Regulatory Commission. "Modifications to Fitness-for-Duty Program Requirements." *Federal Register*.

59 FR 7372. (February 15, 1994). National Highway Traffic Safety Administration. "Highway safety programs; model specifications for devices to measure alcohol in bodily fluids." *Federal Register*.

59 FR 42996-43019. (August 19, 1994.) U.S. Department of Transportation. "Procedures for transportation workplace drug and alcohol testing programs." *Federal Register*.

Abrams, D., Elder, J. and Carleton, R. (1986). Social learning principles for organizational health promotion: an integrated approach. In M. F. Cataldo and T. J. Coates (Eds.), *Health and Industry: A Behavioral Medicine Perspective*. New York, NY: John Wiley and Sons.

Ajzen I and Fishbein M. (1980). *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Al-Lanqawi, Y., Moreland, T.A., McEwen, J., Halliday, F., Durnin, C.J. and Stevenson, I.H. (1992). "Ethanol kinetics: Extent of error in back extrapolation procedures." *British Journal of Clinical Pharmacology*, 34(4): 316-321.

American Management Association. (1993). *Workplace Drug Testing and Drug Abuse Policies: Summary of Key Findings*. New York, NY: Author.

Americans with Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. 328, 42 USC. §12101 et seq. (1990).

Ames, G. (1993). "Research and strategies for the primary prevention of workplace alcohol problems." *Alcohol, Health, and Research World*, 17(1): 19-27.

Bandura A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Barnes, V., Fleming, I., Grant, T., Hauth, J., Hendrickson, J., Kono, B., Moore, C., Olson, J., Saari, L., Toquam, J., Wieringa, D., Yost, P., Hendrickson, P., Moon, D. and Scott, W. (1988). *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues* (NUREG/CR-5227). Washington, DC: Nuclear Regulatory Commission.

Bates, M.E., Brick, J. and White, H.R. (1991). "The correspondence between saliva and breath estimates of blood alcohol concentration: Advantages and limitations of the saliva method." *Journal of Studies on Alcohol*, 54(1): 17-22.

Bennett, R. (1993, March/April). "Performance safety and the EAP." *EAP Digest*: 20-23.

Bickerton, R. (1990, March). "The right to privacy, the need to know: Are they natural enemies?" *EAPA Exchange*: 42-44.

Bittner, A.C., Jr., Toquam, J.L., Forslund, C. and Fleming, T. (1993). *TCRP Project Fitness-for-Duty Testing in the Transit Workplace: Task 2, A Literature Review*. Seattle, WA: Battelle Human Affairs Research Centers.

Bray, R.M., Kroutil, L.A., Luckey, J.W., Wheless, S.C., Iannacchione, V.G., Anderson, D.W., Marsden, M.E. and Duntzman, G.H. (1992). *1992 Worldwide Survey of Substance Abuse and Health Behaviors Among Military Personnel* (RTI/5154/06-16-FR). Research Triangle Park, NC: Research Triangle Institute.

Burns, M. and Baselt, R.C. (1993). *Use of a Skin Patch to Monitor Cocaine Abuse*. Paper presented at the Joint Meeting of the Society of Forensic Toxicologists (SFT) and the California Association of Toxicologists (CAT), October 10-16, 1993, Phoenix, Arizona.

Christopher, T.A. and Zeccardi, J.A. (1990). *Evaluation of the QED Saliva Alcohol Test: A New, Rapid, Accurate Device for Measuring Ethanol in Saliva*. Paper presented at the Society of Academic Emergency Medicine Annual Meeting, May 22, 1990.

Clinical Laboratory Improvements Amendments of 1988, Pub. L. No. 100-578, 102 Stat. 2903, 42 U.S.C §263a (1988).

Cody, J.T. and Schwarzhoff, R.H. (1989). "Impact of adulterants on RIA analysis of urine for drugs of abuse." *Journal of Analytical Toxicology*, 13: 277-284.

Cone, E.J. (1990). "Testing human hair for drugs of abuse. Individual dose and time and codeine in plasma, saliva, urine, and beard compared to drug-induced effects on pupils and behavior." *Journal of Analytical Toxicology*, 14: 1-7.

Cone, E.J. and Welch, M.J. (1995). Introduction and Overview: National Institute on Drug Abuse Special Publication on Hair Analysis for Drugs of Abuse. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. U.S. Department of Health and Human Services, National Institutes of Health.

Cone, E.J., Welch, M.J., and Babecki, M.B.G. (Eds.). (1995). *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. U.S. Department of Health and Human Services, National Institutes of Health.

Cone, E.J. and Welch, P. (1989). Time course of detection of 6-acetylmorphine in urine after heroin administration. In L.S. Harris (Ed.), *Problems of Drug Dependence 1989* (NIDA Research Monograph #95: 449). Rockville, MD: National Institute on Drug Abuse.

Cook, R. and Youngblood, A. (1990). Preventing substance abuse as an integral part of worksite health promotion. In M. Scofield (Ed.), *Occupational Medicine: Worksite Health Promotion*, 5(4): 725-738.

Department of Transportation. (1993, December 7). *Reporting of Drug Test Results: Abnormal Test Results and Analysis for Presence of Adulterants*. Washington, DC: Author.

Drake, C. (1993, September/October). "Why the resistance?: EAPs and the male paradox." *EAP Digest*: 32-41.

## General Review

*Drug Detection Report.* (1993a, April 5). Hair testing by an employer stopped for study in Maryland, 3(1): 1-2.

*Drug Detection Report.* (1993b, September 21). Hair testing for drug use makes gains in casino industry, 3(12): 4.

*Drug Detection Report.* (1993c, September 21). Denial of benefits upheld in Nevada after hair testing, 3(12): 4-5.

*Drug Detection Report.* (1993d, October 5). Florida study detects differences in accuracy of on-site tests, 3(13): 4.

*Drugs in the Workplace.* (1993, August). Saliva and hair testing: Pros and cons, 7(9): 1-2.

*Drugs in the Workplace.* (1994, April). EAPs and managed care: Liability ahead?, 8(4): 2-3.

*Drugs in the Workplace.* (1995, September). Shy bladder rule changes proposed, 9(9):5.

DuPont, R.L. (1991). *Drug-Caused Impairment in the Workplace.* Rockville, MD: Bensinger, DuPont Associates, Inc.

Durbin, N., Moore, C., Grant, T., Fleming, T., Hunt, P., Martin, P., Murphy, S., Hauth, J., Wilson, R., Bittner, A., Bramwell, A., Macaulay, J., Olson, J., Terrill, E. and Toquam, J. (1991). *Fitness for Duty in the Nuclear Power Industry: A Review of the First Year of Program Performance and an Update of the Technical Issues* (NUREG/CR-5784). Washington, DC: Nuclear Regulatory Commission.

Durbin, N., Murphy, S., Fleming, T., Macaulay, J., Westra, C., Olson, J. and Christenson, J. (1991). *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports CY 1990* (NUREG/CR-5758, Vol. 1). Washington, DC: Nuclear Regulatory Commission.

Enzymatics. (1990). Summary of technical and performance data on the Enzymatics A-150 and A-350 Saliva Alcohol Tests as presented to the U.S. Food and Drug Administration.

Evans, D.G. (1990). *Drug Testing Law, Technology and Practice.* New York, NY: Clark Boardman Callaghan.

Fehn, J. and Megges, G. (1985). "Detection of 6-Monoacetylmorphine in urine samples by GC/MS as

evidence for heroin use." *Journal of Analytical Toxicology* 9: 134-138.

Field, P.H. and Basteyns, B. (1988). The expert witness. In J.C. Garriot (Ed.), *Medicolegal Aspects of Alcohol Determination in Biological Specimens*: 275-292. Littleton, MA: PSG Publishing Co.

Fleming, T., Westra, C., Field, I., Forslund, C., Durbin, N., Grant, T. and Moffit, R. (1993). *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports CY 1992* (NUREG/CR-5758, Vol. 3). Washington, DC: Nuclear Regulatory Commission.

Gerner, C.J. and Kerns, L.L. (1994, February). "The liability risks of managed care." *Employee Assistance*: 30-32.

Gest, T. (1990, May 28). "Does he or doesn't he? New drug tests target hair." *U.S. News and World Report*.

Gilliland, K. and Schelegel, R.E. (1993). *Readiness to Perform Testing: A Critical Analysis of the Concept and Current Practices.* Washington, DC: DOT/FAA/AM-93.

Glasgow, R. and Terborg, J. (1988). Occupational health promotion programs to reduce cardiovascular risk. *Journal of Consult and Clinical Psychiatry*, 56(3): 365-373.

Gropper, B.A. (1995). Historical Overview: Research and Development on Drug Testing by Hair Analysis. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology.* U.S. Department of Health and Human Service, National Institutes of Health.

H.R. 33, Drug Testing Quality Act, 103rd Congress, 1st Session (1993).

Hall, L., Vacc, N. and Kissling, G. (1991). "Likelihood to use employee assistance programs: The effects of sociodemographic, social-psychological, sociocultural, organizational, and community factors." *Journal of Employment Counseling*, 28: 63-73.

Hirschberg, L. (1993, October). "The changing face of employee assistance. In EAPs: Not what they used to be." *Professional Counselor*: 36-41.

Interstate Brands (1994a). *Interstate Brands Corporation v. Local 441, Retail, Wholesale and Department Store*

*Union, U.S.D.C. for Northern Alabama, CV 93L-2694-S, March 31, 1994.*

Interstate Brands (1994b). *Interstate Brands Corporation v. Local 441, Retail, Wholesale and Department Store Union*, 39 F.3d 1159 (11th Cir, 1994).

Jackson, P.R., Tucker, G.T. and Woods, H.F. (1991). Backtracking booze with Bayes—the retrospective interpretation of blood alcohol data. *British Journal of Clinical Pharmacology*, 31: 55-63.

*Kansas v. Dodd*, 789 P.2d 1212, 1990 Kan. App. LEXIS 188, Court of Appeals of Kansas, March 23, 1990.

Katikaneni, L.D., Sallee, F.R., Ibrahim, H.O. and Hulsey, T.C. (1995). Neonatal Hair Analysis for Benzoylecgonine: A Marker for Gestational Cocaine Exposure, Growth at Birth, and Neurodevelopment Based on Short-Term Followup (4 to 8 Months). *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. U.S. Department of Health and Human Services, National Institutes of Health.

Kato, K., Hillsgrave, M., Weinhold, L., Gorelick, D.A., Darwin, W.D., and Cone, E.J. (1993). Cocaine and metabolite excretion in saliva under stimulated and nonstimulated conditions. *Journal of Analytical Toxicology*, 17: 338-341.

Kidwell, D.A., and Blank, D.L. (1995). Mechanisms of Incorporation of Drugs Into Hair and the Interpretation of Hair Analysis Data. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. U.S. Department of Health and Human Services, National Institutes of Health.

Klein, J. and Koren, G. (1995). Neonatal Hair Analysis: A Tool for the Assessment of In Utero Exposure to Drugs. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. U.S. Department of Health and Human Services, National Institutes of Health.

Lesar, J. (1992, January 30). "Companies try hair test to snare drug abusers." *International Herald Tribune: International Manager*.

*Louisiana v. Kinchen*, 342 So. 2nd 174, Supreme Court of Louisiana, January 24, 1977.

Major, M. (1990). "Employee assistance programs: An idea whose time has come." *Modern Office Technology*, 35(3): 76-81.

McLean, G.A., Wilcox, B.C. and Canfield, D.V. (1991, August). *Selection Criteria for Alcohol Detection Methods* (DOT/FAA/AM-91/12). Washington, DC: Office of Aviation Medicine, Federal Aviation Administration.

Mieczkowski, T. (1992). New approaches in drug testing: A review of hair analysis. *Annals of the American Academy of Political and Social Science*, 521: 132-150.

Mieczkowski, T., Barzelay, D., Gropper, B. and Wish, E. (1991). Concordance of three measures of cocaine use in an arrestee population: hair, urine, and self-report. *Journal of Psychoactive Drugs*, 23(3): 241-249.

Mieczkowski, T., Landress, H.J., Newel, R. and Coletti, S.D. (1993, January). "Testing hair for illicit drug use." *National Institute of Justice Research in Brief*

Miller, M.L., Donnelly, B. and Martz, R.M. (1995). Research at the Federal Bureau of Investigation in Hair Analysis for Drugs of Abuse. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. U.S. Department of Health and Human Services, National Institutes of Health.

Miller, S.M. (1993). "Saliva: New interest in a nontraditional specimen." *Medical Laboratory Observer*, 25(4), 31-35.

Modell, J.G., Taylor, J.P. and Lee, J.Y. (1993). Breath alcohol values following mouthwash use. *Journal of the American Medical Association (JAMA)*, 270(24): 2955-2956.

Moeller, M.R., Fey, P. and Sachs, H. (1995). Hair Analysis for Opiates in Forensic Cases. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. U.S. Department of Health and Human Services, National Institutes of Health.

Moore, C., Barnes, V., Hauth, J., Wilson, R., Fawcett-Long, J., Toquam, J., Baker, K., Wieringa, D., Olson, J. and Christensen, J. (1989). *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues* (NUREG/CR-5227, Supplement 1). Washington, DC: Nuclear Regulatory Commission.

## General Review

*MRO Newsletter.* (1993, August). Bloodborne pathogen standard, 2(8): 4.

*MRO Alert.* (1994, November). Preliminary Report of the AAMRO Technical Committee on Medical and Procedural Issues of "Shy Bladder," 5(9): 7-9.

Mulè, S.J. and Casella, G.A. (1988). Rendering the "poppy-seed defense" defenseless: Identification of 6-monoacetylmorphine in urine by gas chromatography/mass spectroscopy. *Clinical Chemistry*, 34(7): 1427-1430.

Murphy, S., Fleming, T., Westra, C., Field, I., Durbin, N. and Moffitt, R. (1992). *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports CY 1991* (NUREG 5758, Vol. 2). Washington, DC: Nuclear Regulatory Commission.

Naranjo, C.A. and Bremmer, K.E. (1993). Behavioural correlates of alcohol intoxication. *Addiction*, 88: 25-35.

National Highway Traffic Safety Administration. (1992, December). *Laboratory Testing of a Saliva-Alcohol Test Device by Enzymatics, Inc.* (DOT HS 807 893). Washington, DC: Author.

National Institute on Drug Abuse. (1988). *Medical Review Officer Manual: A Guide to Evaluating Urine Drug Analysis.* Rockville, MD: Author.

Neuhaus, C. and Caplan, R. (1992). *Employee Assistance Quarterly*, 7(4): 35-49.

Normand, J., Lempert, R.O., and O'Brien, C.P., Eds. (1994). *Under the Influence? Drugs and the American Work Force.* Committee on Drug Use in the Workplace. Washington, DC: National Academy Press.

NUMARC. (1991). Letter from NUMARC to NRC regarding current rule—fitness-for-duty programs. April 17, 1991.

Osterloh, J. and Becker, C. (1990). Chemical dependency and drug testing in the workplace. *Journal of Psychoactive Drugs*, 22(4): 407-417.

Paul, B.D., Mitchell, J.M. and Mell, L.D., Jr. (1988). Gas chromatography/electron impact mass fragmentometric determination of urinary 6-acetylmorphine, a metabolite of heroin. *Journal of Analytical Toxicology*, 13(2): 2-7.

Petracek, L. (1992). Workplace intervention. *Addiction and Recovery*, 12(4): 10-11.

PharmChem. (1993). *Building a Body of Evidence for Drug Testing Over Time.* Menlo Park, California: Author.

Prochaska, J. (1991). Assessing how people change. *Cancer*, 68: 805-807.

Psychemedics. (1992). *Detection of Drugs of Abuse by Analysis of Hair.* Chicago, Illinois: Author.

Quadrangle Research, LLC. (1995). *MRO Alert*, October/November 1995, 8:1-3.

Rollins, D.E., (1993). *On-Site Drug Testing in the Workplace.* Final Report to the NIDA Committee on On-Site Drug Testing in the Workplace.

Schechter, B. (1992). Drug-free workplace: It's more than testing. *EAP Digest*, 12(6): 36-42.

Schramm, W., Craig, P.A., Smith, R.H. and Berger, G.E. (1993). Cocaine and benzoylecgonine in saliva, serum, and urine. *Clinical Chemistry*, 39(3): 481-487.

Schramm, W., Smith, R.H. and Craig, P.A. (1992). Drugs of abuse in saliva: A review. *Journal of Analytical Toxicology*, 16:1-9.

Schuckit, M.A. (1991). Is there a protracted abstinence syndrome with drugs? *Drug Abuse and Alcoholism Newsletter*, 20(4): 1-3.

Schwartz, J.G., Zollars, P.R., Okorodudu, A.O., Carnahan, J.J., Wallace, J.E. and Briggs, J.E. (1991, March). Accuracy of common drug screen tests. *American Journal of Emergency Medicine*, 9(2): 166-170.

Schwartz, R.H., Clark, H.W. and Meek, P.S. (1993). Laboratory tests for rapid screening of drugs of abuse in the workplace: A review. *Journal of Addictive Diseases*, 12(2): 43-56.

*The Seattle Times.* (July 15, 1994). Job perks can come at a cost of less privacy: A3.

Seymour, R. and Smith, D. (1990). Identifying and responding to drug abuse in the workplace: An overview. *Journal of Psychoactive Drugs*, 22(4): 383-405.

Smith, G.D., Shaw, L.J., Maini, P.K., Ward, R.J., Peters, T.J. and Murray, J.D. (1993). Mathematical modelling of ethanol metabolism in normal subjects and chronic alcohol misusers. *Alcohol and Alcoholism*, 28(1): 25-32.

Society of Forensic Toxicologists. (1992). Revised consensus opinion on applicability of hair analysis for drugs of abuse. *ToxTalk*, 16(3): 2-11.

Stephenson, R.L. (1993). *Update on Hair Testing for Drugs of Abuse*. Presentation at the NII-BJA Fourth Annual Evaluating Drug Control Initiatives Conference, June 29, 1993.

Stetzer, E. (1993, October). "Are all the champions gone? In EAPs: Not what they used to be." *Professional Counselor*: 36-41.

Tompkins, C. (1991). Drug abuse among workers and employee assistance programs. *Drug Abuse Services Research Series: Background Papers on Drug Abuse Financing and Services Research*, no. 1. Rockville, MD: U.S. Department of Health and Human Services: 82-105.

Turk, M. (1993, September). High anxiety. *The New Physician*: 16-22.

U.S. Centers for Disease Control and Prevention. (1992, December 16). "Surveillance for occupationally acquired HIV infection—United States, 1981-1992." *Journal of the American Medical Association*. 268(23): 3294.

Washington State Drug Free Business. (1992a). Wellness programs coexisting with EAPs. *Business Taking Action*. Seattle, WA: Author.

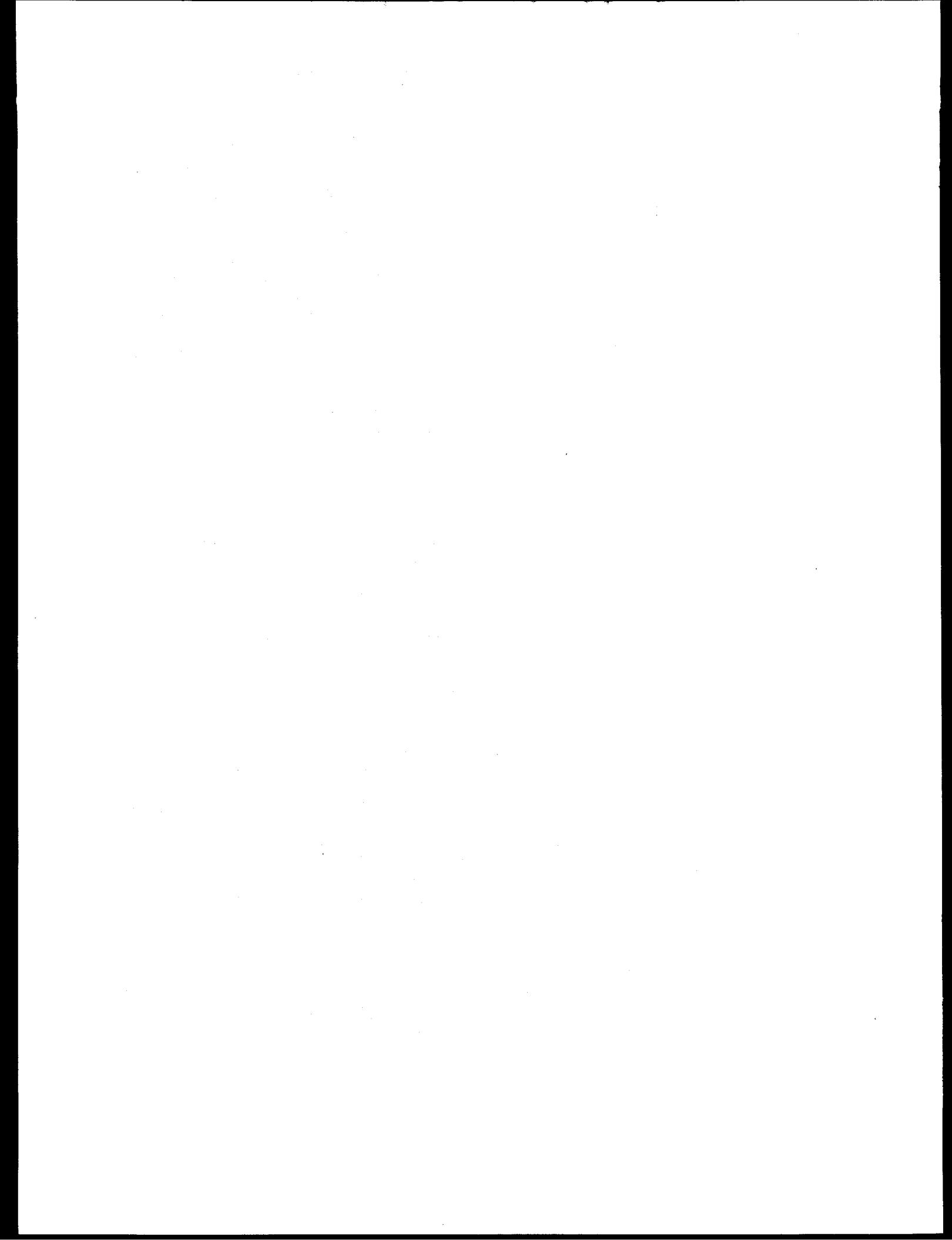
Washington State Drug Free Business. (1992b). EAPs entering managed care field to survive. *Business Taking Action*. Seattle, WA: Author.

Watkins, R.L. and Alder, E.V. (1993). The effect of food on alcohol absorption and elimination patterns. *Journal of Forensic Sciences, JFSCA*, 38(2): 285-291.

Welch, J.M. and Sniegoski, L.T. (1995). Interlaboratory Studies on the Analysis of Hair for Drugs of Abuse: Results From Three Exercises. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. U.S. Department of Health and Human Service, National Institutes of Health.

Westra, C., Forslund, C., Field, I., Gutierrez, J., Durbin, N., Grant, T. and Moffitt, R. (1994). *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports CY 1993* (NUREG/CR 5758, Vol. 4). Washington, DC: Nuclear Regulatory Commission.

*Woolley v. Hafner's Wagon Wheel*, 22 Ill.2d 413 (1961).



# CHAPTER 3: PREVALENCE OF DRUG USE IN THE UNITED STATES

Brent L. Baxter<sup>1</sup>

## 3.1: Introduction

The prevalence of illegal drug use and alcohol abuse among nuclear power industry workers is in part a reflection of such behavior in the wider society. It is therefore useful for nuclear power industry employers to have information on the current prevalence of—and emerging trends in—these behaviors among Americans in general, as well as specific populations of Americans, including American workers and nuclear power workers. This chapter provides a review of major findings on the prevalence of illegal drug use and alcohol abuse by addressing four primary questions:

1. What portion of Americans use illegal drugs or abuse alcohol?
2. What specific drugs are most often used/abused?
3. What are the characteristics (e.g., gender, age, employment status, occupation, occupational status, income) of Americans who use illegal drugs or abuse alcohol?
4. Which patterns of drug use prevalence are most relevant to fitness-for-duty programs in the nuclear industry?

These questions are explored across several contexts of drug use:

1. illegal drug use and alcohol abuse among Americans (Section 3.2);
2. self-reported illegal drug use and alcohol abuse by the American workforce (Section 3.3);
3. self-reported illegal drug use and alcohol abuse in the American workplace (Section 3.4);
4. illegal drug use and alcohol abuse detected in workplace drug testing (Section 3.5).

The chapter's final section (Section 3.6) contains conclusions regarding these findings and a discussion of the implications such findings may have for drug and alcohol testing in the nuclear power industry.

## 3.2: Prevalence of Illegal Drug Use and Alcohol Abuse Among Americans

This section highlights the major findings on the prevalence of illegal drug use and alcohol abuse among (1) Americans in general and (2) several special populations of Americans in which drug use is especially likely to be found: (a) young adults, (b) persons admitted to hospital emergency rooms for drug-related reasons, and (c) "street" users of drugs.

Drug use among Americans in general is explored through the findings of a widely-recognized annual self-report survey:

- National Household Survey on Drug Abuse (NHSDA).

Drug use among selected special populations of Americans is viewed through the findings of three diverse periodic studies:

- Monitoring the Future (MTF) surveys of college students and young adults; and
- Drug Abuse Warning Network (DAWN) reports of drug abuse-related emergency room episodes; and
- Pulse Check profile of drug use and availability according to field reports from drug ethnographers, epidemiologists, treatment providers, and police.

While each of these data sources has methodological limitations (see Appendix E), each undoubtedly contributes a useful perspective on the overall picture of drug use in America. As will be seen, findings from these sources present divergent pictures of recent trends in drug use in the United States. While survey data (from NHSDA) suggest that illegal drug use by Americans in general has decreased significantly in the past decade, the same data suggest a recent leveling off of that decline and modest increases in the use of several particular drugs (notably marijuana and LSD) by young adults. Meanwhile, qualitative studies of drug-related emergency room episodes (DAWN) and field reports

## Prevalence

(Pulse Check) present an alarming picture of dramatically increasing use of some drugs, particularly marijuana and LSD. In addition, while DAWN data suggest a recent increase in cocaine use, the other studies unanimously point to a significant drop in cocaine use. The summary elaborates on these interstudy differences in findings and suggests that they are primarily the result of differences in the methodologies and populations of interest of the respective studies.

### **3.2.1: The National Household Survey on Drug Abuse (NHSDA)**

NHSDA, an annual survey of a sample of the civilian, noninstitutionalized U.S. population aged 12 and older, is the most widely recognized survey of illegal drug use and alcohol abuse in the United States (SAMHSA\_NHSDA92, 1993; SAMHSA\_NHSDA92\_POPUL, 1993). The survey is administered in person in the homes of households selected on the basis of a multistage area probability sample, including oversampling of particular populations of interest (e.g., persons under age 35, blacks, Hispanics, and certain cities) to provide more extensive analysis of those populations. Over 26,000 persons were interviewed in the 1993 survey (SAMHSA\_NHSDA93, 1994). Conducted by the Office of Applied Studies within the Substance Abuse and Mental Health Services Administration (SAMHSA) (formerly NIDA), NHSDA is the most widely recognized survey of self-reported drug use in the United States and is the only drug use survey that is representative of the entire U.S. population (SAMHSA\_NHSDA91, 1993). Other available data sources tend to restrict their samples to specific populations, such as young persons (e.g., the various Monitoring the Future surveys), arrestees (e.g., the Drug Use Forecasting Program) or fatalities (DAWN data from medical examiner reports). For this reason, and despite the survey's limitations, this review relies heavily on NHSDA results to assess illegal drug use and alcohol abuse prevalence among the American public. (See Appendix E for further discussion of NHSDA's methodological strengths and weaknesses.)

NHSDA findings typically report use of illegal drugs and alcohol in the following drug categories:

- any illegal drug (in general);
- marijuana (including hashish);
- cocaine (including crack cocaine);
- inhalants;

- hallucinogens (including PCP);
- heroin;
- nonmedical use of any pharmaceutical;<sup>2</sup> and
- alcohol.

NHSDA presents each respondent's pattern of illegal drug use and alcohol abuse in terms of two general aspects: recency of use and frequency of use.

Recency of use is assessed by three measures:

- ever used ("lifetime use");
- used in past year ("annual use"); and
- used in past month (i.e., "current use").

Frequency of use, measured for marijuana, cocaine and alcohol, based on average use ("tendency") over the past year, is also assessed by three measures:

- used in the past year but less often than monthly (i.e., "occasional use");
- monthly use; and
- weekly use.

In addition, "heavy alcohol use," defined as "drinking five or more drinks per day on each of five or more days in the past 30 days," is measured.

NHSDA's 1993 results are summarized here and discussed in more detail in Sections 3.2.1.1 to 3.2.1.3.

#### **Illegal Drug Use: Summary of Findings**

- Illegal drug use has declined significantly over the past decade.
- The decade-long steady decline in illegal drug use may be leveling off.
- Illegal drug use within specific drug categories has declined or remained steady.
- The most prevalent illegal drug is marijuana, followed by nonmedical use of pharmaceuticals (especially analgesics) and then cocaine.

- A majority of current illegal drug users is using marijuana only.
- While cocaine use has declined substantially over the past decade, it still enjoys some popularity among younger Americans.
- Use of the most highly publicized illegal drugs—including heroin, PCP, amphetamines, and steroids—has remained relatively rare.

#### Alcohol Use: Summary of Findings

- While alcohol use in general has declined, heavy use has remained stable.

#### Differences Across Population Categories: Summary of Findings

- Illegal drug use and heavy alcohol use are more prevalent among males than females.
- Illegal drug use, as well as heavy alcohol use, is most prevalent among young adults.
- Illegal drug use across educational level is age-specific.
- College graduates are most likely to use alcohol, but least likely to abuse it.

- Use of both illegal drugs and alcohol is most prevalent in the West.

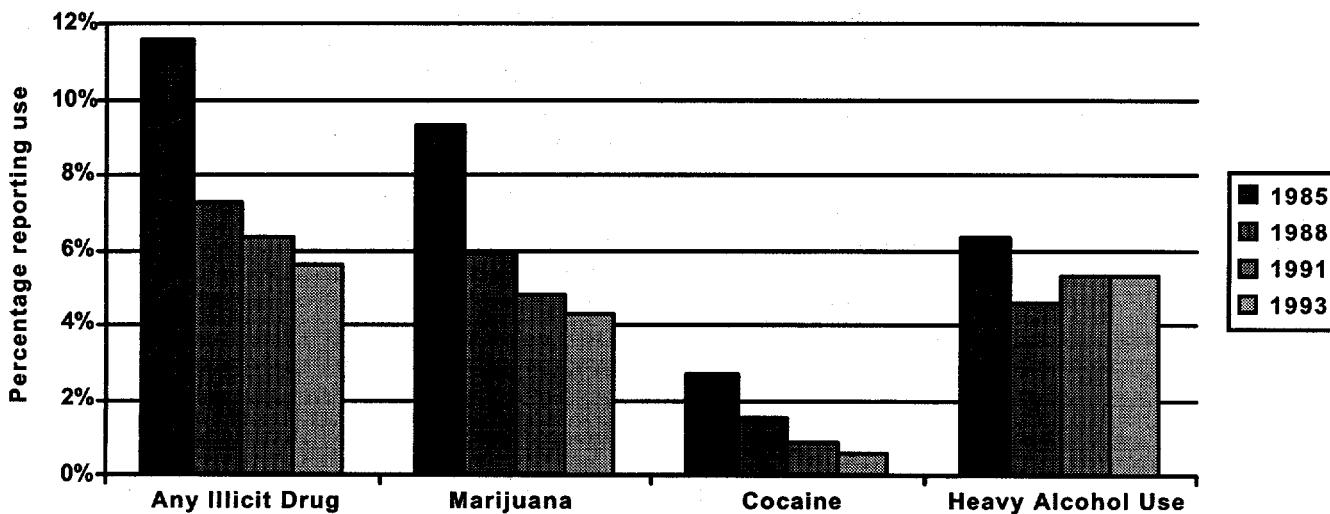
#### 3.2.1.1: Illegal Drug Use: Discussion of NHSDA Findings

- Illegal drug use has declined significantly over the past decade.

The percentage of Americans (aged 12 and older) reporting use of illegal drugs has shown a steady decline since 1985. For example, reported current (past-month) use of any illegal drug declined from 11.6% of Americans in 1985 to 5.6% in 1993 (see Figure 3.1) (SAMHSA\_NHSDA93, 1994). Past-year use has also declined significantly,<sup>3</sup> from 18.6% reporting in 1985 to 11.8% in 1993. In addition, reported overall use of illegal drugs other than marijuana has also declined significantly since 1985.

- The decade-long steady decline in illegal drug use may be leveling off.

The steady decline in overall illegal drug use over the past decade did not persist in 1993. While the percentage of Americans reporting illegal drug use in 1993 declined from 1992, the difference was statistically insignificant. The meaning of this cessation of decline is unclear: “The [1993] data may reflect a temporary interruption of the trend, a leveling off of prevalence, or the start of an



"Heavy Alcohol Use" is defined as drinking 5 or more drinks per day on each of 5 or more days.

Sources: SAMHSA; NHSDA, 1993, 1994.

Figure 3.1: Trends in Americans' Self-Reported Past-Month Drug Use, 1985-1993

## Prevalence

upturn in drug use" (SAMHSA\_NHSDA93, 1994, p. 2). It may in part reflect the persistence—despite anti-drug programs—of a core group of addicts, supplemented in each generation by a more transitional group of young experimental users.

- Illegal drug use within specific drug categories has declined or remained steady.

Without exception, none of the categories of illegal drug use surveyed by NHSDA showed a significant increase in reported recent use (past-year; past-month). This was true for recent (1992-1993) and longer-term (1985-1993) trends.

In addition, most specific categories of illegal drugs showed significant declines in reported current (past-month) use from rates in 1985. The percentages that reported current use in 1993 were: marijuana (4.3%); cocaine (0.6%); inhalants (0.4%); hallucinogens in general (0.2%); PCP (0.1%); nonmedical use of pharmaceuticals (1.3%). Notably, reported marijuana use has declined significantly since 1985 not only in terms of current (past-month) use (see Figure 3.1), but also in occasional use (4.6% in 1993), monthly use (4.3% in 1993), and weekly use (2.4% in 1993) (SAMHSA\_NHSDA93, 1994). Only lifetime use of marijuana has remained steady since 1985. Similarly, reported cocaine use has declined significantly since 1985 not only in terms of current (past-month) use (see Figure 3.1), but also in past-year use, occasional use and monthly use. Only lifetime use and weekly use (0.2% in 1993) of cocaine have remained steady since 1985 (SAMHSA\_NHSDA93, 1994).

- The most prevalent illegal drug is marijuana, followed by nonmedical use of pharmaceuticals and then cocaine.

Reported current (past-month) use of marijuana was 4.3% of Americans in 1993. The next most common illegal drugs were nonmedical use of pharmaceuticals in general (1.3% reporting current use), analgesics in specific (0.7%), and cocaine (0.6%). Likewise, the most common illegal drugs in reported past-year use were marijuana (9.0%), pharmaceuticals in general (3.8%), analgesics in specific (2.2%), and cocaine (2.2%) (SAMHSA\_NHSDA93, 1994).

- Almost half of current illegal drug users are using marijuana only.

Of those Americans reporting use of an illegal drug in 1993, 76% reported that such use included marijuana. In

fact, 48% of illegal drug users used marijuana only (SAMHSA\_NHSDA93, 1994).

- While cocaine use has declined substantially over the past decade, it still enjoys some popularity among younger Americans.

The percentage of Americans reporting past-year cocaine use fell from 5.9% in 1985 to 2.2% in 1993. However, past-year use among young adults (aged 18 to 25) was at 5.0%.

- Use of most highly-publicized illegal drugs—including heroin, PCP, amphetamines, and steroids—has remained relatively rare.

While considerable media attention has been focused on such dangerous drugs as heroin, PCP, amphetamines, and anabolic steroids, reported use of these drugs through 1993 was rare. Reported use of heroin and steroids, whether past-year or current (past-month), was rare in 1993 and has remained stable since 1985. While use of hallucinogens (especially LSD) may be rising among youth (see MTF findings in Section 3.2.2.1), their use (especially of PCP) is rare in the general population. Reported nonmedical use of prescription-type stimulants and anabolic steroids is also rare. Only 0.1% of Americans reported using heroin in 1993, 1.2% using hallucinogens (0.2% using PCP), 1.1% using prescription-type stimulants for nonmedical purposes, and 0.1% using steroids (SAMHSA\_NHSDA93, 1994).

### 3.2.1.2: Alcohol Use: Discussion of NHSDA Findings

- While alcohol use in general has declined, heavy use has remained stable.

By nearly every NHSDA measure (ever used, used in past year, used in past month, monthly use, weekly use), use of alcohol by Americans has decreased significantly since 1985. For example, 49.6% of Americans reported current (past-month) use in 1993, down from 58.3% in 1985; those tending to use alcohol at least weekly decreased from 27.8% in 1985 to 21.5% in 1993 (SAMHSA\_NHSDA93, 1994).

But heavy alcohol use has not decreased. In only two categories did alcohol use not decrease: occasional use, which remained stable (25.0% reported in 1985 and 26.6% in 1993) and, more notably, current (past-month) heavy use, which showed a slight, but statistically insignificant, decrease from 6.3% in 1985 to 5.3% in

1993 (see Figure 3.1) and which has shown no change since 1991 (5.3%) (SAMHSA\_NHSDA93, 1994).

### 3.2.1.3: Differences Across Population Categories: Discussion of NHSDA Findings

The prevalence of illegal drug use and alcohol use often varies across population categories. Some notable examples from 1993 are as follows (SAMHSA\_NHSDA93, 1994):

- Illegal drug use and heavy alcohol use are more prevalent among males than females.

In 1993, 7.4% of males reported current (past-month) illegal drug use, compared with 4.1% of females. Current use among males was more prevalent than female use in every age group.

Gender differences are even more dramatic for heavy alcohol use: 9.5% of males reported current (past-month) heavy alcohol use in 1993, compared with only 1.5% of females. Current heavy use among males was more prevalent than female use in every age group.

- Illegal drug use, as well as heavy alcohol use, is most prevalent among young adults.

Persons aged 18-25 were most likely to report current (past-month) illegal drug use and heavy alcohol use. The percentage reporting current use of any illegal drug in the age categories 12-17, 18-25, 26-34, and 35 and older were, respectively, 6.6%, 13.5%, 8.5%, and 2.8%. The corresponding figures for heavy alcohol use were 1.3%, 10.4%, 7.3%, and 4.2%.

- Illegal drug use across educational level is age-specific.<sup>4</sup>

The relationship between educational status and current (past-month) illegal drug use is highly dependent on age. Current use of any illegal drug by Americans in general (age 18 and older), and by older Americans (age 35 and older) was most prevalent among persons with some college education (but less than a degree), while current use among younger Americans (age groups 18-25; 26-34) was most prevalent among those with less than a high school education.

- College graduates are most likely to use alcohol, but least likely to abuse it.

Current (past-month) alcohol use in 1993 was most prevalent among college graduates and least prevalent among those without a high school degree. Among all Americans age 18 and older, 63.9% of college graduates reported use, compared to 60.9% of those with some college education, 51.6% of high school graduates, and 35.7% of those without a high school degree. This difference held across all age groups (18-25; 26-34; 35 and older).

In contrast, current (past-month) heavy alcohol use in 1993 was least prevalent among college graduates. Among all Americans age 18 and older, 3.9% of college graduates reported heavy use, compared to 7.2% of those with some college education, 5.7% of high school graduates, and 6.5% of those without a high school degree. This pattern held for the two younger age groups (18-25; 26-34).

- Use of both illegal drugs and alcohol is most prevalent in the West.<sup>5</sup>

Current (past-month) illegal drug use, as well as current alcohol use, was most prevalent in the West. Current illegal drug use was least prevalent in the North Central region, while alcohol use was least prevalent in the South.

### 3.2.2: The Monitoring the Future (MTF) Survey of College Students and Young Adults

The annual NIDA-sponsored MTF survey focuses on two of the specific populations in which drug use—especially illegal drug use—is most prevalent: college students and their noncollege peers. Although MTF results may be less representative of drug use patterns in the American population, they provide a detailed examination of drug use prevalence and trends among two population groups whose members are 1) more likely to use drugs and 2) likely to enter, or already be in, the full-time workforce. In addition, the higher rates of illegal drug use found in such groups often allow assessment of trends in the use of less popular drugs, such as heroin and steroids.

MTF surveys are administered by mail to respondents sampled from previous Monitoring the Future surveys of high school seniors (originally selected in nationally representative samples) who subsequently graduated. The surveys contain questions on a wide variety of issues pertaining to lifestyle and values, including drug use.

## Prevalence

### Summary and Discussion of MTF Findings (Compared with NHSDA Findings)

The most recent MTF findings are based on a sample of over 8,200 college students and young adults in 1992 (Johnston, O'Malley and Bachman, 1993). A comparison of MTF findings with those of NHSDA not only provides a comparison of prevalence rates among younger adults and the general population, it also provides a test of the validity of the two surveys by allowing a comparison of prevalence rates among similar populations (young adults) common to both surveys. As indicated in Table 3.1, results from the two surveys indicate remarkably similar prevalence rates (across illicit drugs in general and specific categories of drugs) for comparable population groups: Americans aged 18-25 (from NHSDA) and college students and their noncollege peers (from MTF).

Comparing prevalence rates among young adults with those of the general population, the general results are as expected. As indicated in Table 3.2, prevalence rates for most types of drug use are higher among young adults than among the general population. However, similar patterns of drug use were reported across the two groups. As in the general population, marijuana was clearly the most popular illegal drug used by college students. Furthermore, some well-publicized drugs—e.g., heroin—have notably small prevalence rates among both groups.

Trends in the usage of most types of drugs are also similar across the two groups. As indicated in Table 3.2, since 1985, annual use of illicit drugs, including marijuana, cocaine, stimulants, tranquilizers, and alcohol, is down among both the general population and college students. The prevalence of occasional heavy alcohol use (defined as "having five or more drinks in a row in the

last two weeks") among college students has decreased only slightly over the past decade (from 44.6% in 1985 to 41.4% in 1992), a pattern of stability similar to that of the more rigidly-defined "heavy alcohol use" among the general population, as reflected in NHSDA figures.

There are also a few notable intergroup differences in drug use trends. While reported use of hallucinogens has decreased among the general population over the past decade, use of hallucinogens—especially LSD—has been increasing among college students (and other young adults) for several years (Johnston, O'Malley and Bachman, 1993).

Analysis of the entire series of annual prevalence rates among college students since 1985 (not shown) substantiates two trends: first, while cocaine use among college students has plummeted steadily since 1985, LSD use has risen fairly steadily. Second, the decade-long decline in marijuana use among college students appeared to reverse itself in 1992, with a slight increase in use (from 26.5% in 1991 to 27.7% in 1993).

### 3.2.3: The Drug Abuse Warning Network (DAWN) Data System of Drug-Related Emergency Room Episodes

DAWN, sponsored by NIDA, presents annual compilations of information from case reports of drug-related emergency room (ER) visits by persons (aged seven or older) visiting one of a sample of nonfederal hospitals in over 20 metropolitan areas in the U.S. and a national sample of American hospitals outside these areas. To be included in DAWN data, ER episodes must contain all of the following conditions (SAMHSA\_DAWN\_93, 1994, p. 5):

**Table 3.1: Annual (Past Year) Prevalence Rates, for Selected Drug Categories, Based on Self-Reported Results**

	Any Illicit	Marijuana <sup>a</sup>	Cocaine	Hallucins.	Stimulants <sup>b</sup>	Inhalants	Alcohol
Ages 18-25 <sup>c</sup>	27%	23%	5%	5%	3%	3%	79%
College students <sup>d</sup>	31	28	3	7	4	3	87
Young adults <sup>d</sup>	29	26	6	7	6	4	83

a. Includes hashish.

b. Nonprescribed use only.

c. NHSDA 1993 findings (SAMHSA\_NHSDA93, 1994).

d. MTF 1992 findings (Johnston, O'Malley and Bachman, 1993); "college students" = full-time college students 1-4 years beyond high school; "young adults" = other high school graduates 1-4 years beyond high school.

**Table 3.2: Past-Year Use of Illicit Drugs and Alcohol among the General Population and College Students in Particular**

Drug Categories	Annual Use			
	General Population <sup>a</sup> (age 12+)		Full-Time College Students <sup>b</sup> (1 to 4 years past high school)	
	1985	1993	1985	1992
Any Illicit	18.6%	11.8%	46.3%	30.6%
Marijuana	14.8	9.0	41.7	27.7
Hallucinogens	1.6	1.2	5.0	6.8
LSD	NA	NA	2.2	5.7
PCP	0.5	0.2	NA	NA
Cocaine	5.9	2.2	17.3	3.0
MDMA ("Ecstacy")	NA	NA	NA	2.0
Heroin	0.2	0.1	0.2	0.1
Stimulants	3.9	1.1	11.9	3.6
Crystal methamphetamine ("Ice")	NA	NA	NA	0.2
Tranquilizers	3.4	1.2	3.6	2.9
Anabolic Steroids	NA	0.1	NA	NA
Inhalants	1.5	1.0	3.1	3.1
Alcohol	72.9	66.5	92.0	86.9

NA Not available.

a. Source: SAMHSA\_NHSDA93, 1994.

b. Source: Johnston, O'Malley and Bachman, 1993.

- the patient was treated in the hospital's emergency department;
- the patient's presenting problem(s) was induced by or related to drug use, regardless of when the drug consumption occurred;
- the case involved the nonmedical use of a legal drug or any use of an illegal drug;
- the patient's reason for taking the substance(s) included one of the following: (1) dependence, (2) suicide attempt or gesture, or (3) psychic effects.

In interpreting DAWN findings, it is critical to note that not all the drugs "mentioned" (self-reported or detected) in a particular ER episode are responsible for the medical emergency.

In contrast to self-report surveys on varying types of drug and alcohol use, DAWN data on drug-related hospital ER visits provide a rough assessment of drug abuse that has produced serious medical consequences, thereby serving as an indicator of particularly dangerous forms of drug abuse. Only those episodes judged by the attending medical staff to be caused by the abuse of drugs (including prescription and OTC drugs) are included in DAWN statistics. (It should be noted that alcohol is reported only when present in combination with other drugs, thus the category "alcohol-in-combination.") While DAWN findings do not present an accurate picture of the overall level of drug use, nor even the distribution of the types of drugs used in American society, they may suggest trends in drug use, particularly excessive drug abuse. (See Appendix E for further discussion of DAWN's methodological strengths and weaknesses.)

## Prevalence

### Summary and Discussion of DAWN Findings

The most recent full-year DAWN findings are based on preliminary figures from over 466,000 drug-related ER episodes sampled in 1993 (SAMHSA\_DAWN\_93, 1994), representing the highest levels ever recorded in DAWN data. The five most common drugs in drug-related ER episodes in 1993 were, in order:

- alcohol-in-combination (mentioned in 31% of all episodes);
- cocaine (26%);
- heroin/morphine (13%);
- acetaminophen (7%); and
- marijuana/hashish (6%).

This pattern in the top four types of drugs (same drugs, same order) has persisted since 1988. Marijuana was the sixth most commonly mentioned drug (aspirin was fifth) from 1988 through 1991, rising to fifth place in 1992 when the number of mentions of marijuana jumped sharply.

Reports of trends in DAWN findings usually focus on three common illegal drugs: cocaine, heroin and marijuana. Data from recent years show an overall increase in the number of mentions of each of these drugs. According to 1993 DAWN data, mentions of heroin and marijuana increased dramatically (each increased 86%) since 1990, while mentions of cocaine increased 53% over the same time period. (Meanwhile, mentions of LSD have decreased slightly [down 13%] in the same time period.)

Recent increases in mentions of heroin are particularly notable in cases where the drug was inhaled: from 1988 to 1993, heroin-related episodes resulting from inhalation increased 470% while heroin-related episodes resulting from injection increased only 31%.

Nevertheless, the largest increases in drug-related ER episodes in recent years have taken place in the mentions of particular prescription and OTC drugs (often associated with suicide attempts), in particular the benzodiazepine alprazolam (Xanax®) (up 1,340% since 1982), the anticonvulsants clonazepam (up 644% since 1988) and carbamazepine (up 146% since 1988) and the antidepressant trazodone (up 167% from 1988).

### 3.2.4: The Pulse Check Profile of National Trends in Drug Abuse

The Pulse Check study, conducted by Abt Associates, Inc., and commissioned by the Office of National Drug Control Policy (ONDCP), presents a quarterly report of findings based on periodic telephone contacts with drug ethnographers, epidemiologists, treatment providers, police, and college counselors, primarily in urban areas—"persons around the country who are 'on the front lines' of the drug problem" (ONDCP, 1994, p. 1).

The sample of drug ethnographers and epidemiologists was selected based on each individual's ethnographic experience with drug users or access to ethnographic data. Narcotics police were selected based on recommendations by drug researchers and other contacts in the field. Each quarter, roughly the same group of individuals is contacted to provide comparable data over time. The sample of treatment providers was randomly selected from the files of the National Drug Abuse Treatment Unit Survey. The specific providers included in the sample vary from quarter to quarter.

Data collection is done through loosely-structured phone conversations with the field contacts. Typical topics discussed include drug-specific trends in drug use prevalence, typical characteristics of users/sellers, common methods of drug administration, purchase prices, and levels of drug purity. Unlike more quantitatively-based self-report surveys such as NHSDA, the loosely-structured phone conversations employed in Pulse Check allow respondents (field contacts) to elaborate on their answers and to discuss drug use patterns not anticipated by the researchers. This data source is therefore a useful supplement to other survey data. In the author's own words:

[Pulse Check] provides a current, subjective profile of drug use and availability in contrast to the more delayed, objective profile provided by surveys. The Pulse Check is intended to complement, not replace, traditional data sources (ONDCP, 1994, p. 1).

Pulse Check findings are presented primarily in descriptive summary, supplemented by tables briefly summarizing field reports from specific urban areas around the country. Data provided by treatment providers are summarized by region.

### Summary and Discussion of Pulse Check Findings

Pulse Check data gathered in November, 1994 indicate that cocaine use has leveled off in most areas of the country although it remains the dominant drug in clients presenting for treatment and in law enforcement incidents. While cocaine use has stabilized, the demand for some other drugs, particularly marijuana, hallucinogens, and amphetamines is on the rise.

Marijuana is increasingly popular and widely available in many areas of the country. Ethnographic sources and police contacts report that the number of youth who are using marijuana has increased the most rapidly, although users of all ages are reported. Many providers have also noted that marijuana users entering treatment "just keep getting younger" (ONDCP, 1994, p. 11).

Hallucinogens have been frequently mentioned by these sources as an emerging drug. The very rapid increases reported in early 1994 appear to have leveled off, but interest in naturally occurring hallucinogens such as psilocybin ("mushrooms"), jimsonweed, and a "toad" extract, MDMA, and "club drugs," a term used to describe a variety of legal and illegal substances used at parties or clubs, are also mentioned by contacts. MDMA is primarily popular among young, upper-middle class whites and the gay community in a few West Coast urban areas and in Texas (ONDCP, 1994, p. 15). Universities reported that hallucinogen use, especially LSD, is on the rise but that alcohol continues to be the primary drug problem, followed by marijuana.

Amphetamine and methamphetamine use is emerging in the Middle Atlantic States according to several sources. Methamphetamines are being ingested in pill form, snorted, or injected. In Denver the practice of injecting methamphetamine alone or with cocaine is described as popular, particularly among white young adults (ONDCP, 1994, p. 15).

Heroin use has been expanding into new population groups—young adults, females, and middle-income users. Some sources believe that this is a sign of a new epidemic emerging; others are not clear if it is a significant expansion or a replacement cohort (i.e., new users replacing previous users who either quit using or are no longer alive).

Finally, alcohol continues to be a legal drug nearly unanimously used by illicit drug users. It continues to be the primary problem drug on college campuses and it is

used by the vast majority of those clients presenting for treatment.

While aggregate levels of demand for treatment reflect a number of factors in addition to drug use prevalence, Pulse Check results are an important source of data on regional variation in abuse. Compared to 1993, demand for cocaine treatment increased slightly in all four NRC regions, ranging from a low of 17% of providers reporting an increase in Region IV (West) to a high of 31% of providers reporting an increase in Region II (South). Only in Region IV (West) was there a substantial number of providers (25%) reporting an increase in the demand for marijuana treatment; most reported no change. The demand for heroin treatment increased in all four regions although the number of clients whose primary drug of abuse is heroin remains small, ranging from 7% in Region III (Midwest) to 23% in Region I (Northeast). In Region IV (West), 50% of treatment programs reported an increase in the number of heroin treatment clients. In Regions I and II (Northeast and South), over 40% of programs reported a similar increase. The percentage of programs reporting an increase in Region III (Midwest), the lowest of the four regions, was still nearly 10%.

### 3.2.5: Conclusions Regarding the Prevalence of Drug Use Among Americans

Before drawing conclusions from the variety of findings presented above, it is important to explore the possible reasons for the interstudy differences in findings, particularly differences between the NHSDA findings and those of the more qualitative studies, DAWN and Pulse Check.

#### 3.2.5.1: Examination of Interstudy Differences in Findings

Self-report survey results (NHSDA, MTF) present a different picture from those of more subjective sources (DAWN, Pulse Check). While self-report survey data show a leveling off of a decades-old decline in drug use, the qualitative studies present a more alarming picture of increasing use of illicit drugs, particularly marijuana, heroin and hallucinogens. While these differences in findings may seem to reflect a conflict in findings, it should be noted that these studies differ in several aspects, most notably their study population and methodology. It is likely that the divergence in findings is due primarily to such study characteristics. To the extent this is true, each study's findings can be used to

## Prevalence

complement, rather than negate, those of the other studies.

The primary differences in the studies are:

**Sample Population:** The surveys examine a much wider range of Americans than the qualitative profiles. The NHSDA surveys a nationwide random sample of all Americans (as we noted previously, such a sample probably underrepresents populations with a disproportionately high number of hard-core drug users), while the MTF surveys college students and their same-age cohorts. The qualitative profiles are based on much more restricted samples. DAWN reports examine persons entering hospitals for drug-related emergencies, while Pulse Check reports are based on subjective, relatively unstructured interviews of various field contacts around the country. Thus, while the self-report surveys probably underemphasize hard-core drug use, the qualitative studies are expressly designed to highlight it.

A recent National Research Council report on drug abuse in America describes this distinction as presenting "two worlds" of drug abuse (Gerstein and Green, 1993, pp. 35-36):

In one [world], measured by survey data on individuals in school classrooms and households, illicit drug use is not confined to or even particularly prominent in any one social class, economic stratum, race, or ethnic group . . . . This world of low-intensity consumption shows steady and cumulatively very marked declines in the prevalence of marijuana use since the late 1970s and of cocaine since the middle 1980s; heroin use is so rare as to be barely measurable.

The other world is that of emergency rooms, morgues, drug clinics, juvenile detention center, jails, and prisons, in which indicators of intensive drug consumption . . . are collected . . . [where] the poor predominate . . . marijuana and heroin use are common . . . and cocaine use increased explosively throughout the 1980s and simply leveled off at high levels in the 1990s.

In assessing the needs of a workplace drug testing program, we must be aware that the former of these worlds—that of low-intensity or occasional use—is the predominant one influencing the American workplace, including the nuclear power industry. Below we discuss how three key differences in the methods, the source of data, the precision of the analysis, and the accuracy of the

results affect the conclusions one would reach from these studies.

**Source of data:** The surveys rely on the self-reports of subjects; while this is generally a scientifically-accepted method of data collection (and is administered expertly within these surveys), it justifiably raises some concerns about the validity of the data. In contrast, the qualitative studies use information gathered from sources (ER attendants and authorities on street drug use) that are largely independent of the potential drug users being studied (DAWN data are partially based on self-reports from the ER patients).

**Precision of data gathering/analysis:** The surveys employ questionnaires designed to cover a wide variety of drug types and modes of drug use and to allow extensive and objective analysis of the resulting quantitative data. In contrast, the qualitative studies—DAWN and Pulse Check—are much more subjective in nature, producing diverse and sometimes anecdotal data through relatively unsystematic methods. (Conversely, such methods allow respondents to elaborate on their comments and to explore unanticipated areas of information.)

**Currency of results:** Given the size of the nationwide self-report surveys, their results typically do not appear until at least nine months after the survey is administered. Results from the qualitative reports appear much faster, typically within a few months after data collection.

Some experts are pointing to DAWN findings as evidence that while casual use of illegal drugs has continued to decline, the number of ER visits by chronic users of cocaine, heroin and marijuana has remained stable or increased (Treaster, 1993). According to Harvard University Professor Mark Kleiman, the DAWN findings suggest that "heroin is coming back in a serious way and that marijuana may also be coming back" (Treaster, 1993). However, since NHSDA data suggest that marijuana use is declining and heroin use is stable, increases in ER episodes involving these drugs may be a reflection not of an increase in the number of users but of such factors as the strong popularity of marijuana among illegal drug users and new, more potent strains of heroin (e.g., "China White").

### 3.2.5.2: Conclusions Regarding Study Findings

An appreciation of the differences in focus and methodology among the different studies allows us to suggest several conclusions regarding the study findings:

- The decade-long decline in drug use among the general population may be leveling off.
- Use of some illegal drugs, particularly marijuana and LSD, appear to be increasing among younger Americans.
- While trends in marijuana use vary across age groups, it clearly remains the most popular illegal drug.
- Cocaine use, though declining among the general population, is still fairly prevalent among younger Americans.
- The use of other illegal drugs appears to be less pervasive than is implied by media attention.
- Alcohol abuse is prevalent enough to cause concern.
- While casual use of illegal drugs and alcohol has declined, it appears that a core group of alcoholics and addicts remains.

These conclusions are discussed below:

- The decade-long decline in drug use among the general population may be leveling off.

The steady decline in overall illegal drug use found in NHSDA findings over the past decade did not persist in 1993. This may in part reflect the persistence—despite anti-drug programs—of a core group of addicts, supplemented in each generation by a more transitional group of young experimental users.

- Use of some illegal drugs, particularly marijuana and LSD, appears to be increasing among younger Americans.

MTF findings reveal that among college students, the decade-long decline in marijuana use appears to be reversing itself, while LSD use has risen fairly steadily since 1985. Pulse Check findings suggest that marijuana use is increasing dramatically (“exploding”) among young Americans, while use of hallucinogens and designer drugs—including synthetic hallucinogens—is also increasing among white urban youth.

- While trends in marijuana use vary across age groups, it clearly remains the most popular illegal drug.

While NHSDA results indicate a steady decline in marijuana use among the American public, MTF reports a recent increase in its use among young adults.

Meanwhile, according to NHSDA results, 76% of persons reporting illegal drug use in 1993 used marijuana with 48% using marijuana only (calculated from SAMHSA-NHSDA93, 1994: Table 6A). DAWN data list marijuana as the fifth most commonly mentioned drug in drug abuse-related ER visits. These findings indicate that marijuana is very popular among illegal drug users.

- Cocaine use, though declining among the general population, is still fairly prevalent among younger Americans.

While reported past-month use of cocaine has declined steadily since the mid 1980s, from 2.7% in 1985 to 0.6% in 1993, a considerable portion (5.0%) of Americans aged 18-25 reported using cocaine in 1993. In addition, cocaine was mentioned in over a quarter of DAWN episodes in 1992, remaining the most common illegal drug in drug abuse-related ER visits.

- The use of other illegal drugs appears to be less pervasive than is implied by media attention.

Some drugs that receive considerable media attention, and may in fact be causing considerable problems in some segments of society, appear to be much less prevalent among Americans overall. These drugs include the opiates, hallucinogens (especially PCP), stimulants and steroids. While there is evidence of increasing use of hallucinogens (especially LSD) among young persons, the rate of hallucinogen use among the general population remains low. Finally, while heroin/morphine is mentioned in 13% of DAWN episodes in 1993, and is cited in Pulse Check field reports as an emerging drug, the prevalence of heroin use in the overall population is undoubtedly much lower (NHSDA indicates only 0.1% of Americans reported using heroin in 1993).

- Alcohol abuse is prevalent enough to cause concern.

While reported use of alcohol is down among both the general population and young adults, heavy alcohol use is stable among both populations. In fact, the percentage of Americans reporting current (past-month) heavy alcohol use in 1993 (5.3%) was only slightly less than the percentage reporting current use of any illegal drug (5.6).<sup>6</sup> In addition, DAWN data consistently indicate that despite omitting alcohol-only ER episodes, alcohol (identified as “alcohol-in-combination” with other drugs) is still the most commonly mentioned drug in drug abuse-

## Prevalence

related ER visits. These data suggest that a small but significant percentage (estimated at between 5% and 10%) of the American population is at any one time using illegal drugs or abusing alcohol.

- While occasional use of illegal drugs and alcohol has declined, it appears that a core group of alcoholics and addicts remains.

While occasional use of alcohol has declined among Americans, heavier use of alcohol—often symptomatic of alcoholism—has remained steady, apparently persisting among a core group of users. Likewise, while NHSDA results suggest a decade-long trend toward lower prevalence of illegal drugs among the general population, NHSDA, DAWN and Pulse Check findings suggest the persistence of a fairly stable and relatively small group of heavy users.

### **3.3: Prevalence of Illegal Drug Use and Alcohol Abuse in the American Workforce**

While reported use of both illegal drugs and alcohol by Americans has been decreasing, illegal drug use and alcohol abuse still pose considerable problems in American society, including work-related fitness-for-duty concerns. Sections 3.3 and 3.4 present estimates of illegal drug use and alcohol use and abuse prevalence among American workers from two general perspectives:

1. illegal drug use and alcohol abuse by the workforce and, more specifically,
2. drug and alcohol use in the workplace.

Illegal drug use and alcohol abuse in the workforce will be assessed on the basis of the survey results of self-reported illegal drug use across employment status and industry and self-reported/diagnosed alcohol problems across occupations. Illegal drug use and alcohol use in the workplace will be assessed on the basis of survey results of self-reported on-the-job drug and alcohol use across occupations. A final section (Section 3.5) presents data on drug use and alcohol use and abuse in both the workforce and in the workplace based on drug test results compiled across industry. While drug detection rates and drug use prevalence are clearly not equivalent, these data do provide an interesting comparison of nuclear power industry test results with those of other industries.

#### **3.3.1: Workforce Illegal Drug Use and Heavy Alcohol Use Prevalence Across Employment Status**

National Household Survey on Drug Abuse (NHSDA) findings across employment status suggest that illegal drug use is considerably more prevalent among the unemployed, while differences in heavy alcohol use are smaller and less consistent (see Table 3.3).<sup>7</sup> For example, in 1992, 13.8% of unemployed persons reported current (past-month) illegal drug use, compared to 6.2% of part-time workers and 5.8% of full-time workers (SAMHSA\_NHSDA92, 1993). Similar patterns in illegal drug use occurred in every age category in each year examined. In contrast, patterns in current (past-month) heavy alcohol use varied depending on the age group and year examined. The only apparent trend is toward higher prevalence of heavy use of alcohol among the unemployed (compared to employed workers).

#### **3.3.2: Workforce Illegal Drug Use and Heavy Alcohol Use Prevalence Across Income**

Among full-time workers, there is a relationship between income and reported drug use, especially among males. As shown in Table 3.4, reported past-month use of any illegal drug, as well as past-month heavy alcohol use, tends to be more prevalent in the lower-income brackets and is progressively less prevalent (at least among males) as income rises (Kopstein and Gfroerer, 1990; Normand, Lempert and O'Brien, 1994).

#### **3.3.3: Workforce Illegal Drug Use and Heavy Alcohol Use Prevalence Across Industry**

As shown in Table 3.5, “current” (past-month) illegal drug use—as well as marijuana use and cocaine use measured separately—among full-time male employees was most prevalent in the construction and finance industries. Heavy use of alcohol was also prevalent in construction, followed by retail trade and repair services. Considerable variability across type of drug was found in the transportation, wholesale trade, finance and professional industries, with considerably higher rates of illegal drug use than of heavy alcohol use. (While corresponding data on female workers in some industries were not reliable, illegal drug use by full-time female

**Table 3.3: Prevalence Rates of Self-Reported Drug Use, for Selected Drug Categories, by Employment Status and Age: 1988, 1990, 1992**

Employment status	Past-Month Use of Any Illegal Drug <sup>a</sup>			Past-Month Heavy Use of Alcohol <sup>b</sup>		
	1988	1990	1992	1988	1990	1992
Full-time employed	8.2%	7.0%	5.8%	6.0%	6.8%	6.0%
18-25 years	18.7	14.2	12.2	10.5	13.9	10.6
26-34 years	12.9	8.5	9.8	6.8	8.4	7.7
35+ years	2.4	4.1	2.6	4.0	4.0	4.2
Part-time employed	8.9	8.0	6.2	5.2	5.1	6.2
18-25 years	16.7	15.9	11.7	10.0	9.2	12.6
26-34 years	13.0	14.2	10.1	4.7	5.2	5.5
35+ years	3.6	2.0	1.6	3.1	3.1	2.8
Unemployed <sup>c</sup>	17.4	14.0	13.8	6.4	6.6	11.1
18-25 years	28.2	19.3	23.3	9.0	8.8	14.8
26-34 years	24.8	22.8	18.1	8.2	8.0	14.6
35+ years	4.8	5.7	5.8	*	*	6.9

a. Includes use of marijuana, hashish, inhalants, hallucinogens, cocaine, heroin, and nonmedical use of pharmaceuticals.

b. Having 5 or more drinks on the same occasion 5 or more times in the past 30 days.

c. Excludes persons retired or disabled, homemakers and students.

\* Low precision; no estimate reported.

Source: SAMHSA\_NHSDA92, 1993: Table 16B.

workers was relatively high in manufacturing, retail trade, and professional industries. In fact, in the retail trade industry, overall illegal drug use and use of cocaine were more prevalent among females than males.)

### 3.3.4: Workforce Alcohol Problems Across Occupation

Parker and Harford (1992) analyzed the 1988 National Health Interview Survey (NHIS) to provide estimates of the prevalence of self-reported alcohol use and dependence among workers in various occupations (after adjusting statistically for age differences across occupation). In general, they found that workers in blue-collar occupations were less likely to use alcohol but those that did use alcohol tended to drink larger amounts and to display symptoms of alcohol dependence more often than white-collar drinkers.

Stinson, DeBakey and Steffens (1992) used the same 1988 NHIS data to assess the prevalence of alcohol abuse

and/or dependence across a larger set of more specific occupations. The occupation with the highest rate of alcohol abuse/dependence was bartenders (42%). The industries with the highest rates of alcohol abuse and/or dependence were the forestry and fishing occupations (22%) and the construction and extractive (e.g., mining) trades (21%), with the latter including high rates (20% or greater) among brickmasons and stonemasons (37%); carpenters (26%); plumbers, pipefitters, and steamfitters (21%); and roofers (25%). Other specific occupations with high rates of alcohol abuse/dependence (20% or greater) were operations and systems researchers and analysts; editors and reporters; securities and financial services sales occupations; mail clerks, excluding postal service; attendants, amusement and recreation facilities; aircraft engine mechanics; automobile body and related repairers; heavy equipment mechanics; tool and dye makers; sheet metal workers; garage and service station related occupations; and vehicle washers and equipment cleaners.

## Prevalence

**Table 3.4: Prevalence Rates of Self-Reported Drug Use Among Full-Time Workers, for Selected Drug Categories, by Income and Gender: 1988, 1990**

	Past-Month Use of Any Illegal Drug <sup>a</sup>		Past-Month Heavy Use of Alcohol <sup>b</sup>	
	1988	1990	1988	1990
\$30,000 or over				
Male	10.5%	7.8%	9.6%	13.9%
Female	14.4	6.2	*	*
\$20,000 to \$29,999				
Male	17.8	11.0	12.8	11.8
Female	6.4	9.7	*	3.5
\$12,000 to \$19,999				
Male	19.4	11.7	10.5	17.2
Female	12.7	5.8	5.6	2.2
Less than \$12,000				
Male	26.8	21.7	18.5	19.1
Female	10.2	7.8	*	3.2

a. Includes use of marijuana, hashish, inhalants, hallucinogens, cocaine, heroin, and nonmedical use of pharmaceuticals.

b. Having 5 or more drinks on the same occasion 5 or more times in the past 30 days.

\* Low precision; no estimate reported.

Sources: Kopstein and Gfroerer, 1990: Table 3 (1988 figures) and Normand, Lempert and O'Brien, 1994: Table 3.3 (1990 figures).

**Table 3.5: Prevalence Rates of Self-Reported Drug Use Among Full-Time Male Workers, Aged 18-34, for Selected Drug Categories, by Selected Industries: 1988, 1990**

Industry	Past-Month Use of Any Illegal Drug <sup>a</sup>		Past-Month Heavy Use of Alcohol <sup>b</sup>	
	1988	1990	1988	1990
Construction	28.1%	20.0%	20.9%	26.4%
Manufacturing	14.8	9.7	11.1	12.7
Transportation	18.4	10.5	9.5	20.4
Wholesale Trade	20.6	13.8	7.7	20.5
Retail Trade	13.0	16.8	16.0	14.3
Finance	25.3	11.3	7.7	16.7
Repair Services	22.7	16.1	16.0	9.3
Professional	21.6	9.0	11.4	9.9

a. Includes use of marijuana, hashish, inhalants, hallucinogens, cocaine, heroin, and nonmedical use of pharmaceuticals.

b. Having 5 or more drinks on the same occasion 5 or more times in the past 30 days.

Sources: Kopstein and Gfroerer, 1990: Table 5 (1988 figures) and Normand, Lempert and O'Brien, 1994: Table 3.4 (1990 figures).

In their recent review of cross-occupation and occupation-specific studies dating back as far as 1945, Martin, Kraft and Roman (1994) found evidence that workers in certain occupations are more likely to have alcohol problems. These occupations include food and beverage servers, sales workers, brewery workers, members of the armed forces, sailors and seamen, protective service workers, workers in domestic services, transport operatives, journalists and workers in the newspaper trades, extractive workers, and workers in construction trades. Many of these occupations display alcohol problems cross-nationally. While many of these studies lack rigorous measures of alcohol abuse, studies that do use more rigorous measures (e.g., DSM-III diagnostic criteria) tend to corroborate these findings (Stinson, DeBakey and Steffens, 1992).

In general, these data show a relatively high prevalence of alcohol abuse across a range of occupations, including occupations that are commonly found in the nuclear power industry; e.g., pipefitters, steam fitters and carpenters.

### 3.4: Prevalence of Illegal Drug Use and Alcohol Abuse On the Job in the American Workplace

The use of illegal drugs and the abuse of alcohol while on the job are particularly serious fitness-for-duty concerns. The *Under the Influence* report<sup>8</sup> provides a review of fifteen studies of self-reported workplace illegal drug use and alcohol use and abuse prevalence (Normand, Lempert and O'Brien, 1994, pp. 89-95). In the report's summary, the authors advocate caution in drawing conclusions from such studies, which are often limited to a particular industry or organization:

These studies cannot be readily compared because they vary greatly with respect to samples, job types, periods of assessment, definitions of *at work*, and type of substance. At best, these studies provide only preliminary estimates of the extent of alcohol and other drug use on the job or being high at work (Normand, Lempert and O'Brien, 1994, pp. 89-90; original emphasis).

Another recent review of the literature agrees:

... estimates of the prevalence of employee substance abuse vary significantly depending on the population studied and the measures

utilized . . . . Given the range of available prevalence estimates, it is unlikely that a consensual estimate of [workplace] alcohol or drug abuse rates can be obtained from the research literature (Martin, Kraft and Roman, 1994, p. 10).

Still, two recent reviews of the literature (Martin, Kraft and Roman, 1994; Normand, Lempert and O'Brien, 1994) suggest that studies limited to specific industries, organizations or occupations have produced some general conclusions about workplace drug use. These findings are generally consistent with findings of large-scale surveys on drug use:

- Drug use in the workplace "has been found to be inversely related to occupational status" (Martin, Kraft and Roman, 1994, p. 10); that is, the higher the status of an occupation, the lower the reported drug use on the job.
- Men are more likely than women to use drugs in the workplace (Normand, Lempert and O'Brien, 1994), even when controlling for occupation type (Martin, Kraft and Roman, 1994).
- Younger workers are more likely than older workers to use drugs in the workplace (Normand, Lempert and O'Brien, 1994), even when controlling for occupation type (Martin, Kraft and Roman, 1994).

In terms of specific drugs, alcohol was usually found to be the most common drug used in the workplace. Studies with contradictory findings tended to have flawed or limiting methodology:

The findings in several studies that more respondents admit to using marijuana than alcohol in the workplace is surprising, since alcohol is in general the most frequently used drug at the work site. This discrepancy is probably a sample artifact, as in [one study] in which respondents were limited to those who had used the illicit drug (Normand, Lempert and O'Brien, 1994, pp. 96-97).

Table 3.6 shows trends from 1987 to 1991 in workplace drug use prevalence across gender and specific drugs, for workers aged 19-28. The findings reveal a decreasing use of every substance by males and most substances by females (the trend in tranquilizer use by females is unclear). The most common drugs used in the workplace by these workers in 1991 were alcohol (for both genders),

## Prevalence

**Table 3.6: Trends in Alcohol and Other Drug Use at Work, Past 12 Months, Young Adults Aged 19-28 Years**

Substance Sex	Year of Survey				
	1987	1988	1989	1990	1991
Alcohol					
Male	12.2%	7.0%	9.0%	8.8%	7.6%
Female	7.7	6.3	5.4	6.0	4.6
Marijuana					
Male	8.0	5.6	5.1	4.6	4.5
Female	1.5	1.6	1.5	1.2	0.9
Cocaine					
Male	2.9	2.7	1.5	1.6	0.5
Female	1.0	2.3	0.5	1.1	0.3
Amphetamines					
Male	2.9	2.8	1.5	1.6	0.4
Female	3.0	2.0	2.0	1.3	0.8
Tranquilizers					
Male	0.8	0.3	0.5	0.3	0.2
Female	1.0	0.5	0.1	0.4	0.3

Source: Normand, Lempert, and O'Brien, 1994:96. Reprinted with permission from **UNDER THE INFLUENCE? DRUGS AND THE AMERICAN WORKFORCE**. Copyright 1994 by the National Academy of Sciences. Courtesy of the National Academy Press, Washington, D.C.

followed by marijuana for males and marijuana and amphetamines for females. Males use more alcohol and marijuana than do females; gender differences in the use of other drugs are less pronounced.

### 3.4.1: Workplace Drug Use Prevalence Across Occupation

Table 3.7 shows workplace drug use prevalence across occupational category, gender and specific drugs, for workers aged 19-28 in 1991. The findings show that alcohol is the most common drug used on the job by younger workers. The highest prevalence of alcohol use (by both genders) is in skilled (trade) and manager/sales representative occupations, followed by semiskilled occupations. Highest prevalence of illegal drug use (by both genders) tends to be in semiskilled and skilled (trade) occupations.

### 3.4.2: Workplace Drug Use Prevalence in the Nuclear Industry

Other than the NUREG/CR series of reports,<sup>9</sup> the latest of which is summarized in Section 3.5.1, few studies

have examined workplace drug use among nuclear power facility employees.<sup>10</sup> An examination of drug test results at a nuclear power facility in California during implementation of a drug screening program between 1984 and 1988 (Osborn and Sokolov, 1989; see also Osborn and Sokolov, 1990) was summarized as follows:

3% of drug test failure in November, 1988, less than 2% per month thereafter; higher for contract than for permanent employees; 40% of positives for marijuana, 31% for amphetamines, 29% for cocaine (Normand, Lempert and O'Brien, 1994, p. 84).

Perhaps most notable is the high rate of positive results for amphetamines, which is due in part to the location of the facility. Parts of California are known to have particularly high rates of amphetamine use (see Section 3.5.1 on Nuclear Regulatory Commission (NRC) test results for further evidence of this). On another issue, Osborn and Sokolov (1989, 1990) suggest that contract workers (predominantly in the construction crafts) have higher rates of test failure because they tend to have "less to lose" and can obtain employment elsewhere.

**Table 3.7: Use of Alcohol and Other Drugs at Work in Past 12 Months, Young Adults Aged 19-28 by Occupational Category**

Sex	Occupation	Alcohol	Marijuana	Cocaine	Amphetamines	Tranquilizer
<b>Males</b>						
	Semiskilled <sup>a</sup>	9.6%	9.5%	2.6%	3.4%	0.8%
	Clerical, sales <sup>b</sup>	4.2	3.4	1.8	1.4	0.5
	Police, fire <sup>c</sup>	3.6	1.3	0.0	0.9	0.0
	Military <sup>d</sup>	5.8	1.1	0.7	1.4	0.0
	Skilled <sup>e</sup>	11.3	9.2	3.6	2.4	0.5
	Manager, sales rep. <sup>f</sup>	10.1	2.8	0.8	1.4	0.4
	Professional <sup>g</sup>	8.5	2.8	1.4	0.5	0.3
<b>Females</b>						
	Semiskilled <sup>a</sup>	6.8	2.7	1.2	3.0	0.5
	Clerical, sales <sup>b</sup>	6.0	1.6	1.6	1.9	0.2
	Police, fire <sup>c</sup>	0.0	0.0	0.0	3.2	0.0
	Military <sup>d</sup>	5.4	0.0	0.0	3.2	0.0
	Skilled <sup>e</sup>	7.0	4.8	1.0	4.0	2.9
	Manager, sales rep. <sup>f</sup>	7.9	0.5	1.1	0.0	0.9
	Professional <sup>g</sup>	5.0	0.2	0.1	0.6	0.4

- a. Semiskilled: laborer (car washer, sanitary worker, farm laborer); service worker (cook, waiter, barber, janitor, gas station attendant, practical nurse, beautician); operative or semiskilled worker (garage worker, taxicab, bus or truck driver, assembly line worker, welder).
- b. Clerical, sales clerk: sales clerk in a retail store (shoe salesperson, department store, drug store); clerical or office worker (bank teller, bookkeeper, secretary, typist, postal clerk or carrier, ticket agent).
- c. Police, fire: protective service (police officer, fireman, detective).
- d. Military: military service.
- e. Skilled: craftsman or skilled worker (carpenter, electrician, brick layer, mechanic, machinist, tool and die maker, telephone installer).
- f. Manager, sales rep.: farm owner, farm manager; owner of a small business (restaurant owner, shop owner); sales representative (insurance agent, real estate broker, bond salesman); manager or administrator (office manager, sales manager, school administrator, government official).
- g. Professional: professional with doctoral degree or equivalent (lawyer, physician, dentist, scientist, college professor).

Source: Normand, Lempert and O'Brien, 1994:97. Reprinted with permission from *UNDER THE INFLUENCE? DRUGS AND THE AMERICAN WORKFORCE*. Copyright 1994 by the National Academy of Sciences. Courtesy of the National Academy Press, Washington, D.C.

Original source: Unpublished data from the High School Senior Surveys (1992).

### 3.5: Prevalence of Drug Use Detected in Workplace Drug Testing

This section describes the results of workplace drug testing. Results of workplace testing programs are addressed separately because the relationship between drug detection rates and drug use prevalence in the workforce population is often unclear. Several factors, most notably the differential detection time across drugs and type of test, affect the relationship between the positive test rate of testing programs and the prevalence of use in the workplace where testing occurs. Certain types of tests (e.g., post-accident testing) provide a reasonable assessment of drug use on the job, while others (e.g., random testing) may best reflect drug use in

the workforce, and others (e.g., for-cause testing) may reflect either type of drug use. As summarized in *Under the Influence*:

Urine testing done with appropriate quality control procedures provides an objective indicator of recent drug use that is not subject to problems of underreporting, as are self-reports of use. However, because this kind of testing measures only a subset of drugs and detects only very recent drug use, the results may not be directly comparable with self-reported data, which are typically based on longer reference periods such as a month or a year. Furthermore, drug test findings are not based on probability samples, hence estimates of drug use prevalence

## Prevalence

derived from them may be subject to unknown biases (Normand, Lempert and O'Brien, 1994, p. 74).

An advantage of workplace drug testing is the increased potential to distinguish use of legal drugs from abuse of such drugs. For example, and most notably, while self-reported use—perhaps even heavy use—of alcohol may not have fitness-for-duty implications, a positive test result for alcohol may indicate a fitness-for-duty problem that may be from on-the-job alcohol abuse.

In this section, the results from four primary sources of information derived from worksite drug testing programs are discussed:

- NRC results from nuclear power licensees with fitness-for-duty (FFD) programs (Westra et al., 1994);
- Drug Testing Index (SmithKline Beecham, 1994);
- NIDA-certified laboratories testing for organizations (as summarized in Normand, Lempert and O'Brien, 1994; pp. 73-81); and
- American Management Association (AMA) survey of firms testing for drugs (AMA, 1994).

Although all of these sources report results of workplace drug testing, the results are not fully comparable for a number of reasons. First, although all programs test for the five drugs covered under the DHHS guidelines ( opiates, marijuana, cocaine, PCP, and amphetamines), some NRC licensees also test for additional drugs (primarily barbiturates and benzodiazepines), some licensees use lower (i.e., more rigorous) cutoff levels than those specified by DHHS guidelines, and all of the NRC regulated FFD programs include breath testing for alcohol. Second, some sources, such as the NRC program results, report positive rates based on results that have been reviewed by an MRO. Other sources, such as the Drug Testing Index, reflect the percent of samples that are determined to be positive by GC/MS testing by a laboratory. MROs report as positive only those results that are determined to be confirmed as positive by laboratory GC/MS testing and for which there is no legitimate, medical explanation.<sup>11</sup>

### 3.5.1: Test Results Under NRC's FFD Rule

This section provides a summary of drug test results on workers covered by the nuclear power industry's fitness-

for-duty rule (Westra et al., 1994). Of over 242,000 tests in 49 utilities in 1993, less than 1% (0.62%) showed positive results for any illegal drug or alcohol. For-cause tests produced the highest rate of positive results (21.70% were positive). Although NRC reports do not contain information on type of substance by type of test, anecdotal evidence suggests that the majority of the for-cause positive test results are from alcohol. Since other programs may not include testing for alcohol, the higher positive percentage of for-cause tests in the nuclear industry may be attributable to this difference. (This rate of detection may also reflect a relatively high degree of management attention given to behavioral observation techniques.) Pre-access and follow-up tests each produced less than 2% positive results, while random tests—the most common type of test—produced 0.23% positive results.

Among positive results for alcohol and drugs covered by the SAMHSA guidelines, marijuana was the most common drug detected (49.56% of all positive results), followed by cocaine (23.41%) and alcohol (22.65%). Amphetamines (3.24%), opiates (less than 1%), and PCP (less than 1%) were rarely detected. Within reporting units that also tested for barbiturates and benzodiazepines, each of these two drug categories accounted for less than 1% of positive results.

In each of the NRC regions, marijuana was the most common drug detected, accounting for over 40% of the positive test results for the five drugs included under the SAMHSA guidelines and alcohol in every region. Cocaine was the second most common drug detected in the two eastern regions, while alcohol had that distinction in the midwest and southwest. Notably, amphetamines was the second most common drug detected in the far west, accounting for over 22% of the positive results, while amphetamines accounted for less than 2% of positive results in other regions of the country.

In terms of trends in NRC test results, while the overall positive test rate declined sharply between 1990 (0.87%) and 1991 (0.66%), it has remained essentially steady from 1991 through 1993 (0.62%). However, the positive rate among tests other than pre-access has continued to decline, from 0.47% in 1991 to 0.37% in 1993. This suggests that the absence of further decline in the overall positive rate is due in large part to the increasing portion of NRC tests that are pre-access (Westra et al., 1994).

The distribution of positive test results across drug type has changed little since 1990. The portion of positive test results caused by marijuana has increased somewhat

(from about 47% in 1990 to 50% in 1993), while the portion caused by cocaine has fallen (from about 29% to 23%). The portion of positive results caused by alcohol has risen slightly (from 19% to 21%). The portions caused by other drugs have remained small (amphetamines at less than 4%, opiates and PCP at less than 2%).

### 3.5.2: Drug Testing Index Results

The Drug Testing Index summarizes the results of employee drug tests conducted for numerous companies across various industries by the clinical testing laboratory SmithKline Beecham. The latest results, based on more than 11 million tests conducted between 1987 and 1993, show that after declining steadily from 18.1% in 1987 to 8.8% in 1991, the drug-positive rate has leveled off, showing slight declines in 1992 (to 8.76%) and 1993 (to 8.41%) (SmithKline Beecham, 1994) (see Table 3.8). The nuclear power industry rate since 1990 is also provided in Table 3.8 for comparison. As noted earlier, the nuclear power industry rates include positive results for alcohol, while Drug Testing Index results do not. Approximately 25% of positive test results in the nuclear industry are attributed to alcohol (Westra et al., 1994). It should also be noted that the nuclear power industry rates reflect further screening by MROs to determine if results

were based on responsible use. Data collected from a sample of utilities suggest that about 20% of laboratory-confirmed positives would be estimated to be negated based on MRO determination of legitimate use in 1993. The MRO determination of legitimate use is usually with respect to opiates (Durbin et al., 1991b).

Drug Testing Index positive rates are much lower for those workers categorized in the SmithKline Beecham report as "Safety-Sensitive Transportation Workers" (SSTWs) (see Table 3.9). For example, in 1993 there was a positive rate of 2.84% for SSTWs, compared to 8.41% for all other categories of American workers. The positive test rate for SSTWs is much closer to the overall 1993 positive test rate of 0.62% for the workers covered by the nuclear power industry's fitness-for-duty rule (Westra et al., 1994).

Furthermore, the positive test rates are similar across testing categories, and across drug categories, when these safety-sensitive populations are compared. As can be seen in Table 3.9, the nuclear power industry has, on average, a lower positive test rate than is reported for SSTWs as reported by the SmithKline Beecham Drug Testing Index (1994). The exception is in for-cause testing, where the nuclear power industry has a higher percent positive.

**Table 3.8: Positive Test Rates: Comparison of Nuclear Power Industry Results and SmithKline Beecham Drug Testing Index Results**

Year	SmithKline Beecham Drug Testing Index Positive Test Rate <sup>a</sup>	Nuclear Power Industry Positive Test Rate <sup>b</sup>
1987	18.1%	NA
1988	13.6	NA
1989	12.7	NA
1990	11.0	0.87%
1991	8.8	0.66
1992	8.76	0.68
1993	8.41	0.62

NA Not available.

a. Source: SmithKline Beecham, 1994.

b. Source: Durbin et al., 1991b; Murphy et al., 1992; Fleming et al., 1993; Westra et al., 1994.

## Prevalence

**Table 3.9: Positive Test Rates by Testing Category: Comparison of Nuclear Power Industry Results and SmithKline Beecham Drug Testing Index Results for Safety-Sensitive Transportation Workers (SSTWs)**

Testing Category	SmithKline Beecham Drug Testing Index Positive Test Rate for SSTWs 1993 <sup>a</sup>	Nuclear Power Industry Positive Test Rate 1993 <sup>b</sup>
Pre-employment	3.38%	1.04%
Random	2.35	0.23
Postaccident	3.25	0.00
For-cause (not including post-accident)	9.85	27.21
Follow-up	NA	1.35
Total	2.84	0.62

NA Not available.

a. Source: SmithKline Beecham, 1994.

b. Source: Westra et al., 1994.

The nuclear power industry also tests for alcohol, which represented a substantial number of positive results (22.7%). In order to compare with the Drug Testing Index, which did not include information about the number of positive tests caused by alcohol, all positive tests caused by alcohol are excluded from the nuclear power industry findings shown in Table 3.10.

Again, it should be noted that the nuclear power industry rates reflect MRO review while the SmithKline Beecham rates do not. Table 3.10 shows the estimated GC/MS positive result distribution across drug types. This estimated distribution of drug types is similar to the distribution in the Drug Testing Index.<sup>10</sup>

**Table 3.10: Distribution of Positive Test Results by Drug Category: Comparison of Nuclear Power Industry Results and SmithKline Beecham Drug Testing Index Results for Safety-Sensitive Transportation Workers (SSTWs)**

Drug Category	SmithKline Beecham Drug Testing Index Positive Test Results for SSTWs, 1993 <sup>a</sup>	Nuclear Power Industry MRO-Confirmed Positives, 1993 <sup>b</sup>	Nuclear Power Industry Estimated GC/MS-Confirmed Positives, 1993 <sup>c</sup>
Marijuana	51.2%	64.1%	51%
Cocaine	28.3	30.3	24
Opiates	13.0	1.1	21
Amphetamines	5.3	4.2	4
PCP	2.2	0.4	0.3

a. Source: SmithKline Beecham, 1994.

b. Source: Westra et al., 1994 (calculated based on numbers reported; excludes positive results for alcohol).

c. See Endnote 11 (Table 11b) for an illustration of how these numbers were derived.

### 3.5.3: NIDA Test Results

This section describes drug test results from seven NIDA-certified laboratories that perform drug tests for organizations (previously unpublished results summarized in Normand, Lempert and O'Brien, 1994). As noted in Normand, Lempert and O'Brien (1994), a discussion of these results has several limitations, due primarily to the diversity of testing and test-reporting conditions across employers. First, the positive rates are somewhat inflated, since some employers submit an unknown number of positive blind quality control samples containing drugs; the extent of this bias is unknown. (NIDA guidelines call for 20% of the blind specimens to be "spiked.") Second, many employers test only for the five drugs covered under the SAMHSA guidelines (which exclude such drugs as alcohol, barbiturates and benzodiazepines). Third, employers following the SAMHSA protocol will tend to have higher cutoff levels for positive results than will other employers. Fourth, employers vary in the extent to which they provide drug program elements to deter drug use, such as drug awareness education, supervisor training in behavioral observation and willingness to confront, strict sanctions, and treatment opportunities.

About four percent of the nearly 2 million test results between October, 1990 and March, 1992 were positive, with the most commonly-detected drugs being marijuana at 2% positive, followed by cocaine at 1%, opiates at 0.6% and benzodiazepines at 0.5% (Normand, Lempert and O'Brien, 1994, p. 75). Of the different types of testing, postaccident testing showed the highest rates of positive results (5%). Results from random testing revealed 3% positive results.

Regional differences in positive results were also statistically significant, with the West region having the highest positive rates overall and for all categories of drugs tested except cocaine, which was most prevalent in the Northeast.<sup>12</sup> Across industries, by far the highest rates of positive tests were found in the construction industry (almost 6% positive), followed by the retail and wholesale trade industries. As noted in Section 3.3.3 above and in Normand, Lempert and O'Brien (1994), these are the same industries with the highest rates of self-reported use by males aged 18-34 in the 1990 NHSDA survey.

However, the positive rates among construction workers are much lower than their comparable NHSDA self-report rates. Besides the obvious differences in instruments—NHSDA includes self-reported use of a

larger pool of drugs over a longer reference period—this difference might also be due to a difference in workers sampled: ". . . the NHSDA includes responses from casual construction workers working for small companies; those tested were more likely to be working on government projects or for larger organizations" (Normand, Lempert and O'Brien, 1994, p. 80).

### 3.5.4: AMA Test Results

The American Management Association (AMA) conducts an annual survey on workplace drug testing through a questionnaire mailed to each of its member companies. The latest survey results, covering 1993, are based on 794 returned questionnaires. A large majority (87.2%) of the companies responding test for employee drug use, compared with only 21.5% in 1987. Nearly half (48.1%) of the companies responding are in the manufacturing business (AMA, 1994).

The percentage of these companies' drug tests showing positive results decreased steadily over the past few years before leveling off in 1993. Of the over 96,000 tests given to employees in 1993, 2.5% were positive, compared to 8.1% in 1989. Of the over 97,000 tests given to new hires in 1993, 4.3% were positive, compared to 11.4% in 1989 (AMA, 1994).

It should be noted that such dramatic declines in positive drug results do not necessarily reflect less drug use by employees. They may also reflect such factors as a widening testing pool, the relative increase in periodic and random testing with its relatively lower positive rates (as opposed to for-cause testing with relatively higher positive rates), the effects of previous screening of applicants (e.g., pre-employment testing), the reduction in the number of drug and alcohol abusers due to their identification and removal based on previous tests, and the self-exclusion of persons withdrawing their applications due to pre-employment testing (AMA, 1993, 1994). In addition, these changes may reflect an increase in successful subversion of the testing process.

### 3.5.5: Summary of Workplace Drug Test Results

In 1993, the overall positive test rate in the nuclear power industry (0.62%) was considerably less than rates in other industries, as reported by the Drug Testing Index (8.41%), NIDA (4%) and AMA (2.5%), as well as rates for "Safety-Sensitive Transportation Workers" (SSTWs) (2.84%) as compiled in the Drug Testing Index.<sup>13</sup>

## Prevalence

Across drug type, marijuana was the most common drug detected in all of the samples.

Across test type, the nuclear industry had substantially lower positive test rates than other industries in all categories except for-cause, where the nuclear industry rate (21.70%) was relatively high. The nuclear power industry's higher positive rate from for-cause testing may reflect the inclusion of alcohol or may reflect managerial training to accurately detect subordinates under the influence of illegal drugs or alcohol. The lower positive test rates in the other three categories (pre-employment, random, and postaccident) may indicate that the fitness-for-duty program implemented by the nuclear power industry is having a positive effect on reducing drug use. However, the fact that even the pre-employment positive rate is lower suggests that people attracted by work in this industry may be less likely to use drugs than are the SSWTs included in the Drug Testing Index.

### 3.6: Chapter Conclusions and Implications for Drug Testing in the Nuclear Power Industry

#### 3.6.1: Chapter Conclusions

This chapter contains a summary of drug use prevalence in the United States (and particular subpopulations), based on results of self-report surveys, emergency room episodes, and workplace drug testing. The findings suggest the following conclusions:

Self-report data indicate that the use of illegal drugs and alcohol have declined significantly among the general population over the past decade.

However, several factors suggest that a core group of alcoholics and drug addicts, as well as a revolving group of younger, more casual experimenters, remain. These factors include (1) a recent leveling off of the decline in drug use among Americans in general, (2) a recent increase in the number of drug-related emergency room episodes, and (3) increasing evidence from field reports of increasing numbers of heroin users.

While trends in marijuana use are age-specific (it appears to still be declining among the general population, but is increasing among younger Americans), it clearly remains the most popular illegal drug. In fact, almost half of illegal drug users reported using marijuana only.

Marijuana is also the most common drug detected in tests of nuclear power workers.

Cocaine use, though declining among the general population, is still prevalent enough to pose serious fitness-for-duty concerns, especially among younger adults. Cocaine is not only the second most common drug detected in drug-related emergency room episodes, but it is also the second most common drug detected in tests of nuclear power workers.

Illegal use of many highly-publicized drugs—including heroin, PCP and steroids—has remained relatively rare among the general population. While heroin is the third most common drug detected in drug-related emergency episodes, this does not reflect its relative prevalence among American workers, including those in the nuclear industry; opiates are infrequently confirmed by MROs in tests of nuclear power workers.

Alcohol presents significant fitness-for-duty concerns. While alcohol use in general has declined among the general population, the prevalence of heavy use has remained stable. Heavy alcohol use is relatively common in some occupations commonly found in the nuclear power industry—e.g., pipefitters, steam fitters, and carpenters. Alcohol is the third most common drug detected in tests of nuclear power workers, and the most common drug detected in tests of supervisors in nuclear power plants.

Amphetamine use, while low among the general population and nuclear power workers in general, is the second most common drug detected in tests of nuclear power workers in the western United States (former Region V).

Self-report and drug-test data suggest that some demographic and occupational groups tend to have higher rates of illegal drug use and heavy alcohol use; males, young adults, lower-income workers, and construction workers. Illegal drug use also tends to be more prevalent among the unemployed. College students and white-collar workers are more likely to use, but less likely to abuse, alcohol.

In terms of region, use of both illegal drugs and alcohol are most prevalent in the West. Notably, the positive test rate of nuclear power workers is second highest in the West (former Region V), despite that region having the lowest portion of tests given to short-term contractors.

### 3.6.2: Implications for Drug Testing in the Nuclear Power Industry

The findings on drug use prevalence, as summarized in this chapter, suggest the following implications for drug testing in the nuclear power industry:

#### 3.6.2.1: Panel of Drugs

##### Recommendation

NRC fitness-for-duty programs should place particular emphasis on five types of drugs: alcohol, marijuana, cocaine, amphetamines, and prescription-type pharmaceuticals (including stimulants, sedatives, tranquilizers, and analgesics).

##### Supporting Evidence

- Findings presented in this chapter suggest that alcohol abuse in the workplace poses a formidable threat to fitness-for-duty program goals.

Despite the popular attention paid to illegal drugs, alcohol clearly remains the most commonly used, and abused, mind-altering drug in America. This is true for both the general population and employed workers. Survey findings show that while illegal drug use has declined among the general population over the past decade, heavy alcohol use has remained stable.

Furthermore, test results showed alcohol to be the third most common drug detected among nuclear power workers. This is particularly troubling, given that a positive test result for alcohol may indicate on-the-job alcohol impairment.

- Marijuana remains the most popular illegal drug, both among the general population and in the nuclear power workplace.

While trends in marijuana use are age-specific (it appears to be declining among the general population, but increasing among young persons), it clearly remains the most popular illegal drug (in fact, almost half of self-reported illegal drug users reported using marijuana only). Marijuana was the most common drug detected in tests of nuclear power workers.

- Cocaine use, though declining among the general population, is still prevalent enough (especially

among young adults) and dangerous enough to pose serious fitness-for-duty concerns.

Cocaine is the third most prevalent illegal drug used among the general population, and consistently the most common illegal drug in drug abuse-related emergency room visits. A considerable portion (5.0%) of Americans aged 18-25 reported using cocaine in 1993.

- Amphetamine use is particularly high among nuclear power plant workers in Western states.

Amphetamine was the second most common drug detected in the former Region V (Western states), producing over 22% of the positive test results.

- The increased prevalence in the use and abuse of pharmaceuticals, combined with common complacency regarding their use, pose fitness-for-duty concerns.

The use of pharmaceuticals, especially when used counter to medical prescription, should raise concerns. NHSDA results show nonmedical use of a pharmaceutical as the second most prevalent form of illegal drug use. The authors of *Under the Influence* note a need for concern over the particular nature of this problem (especially among women), given Americans' increased tendency to treat mental depression through psychotropic means (Martin, Kraft and Roman, 1994). Components of NRC fitness-for-duty programs—drug education, supervisor training, treatment opportunities—should reflect the fitness-for-duty risks posed by these relatively prevalent forms of drug abuse.

This recommendation does not imply that PCP and opiates should be dropped from the panel of drugs. Instead, it suggests that program emphasis should be placed on those drugs where there is the most evidence of abuse. This could and most likely will change over time. Fitness-for-duty program managers need to continue to adjust program elements to reflect the most prevalent local issues.

#### 3.6.2.2: Balanced Approach to Illegal Drug Use and Abuse of Legal Drugs

##### Recommendation

Care should be taken in FFD programs to avoid overemphasis on "hard-core" illegal drugs and complacency about legal drugs.

## Prevalence

### Supporting Evidence

In general, it should not be assumed that typical media accounts of drug abuse reflect the most common forms of drug abuse in the workplace. Abuse of illegal "hard drugs" (e.g., heroin, cocaine) focused upon by the mass media have tended to be concentrated in the non-working population. Drug use among working Americans has tended to involve different drugs: alcohol (predominant in all sectors of American society), marijuana, and, increasingly, the abuse of prescription-type pharmaceuticals. In addition, amphetamine abuse is a problem in some industries, including the nuclear power industry.

It is notable that two of the five primary drugs of concern listed in the recommendation regarding the panel of drugs (alcohol and pharmaceuticals) are legal substances. More than any other drug, alcohol poses considerable special fitness-for-duty concerns. In addition to its prevalence, alcohol, because of its legality, is more accessible and more tolerated than any illegal drug. Despite this, it is clear that alcohol, when abused, can significantly affect work performance—from increased absenteeism to diminished work performance—and pose serious fitness-for-duty risks.

As with alcohol, pharmaceutical abuse is often produced—and compounded—by complacency about legal drugs:

One cannot simply assume that because these drugs are used as medications, they have little or no impact on job behaviors. Further, such usage can become abuse, sometimes without the knowledge of the individual users or their significant others (Martin, Kraft and Roman, 1994, p. 25).

### 3.6.2.3: Program Flexibility with Regard to Workforce Differences

#### Recommendation

Fitness-for-duty programs should have the flexibility to respond to regional and workforce differences in drug use prevalence.

#### Supporting Evidence

- Findings presented in this chapter suggest that the prevalence—and forms—of drug abuse will vary

across different nuclear power plants, depending on regional and workforce characteristics.

For example, drug testing of nuclear power workers reveal that amphetamine use is much more prevalent in the Western U.S. than in other regions. Test results suggest that alcohol abuse is the most prevalent drug problem among power plant supervisors.

#### 3.6.2.4: Currency of Data on Drug Use Prevalence

#### Recommendation

Fitness-for-duty programs should periodically incorporate current information on drug use prevalence and trends, while stressing the judicious interpretation of such data.

#### Supporting Evidence

Timely information on drug use prevalence can aid fitness-for-duty programs by keeping program managers informed about the latest patterns of drug use both regionally and nationwide. Given the generally clandestine nature of drug abuse, and the variability of drug use prevalence across time and population groups, no single source of data can provide a complete and accurate picture of drug use patterns. Thus, it is essential to obtain the most recent information from a variety of sources that use different methodologies to study different population groups.

It must also be stressed that each set of findings should be interpreted with the specific study's methodology and population of interest in mind. Specific questions to ask include:

- Who was being studied/surveyed (e.g., a random sample of Americans or only those admitted to an emergency room)?
- How was the information obtained (e.g., through a self-report survey or through mandatory drug testing)?
- How precise were the data-collection methods (e.g., an organized series of carefully-worded survey questions or a loosely-structured, open-ended conversation)?
- How recent is the information?

The four primary periodic studies detailed in this chapter—NHSDA, MTF, DAWN and Pulse

Check—provide a diverse set of assessments of Americans' drug use prevalence. Recent reports from each of these studies can be obtained through the

National Clearinghouse for Alcohol and Drug Information, P.O. Box 2345, Rockville, MD 20847-2345 (800-729-6686).

### Endnotes

1. Brent L. Baxter, Ph.D., Alcohol and Drug Abuse Institute, University of Washington, Seattle, Washington
2. Although NHSDA uses the term "psychotherapeutic" for this category of drugs, we believe that it should be broadened to include all pharmaceuticals, including prescription-type or over-the-counter stimulants, sedatives, tranquilizers, or analgesics. See Chapter 4 for a discussion of nonmedical use and abuse of legal drugs.
3. The description of drug use changes as "significant" refers to statistical significance as determined by the original researchers.
4. Tests for statistical significance were not calculated for differences in drug use across education level.
5. Tests for statistical significance were not calculated for differences in drug use across region.
6. Comparing these figures is reasonable. While self-reporting provides conservative figures for illegal drug use, it undoubtedly does the same to some extent for alcohol abuse.
7. Tests for statistical significance were not calculated for differences in drug use across employment status.
8. Another recent analysis of the prevalence of drug use in American society and the American workforce (hereafter referred to as the *Under the Influence* report) used many of the same or comparable/similar sources (Normand, Lempert and O'Brien, 1994) and produced similar findings. The *Under the Influence* report, commissioned by the National Research Council, provides an extensive review of evidence on drug use prevalence from a variety of sources: several large-scale self-report surveys (including the National Household Survey on Drug Abuse (NHSDA) and the High School Senior surveys of the Monitoring the Future series), supplemented by results of workplace drug testing, organization-specific prevalence studies, and epidemiological studies (including the Drug Abuse Warning Network (DAWN) of drug abuse-related emergency room ER episodes).

Conclusions arrived at in the *Under the Influence* report are in striking accordance with the conclusions presented in this chapter:

- The prevalence of illicit drug use among members of the general population and the workforce has been decreasing, but continues to affect a sizable proportion of the population, especially young adults.
- Illicit drug use may be decreasing among occasional users, but it may be stable or even increasing among hard-core users who are generally not well-represented in surveys.
- Illicit drug use is more common among unemployed than employed persons, and weekly alcohol use is highest among young employed persons.
- Illicit drug use is relatively high among male workers in certain industries, such as construction, but relatively low among professionals.

The report argues that given such long-term trends in decreasing drug use, short-term decreases may not be fully attributable to recent concentrated national efforts to discourage use. In addition, the report notes the paucity of studies of drug use specifically among the workforce, and encourages the development of such studies, including the expansion of NHSDA to include more specific items on employment.

9. Barnes et al., 1988; Moore et al., 1989; Durbin et al., 1991a; Durbin et al., 1991b; Murphy et al., 1992; Fleming et al., 1993; Westra et al., 1994.
10. For a summary of organization-specific prevalence studies (which are inconsistent in quality and comparability and, in sum, provide insufficient data on the prevalence of drug use in the American workforce), see Normand, Lempert and O'Brien, 1994, pp. 81-89.
11. The number of MRO-declared positives in the nuclear industry is lowered by approximately 20% if only the five dangerous drugs in the DHHS panel are considered (i.e., if alcohol and other drug results are excluded). However, it is estimated that an additional 25% of GC/MS-confirmed positives for the five drugs are excluded from MRO-declared positives, having been determined by MROs to be due to legitimate use of the drugs or due to invalid test results. Most common among these are opiate positives (determined to be from eating poppy seeds or from medically prescribed codeine) or amphetamine positives. There are also a very small number of invalid results. As shown in Table 11a below, these differences between the NRC results and other reported results such as the Drug Testing Index, largely offset each other; excluding positive tests for alcohol from the NRC results reduces the number of positives to 80% of the total and estimating the number of GC/MS-confirmed positives based on the number of MRO-declared positives raises the number back to near 100% of the total (125% of 80%). NRC industry results are still somewhat different than Drug Testing Index results because some licensees use lower cut off levels and test for additional drugs.

Table 11a

	Number Tested*	MRO-Declared Positives*	MRO-Declared Positives Excluding Alcohol*	Estimated GC/MS-Confirmed Positives Excluding Alcohol†
Number	242,966	1,512	1,219	1,547
Percentage of MRO-Declared Positives	NA	NA	81%	102%
Positive Test Rate	NA	0.62%	0.50%	0.64%

\* Source: Westra et al., 1994.

† See Table 11b.

The following table, based on a study of 10 nuclear power plants, shows the relationship between MRO-declared positives and GC/MS-confirmed positives in the nuclear power industry.

Table 11b

Drug	Nuclear Power Industry MRO-Declared Positives, 1993*	Ratio of GC/MS-Confirmed Positives to MRO-Declared Positives†	Estimated GC/MS-Confirmed Positives‡	Percentage Distribution of Estimated GC/MS-Confirmed Positives
THC	781	1.004	784	51%
Cocaine	369	1.000	369	24%
Opiates	13	25.54	332	21%
Amphetamine	51	1.111	57	4%
PCP	5	1.000	5	0.3%
Total	1,219	NA	1,547	100%

\* Source: Westra et al., 1994.

† Calculated from numbers reported from a study of 10 nuclear power plants. Source: Durbin et al., 1991b, Appendix B.

‡ The product of the first two columns.

12. It should be noted that these reported test results across region (as well as industry) do not control for any differences in the distribution of types of test (pre-employment, random, etc.) across region (and industry).
13. Although these rates are not directly comparable, analyses of what the NRC rate would be under conditions similar to those of the Drug Testing Index indicate that the nuclear industry rate would continue to be approximately 0.6% (see Endnote 11).

## References

American Management Association (AMA). 1993. *1993 AMA Survey: Workplace Drug Testing and Drug Abuse Policies*. New York, NY: AMA.

American Management Association (AMA). 1994. *1994 AMA Survey: Workplace Drug Testing and Drug Abuse Policies*. New York, NY: AMA.

Barnes, V., I. Fleming, T. Grant, J. Hauth, J. Hendrickson, B. Kono, C. Moore, J. Olson, L. Saari, J. Toquam, D. Wieringa, P. Yost, P. Hendrickson, D. Moon, and W. Scott. 1988. *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues* (NUREG/CR-5227). Seattle, WA: Battelle Human Affairs Research Centers.

Durbin, N., S. Murphy, T. Fleming, J. Macaulay, C. Westra, J. Olson, and J. Christensen. 1991a. *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports — CY 1990* (NUREG/CR-5758, Vol. 1.). Seattle, WA: Battelle Human Affairs Research Centers.

Durbin, N., C. Moore, T. Grant, T. Fleming, P. Hunt, R. Martin, S. Murphy, J. Hauth, R. Wilson, A. Bittner, A. Bramwell, J. Macaulay, J. Olson, E. Terrill, and J. Toquam. 1991b. *Fitness for Duty in the Nuclear Power Industry: A Review of the First Year of Program Performance and an Update of the Technical Issues* (NUREG/CR-5784). Seattle, WA: Battelle Human Affairs Research Centers.

Fleming, T., C. Westra, I. Field, C. Forslund, N. Durbin, T. Grant, and R. Moffitt. 1993. *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports — CY 1992* (NUREG/CR-5758, Vol. 3.). Seattle, WA: Battelle Human Affairs Research Centers.

General Accounting Office. 1991 (May). *Employee Drug Testing: Status of Federal Agencies' Programs* (GAO/GGD-91-70). Washington, D.C.: GAO.

General Accounting Office. 1993 (June). *Drug Use Measurement: Strengths, Limitations, and Recommendations for Improvement* (GAO/PEMD-93-18). Washington, D.C.: GAO.

Gerstein, Dean R. and Lawrence W. Green (Eds.). 1993. *Preventing Drug Abuse: What Do We Know?* Washington, D.C.: National Academy Press.

Hser, Y., M.D. Anglin, T.D. Wickens, M. Brecht and J. Homer. 1992 (May). *Techniques for the Estimation of Illicit Drug-Use Prevalence: An Overview of Relevant Issues*. Washington, D.C.: National Institute of Justice.

Johnston, Lloyd D., Patrick M. O'Malley and Jerald G. Bachman. 1992. *Smoking, Drinking, and Illicit Drug Use Among American Secondary School Students, College Students, and Young Adults, 1975-1991, Vol. I and II*. Rockville, MD: National Institute on Drug Abuse.

Johnston, Lloyd D., Patrick M. O'Malley and Jerald G. Bachman. 1993. *National Survey Results on Drug Use from the Monitoring the Future Study, 1975-1992. Volume II: College Students and Young Adults*. U.S. Department of Health and Human Services.

Kopstein, A. and J. Gfroerer. 1990. "Drug use patterns and demographics of employed drug users: Data from the 1988 National Household Survey on Drug Abuse." Pp. 11-24 in *Drugs in the Workplace: Research and Evaluation Data* (Vol. II) (S.W. Gust et al., Eds.). Washington, D.C.: U.S. Department of Health and Human Services.

Martin, Jack K., Joan M. Kraft and Paul M. Roman. 1994. "Extent and impact of alcohol and drug use problems in the workplace: A review of the empirical evidence." Pp. 3-32 in *Drug Testing in the Workplace*, Scott Macdonald and Paul Roman (Eds.). New York: Plenum Press.

Moore, C., V. Barnes, J. Hauth, R. Wilson, J. Fawcett-Long, J. Toquam, K. Baker, D. Wieringa, J. Olson, and J. Christensen. 1989. *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues* (NUREG/CR-5227, Supplement 1). Seattle, WA: Battelle Human Affairs Research Centers.

Murphy, S., T. Fleming, C. Westra, I. Field, N. Durbin, and R. Moffitt. 1992. *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports — CY 1991* (NUREG/CR-5758, Vol. 2). Seattle, WA: Battelle Human Affairs Research Centers.

Normand, Jacques, Richard O. Lempert and Charles P. O'Brien (Eds.). 1994. *Under the Influence? Drugs and the American Work Force*. Washington, D.C.: National Academy Press.

Office of National Drug Control Policy (ONDCP). 1994 (December). *Pulse Check: National Trends in Drug Abuse*. Washington, D.C.: ONDCP.

Osborn, Carl E. and Jacque J. Sokolov. 1989. "Drug use trends in a nuclear power company: Cumulative data from an ongoing testing program." Pp. 69-80 in *Drugs in the Workplace: Research and Evaluation Data* (S.W. Gust and J.M. Walsh, Eds.). NIDA Research Monograph 91. Rockville, MD: National Institute on Drug Abuse.

Osborn, Carl E. and Jacque J. Sokolov. 1990. "Drug use trends in a nuclear power facility: Data from a random screening program." Pp. 63-77 in *Drugs in the Workplace: Research and Evaluation Data* (Vol. II) (S.W. Gust et al., Eds.). NIDA Research Monograph 100. Rockville, MD: National Institute on Drug Abuse.

Parker, Douglas A. and Thomas C. Harford. 1992. "The epidemiology of alcohol consumption and dependence across occupations in the United States." *Alcohol Health and Research World* 16(2):97-105.

SAMHSA\_DAWN\_92. 1993 (September). *Estimates from the Drug Abuse Warning Network: 1992 Estimates of Drug-Related Emergency Room Episodes*. SAMHSA, Office of Applied Studies.

SAMHSA\_DAWN\_93. 1994 (December). *Preliminary Estimates from the Drug Abuse Warning Network: 1993 Preliminary Estimates of Drug-Related Emergency Department Episodes: Advance Report Number 8*. SAMHSA, Office of Applied Studies.

SAMHSA\_NHSDA\_News. 1993 (Summer). "Household survey shows drug use continues to decline." *SAMHSA News* 1(3), 11.

SAMHSA\_NHSDA91. 1993. *National Household Survey on Drug Abuse: Main Findings 1991*. SAMHSA,

Office of Applied Studies, National Household Survey on Drug Abuse.

SAMHSA\_NHSDA92. 1993 (June). *Preliminary Estimates from the 1992 National Household Survey on Drug Abuse: Selected Excerpts*. SAMHSA, Office of Applied Studies, National Household Survey on Drug Abuse.

SAMHSA\_NHSDA93. 1994 (July). *Preliminary Estimates from the 1993 National Household Survey on Drug Abuse: Advance Report 7*. SAMHSA, Office of Applied Studies, National Household Survey on Drug Abuse.

SAMHSA\_NHSDA92\_POPUL. 1993 (October). *National Household Survey on Drug Abuse: Population Estimates 1992*. SAMHSA, Office of Applied Studies, National Household Survey on Drug Abuse.

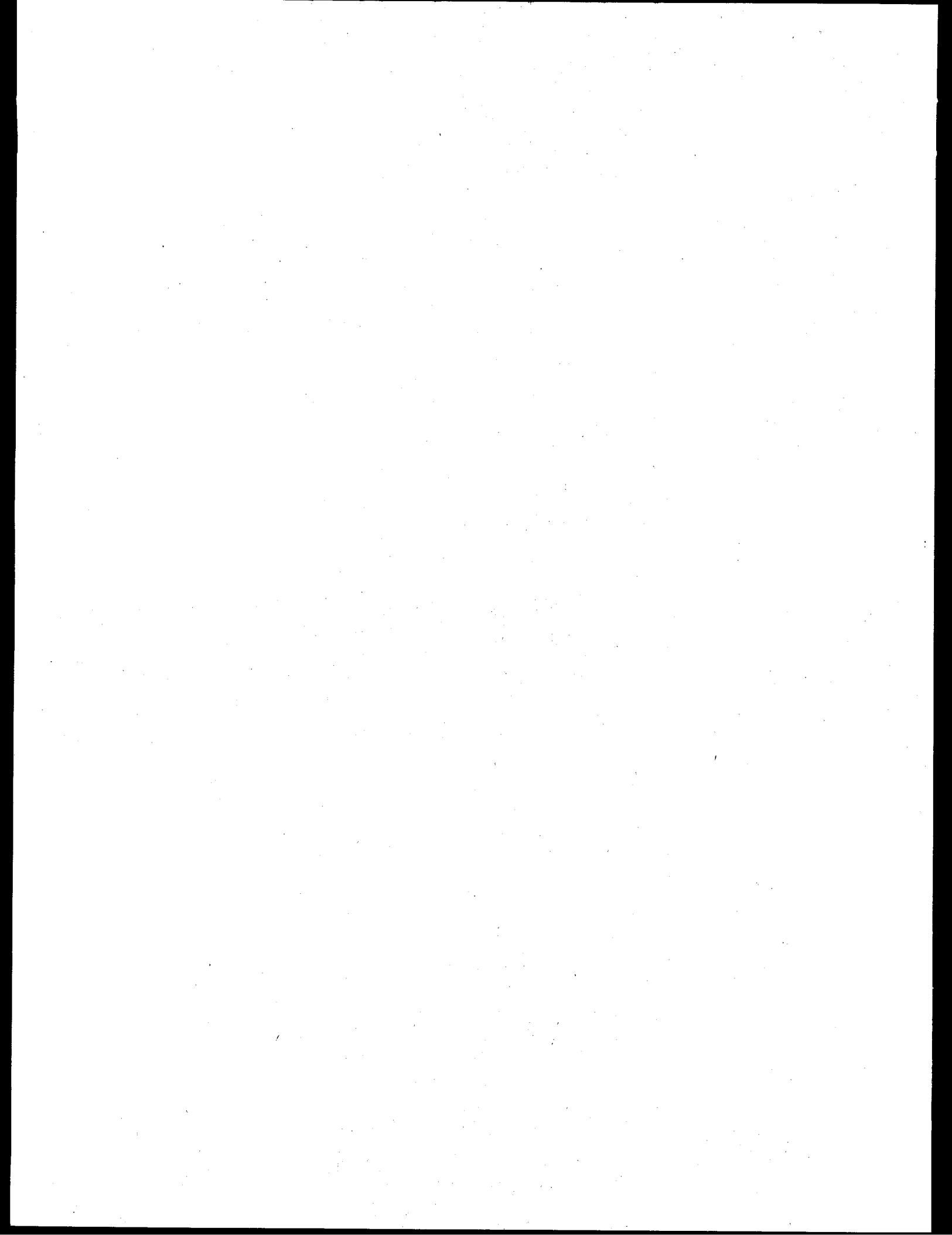
SmithKline Beecham Clinical Laboratories. 1993 (Feb. 10). *SmithKline Beecham Drug Testing Index Shows Trend Flat for U.S. Work Force*. King of Prussia, PA: SBCL Communications.

SmithKline Beecham Clinical Laboratories. 1994 (Feb. 17). *SmithKline Beecham Announces Decline in Workplace Positive Drug Tests*. Collegeville, PA: SBCL Communications.

Stinson, Frederick S., Samar Farha DeBakey and Rebecca A. Steffens. 1992. "Prevalence of DSM-III-R alcohol abuse and/or dependence among selected occupations: United States, 1988" (Epidemiologic Bulletin No. 30). *Alcohol Health and Research World* 16(2):165-172.

Treaster, Joseph B. 1993. "U.S. reports rise in drug emergencies." *New York Times*, October 5, 1993.

Westra et al. 1994. *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports — CY 1993* (NUREG/CR-5758, Vol. 4). Seattle, WA: Battelle Human Affairs Research Centers.



## CHAPTER 4: LEGAL DRUGS

### Carlyn Orians, and Charlene Forslund<sup>1</sup>

#### 4.1: Introduction

The use and abuse of legal drugs have potentially significant implications for employee fitness for duty in the nuclear power industry. (Although alcohol is a legal drug, this section is limited to a discussion of prescription and over-the-counter (OTC) drugs.) Because so much attention is typically focused on illegal drug use and alcohol abuse, the risks posed by legal drugs are frequently overlooked or underemphasized. This chapter addresses the types of misuse and abuse that can occur; potential impairing effects of legal drugs, alone or in combination with other food or drugs; and trends in the use of legal drugs, particularly the new antidepressants and anabolic steroids.

Legal drugs discussed in this section include both prescription and OTC drugs; however, distinguishing between these types of legal drugs is not useful in assessing the fitness-for-duty (FFD) risks. As reviewed previously (Moore et al., 1989), the classification of a drug as prescription or OTC has very little to do with the drug's potential impairing effects. Many OTC products contain active agents that are similar or identical to prescription drugs; the only difference may be the quantity or strength of the dosage. Hence, in many cases the same impairment would result from taking large quantities of OTC drugs as from taking a prescription drug.

The use, misuse, and abuse of prescription and OTC drugs can create considerable risks to employee fitness for duty. These risks are of particular concern because

- prescription and OTC drugs are commonly used,
- the misuse and abuse of these drugs often goes undetected (sometimes even by the user), and
- too little is known about the impairing effects of legal drug use, since most studies of legal drugs have focused on their toxicity rather than their impairing effects (Moore et al., 1989, p. 3-3).

Two current trends are expected to further increase the use of OTC drugs. First, the Food and Drug Administration (FDA) is approving drugs for OTC distribution at an ever-increasing rate, thereby increasing the number and types available to consumers. Second, stress-producing lifestyles that have led many Americans to seek prescription drugs (especially benzodiazepines) for relief are even more likely to encourage the use (and

misuse) of these cheaper, more readily available OTC drugs (Cooper, 1993). Due to the fact that the general public tends to think of these legal drugs as relatively harmless and clearly distinct from dangerous, illegal drugs, it is particularly important that workers be well informed about what constitutes potential misuse and abuse of these drugs and their potential impairing effects.

It is also important for managers of drug testing programs to be informed of special issues regarding testing for legal drugs. First, designing detection programs is more difficult for drugs that are not included under the Substance Abuse and Mental Health Services Administration (SAMHSA) guidelines for workplace drug testing. Second, the Americans with Disabilities Act (ADA) and other legal protections of employee privacy are particularly applicable to legal drugs (see Chapter 2 [Section 2.4] of this report). (See also Moore et al., 1989, Section 3-3.) In keeping with issues related to employee privacy and the ADA, it is best not to ask employees to divulge their use of legal drugs during the collection process. A Medical Review Office (MRO) should privately discuss any use detected in the testing process with the employee. Use of legal drugs should only be reported to the employer if the use is determined to be in violation of the employer's policy or constitutes a hazard to the individual or to others.

#### 4.2: The Use, Misuse, and Abuse of Legal (Prescription and OTC) Drugs

In assessing the risks posed by legal drugs, it is essential to recognize that problems can result from several modes of use. While the intentional abuse of an illegally-obtained prescription drug exists, the use of legal drugs can produce FFD risks through a number of additional routes. Finch (1993) distinguishes three primary categories of legal (prescription) drug use: appropriate use, misuse, and abuse. All of these categories of use can create an FFD concern; however, the implications regarding FFD policy violations differ across categories. These categories apply equally to OTC drugs.

- Appropriate use is the "use of a medication as [properly] prescribed and only for the condition indicated as appropriate use" (Finch, 1993, p. 232). It is important to note that even when appropriately used (i.e., according to prescription or manufacturer's instructions) some legal drugs pose FFD risks (see discussion of impairment, below).

- Misuse (i.e., inappropriate use) is “the use of a medication for a reason other than that for which it was prescribed or in doses or frequencies other than prescribed . . . ” (Finch, 1993, p. 233). In addition, taking the drug in combination with an inappropriate substance (e.g., alcohol) constitutes misuse. Misuse may be unintentional. Unintentional misuse is more typical among elderly patients and mental health patients. Intentional misuse is differentiated from abuse based on the frequency of occurrence and the effects; misuse, as opposed to abuse, occurs in infrequent, low doses, not sufficient to produce significant dysfunction. Examples of misuse include changing dosage or discontinuing the prescription prematurely without physician permission; self-medicating through the prescription of a friend or relative; and using previously-prescribed, often expired, drugs for a current medical condition (Finch, 1993; National Council on Patient Information and Education, 1987). Other forms of misuse include failing to fill the prescription, taking medication at wrong intervals, forgetting doses, taking doses of incorrect size, and drug-mixing (National Council on Patient Information and Education, 1987). Examples of OTC misuse include taking OTC drugs in larger than recommended dosages, in unsafe combinations with other substances, persistently over a period of time (resulting in dependency or toxic reaction), or as a substitute for another prescribed drug (Barnes et al., 1988, p. 2-27). Dangers of prolonged misuse of OTC drugs include not only complications arising from such use, but also the possibility of such use masking the presence of a serious medical condition (Cooper, 1993). Studies indicate that noncompliance is highest among persons taking medications for chronic diseases. Patients who suffer from chronic decreases have a greater tendency to develop lifelong patterns of prescription misuse (National Council on Patient Information and Education, 1987).
- Abuse is “the use of a medication outside the normally acceptable standard for that drug in such a way that negative consequences occur . . . [in spite of which] . . . the individual continues to use the drug, implying a loss of control over [its] use . . . . Use may be characterized by preoccupation with the drug (including drug-seeking behavior . . . ), escalating use, and a chronic, relapsing course,” which may include increased tolerance and physical dependence (Finch, 1993, p. 233). Physical dependence involves “. . . the precipitation of a withdrawal syndrome when the drug is withdrawn acutely” (this may include increased tolerance) (Finch, 1993, p. 233).

### 4.3: Potential Impairing Effects of Legal Drugs

Some legal drugs can have intended effects or side effects which pose FFD risks in particular jobs within the nuclear industry. When misused or abused, legal drugs—even some of those regarded as benign, such as caffeine—can pose serious safety risks. As described in previous reviews of legal drugs (Moore et al., 1989; Barnes et al., 1988), the primary categories of legal drugs and their potential impairing effects are as follows:

- Central nervous system (CNS) stimulants include amphetamines, synthetic amphetamine-like drugs, caffeine, and OTC appetite suppressants (especially those containing phenylpropanolamine—PPA). CNS stimulants are used to provide increased alertness and confidence, decreased fatigue, mood elevation, and weight loss. With controlled doses of amphetamines, adverse effects on physical performance are minimal. In fact, the stimulation may heighten short-term physical and simple cognitive tasks to a minor degree. Complex cognitive task performance is not enhanced. Larger doses, however, negatively affect cognitive and task performance. Effects include an exaggerated sense of well-being, high energy, restlessness, urgency, over-ideation, and gross temporal distortion. Secondary effects, such as hangovers, rebound depressions, and insomnia, also affect cognitive and task performance (Barnes et al., 1988, p. 2-16).
- CNS depressants include barbiturates and their modern, more popular replacements, benzodiazepines (minor tranquilizers). The latter can be either long-acting (e.g., Valium®, Librium®) or short-acting (e.g., Halcion®, Xanax®). Other sedatives include beta-blocking agents found in hypertension medications and antihistamines found in cough and cold medicines, allergy preparations, anti-motion sickness relievers, and analgesics. Barbiturates are used to provide sedation, sleep inducement, and prevention of epileptic seizures; benzodiazepines are used to treat such conditions as tension, anxiety, and psychosomatic disorders. When used for legitimate purposes and at therapeutic levels, these drugs have been shown to impair sensory functioning, psychomotor performance, short-term memory, learning, and communication skills. In addition, chronic use can lead to dependence. Symptoms of dependence on minor tranquilizers include anxiety, insomnia, agitation, anorexia, tremor, muscle

twitching, nausea, and hypersensitivity to sensory stimuli. Withdrawal from barbiturates can cause severe effects such as delirium and grand mal convulsions (Barnes et al., 1988, pp. 2-21 to 2-24).

Moore et al. (1989) examined the legal drugs that may constitute a fitness-for-duty concern and concluded that, in addition to the illegal drugs included in Section 2.1(a) of Appendix A to 10 CFR Part 26, barbiturates and benzodiazepines were likely candidates for inclusion in the NRC drug testing panel. Later reviews (Durbin et al., 1991) supported this recommendation. A U.S. Department of Health and Human Services (DHHS) consensus report provided mixed recommendations regarding testing for barbiturates and benzodiazepines (U.S. Department of Health and Human Services [DHHS], 1990). This report suggests that these drugs be considered for inclusion in testing protocols when they can be justified as special problems in particular workplace environments. They are not, however, included in the DHHS guidelines at this time. Data gathered from licensee programs indicate that these drugs do pose problems in the nuclear workplace. Durbin et al. (1991) reported that, for the 39 nuclear sites that tested for a broader panel of drugs than required under the NRC rule, all 39 tested for benzodiazepines and 32 tested for barbiturates. At these sites, benzodiazepines tests accounted for 2.9 percent and barbiturates accounted for 2 percent of the confirmed positive tests. These percentages are comparable to the confirmed positive tests found when testing for opiates.

The primary reason for concern about benzodiazepines and barbiturates is the potential for serious misuse and abuse because they are so readily available and are open to self-medication. A number of experts on drug testing in the workplace strongly support the inclusion of these drugs in testing programs, at least in for-cause testing (Jatlow et al., 1987, as referenced in Durbin et al., 1991, p. 99). Furthermore, there is evidence that barbiturates and benzodiazepines are used to cover up a drinking problem; some benzodiazepines such as chlordiazepoxide and diazepam are known to cover the smell of alcohol on one's breath (Jatlow et al., 1987, as referenced in Durbin et al., 1991, p. 99).

#### 4.4: Dangerous Legal Drug Combinations

Combinations of certain legal drugs should be of particular concern because a worker may unknowingly misuse legal drugs by taking the correct dosage of a

prescribed or OTC drug in combination with certain other drugs, foods, or beverages, particularly alcohol (National Council on Patient Information and Education, 1987; Office for Substance Abuse Prevention, 1988).

The legal drug combination of most concern, given its likelihood and severity, is the mixing of certain legal drugs and alcohol. Alcohol use can cause immediate and significant problems when mixed with any of a wide variety of legal drugs including analgesics, sedatives, antidepressants, cold and allergy remedies, anti-infectives, hypnotics (sleep aids), cardiovascular drugs, and diabetes medications (Office for Substance Abuse Prevention, 1988). Problems, usually caused by interference with processing by the liver, are of two types: 1) the liver slows down its processing, thereby allowing alcohol and/or other drugs to accumulate in the body; 2) among alcoholics, the liver continues to process at an increased rate, flushing legal drugs through the body too quickly, thereby negating the medicinal purpose (Office for Substance Abuse Prevention, 1988). Of particular concern is the combination of alcohol and sedatives, especially benzodiazepines and fast-acting barbiturates (Moore et al., 1989, pp. 3-10 and 3-11). In fact, "... the combination of fast-acting barbiturates and alcohol is potentially one of the most harmful combinations of drugs possible" (Moore et al., 1989, p. 3-10). Even small amounts of alcohol, amounts which many individuals would regard as trivial, can be dangerous when consumed in conjunction with certain legal drugs, especially sedatives (Moore et al., 1989, pp. 3-10 and 3-11).

In addition, certain nonalcoholic beverages, food, and tobacco can cause problems. For example, taking an antibiotic with milk can slow down the processing of the drug, thereby reducing its effectiveness. Mixing antidepressant medications known as monoamine-oxidase (MAO) inhibitors (such as Marplan®, Nardil®) with foods containing the chemical tyramine (e.g., Parmesan cheese, red wine, chocolate and bananas) can cause a serious reaction, including death. Tobacco contains over 400 chemicals and can either decrease a prescription drug's effectiveness or increase the possible adverse health consequences of some prescription drugs (e.g., some heart medications, antidepressants, antiasthma medications, and birth control pills).

#### 4.5: Legal Drugs of the "90s"

Recently, there has been a growing concern over two types of legal drugs: the new antidepressants and steroids. The biggest concern with respect to potential FFD risks is the rapidly increasing use of the new

## Legal Drugs

antidepressant drugs, particularly Prozac®. Although steroids are not new, their use has increased in the past decade. The use of steroids, while increasing among professional athletes and even serious amateurs, is probably shorter term and limited to more clearly definable segments of the general population than is the use of Prozac®.

### 4.5.1: Prozac® and Other New Antidepressants

A new group of selective serotonin reuptake inhibiting (SSRI) antidepressants (including paroxetine, sertraline, and fluoxetine) have recently become available in the U.S. The most well-known and frequently prescribed is fluoxetine (e.g., Prozac®). Since its introduction in 1987, Prozac® has received vast media attention and has become one of the most commonly prescribed drugs in the U.S. (Breggin and Breggin, 1994).

There are several reasons for the growing popularity of Prozac® and the other SSRIs. In spite of the fact that recognition of psychiatric illnesses is still poor, major depression has been reported to be more commonly diagnosed than hypertension in primary care practice, affecting about 11 to 36 percent of patients (Guze and Gitlin, 1994). Multiple community studies have estimated that 2 to 4 percent of the population suffers from major depression (Robins et al., 1984; Weissman and Myers, 1978). Although the new antidepressants have not been proven to be more effective than traditional antidepressant agents in treating depression, their side effect profiles (i.e., fewer and less severe side effects) represent a medical advance (Guze and Gitlin, 1994). For this reason, SSRIs are replacing traditional antidepressants, particularly in the treatment of mild to moderate depression. In addition to depression, SSRIs are being used to treat bulimia nervosa, personality disorders, and obsessive-compulsive disorder. These conditions tend to require higher dosages than those prescribed to treat depression. The prevalence of SSRIs makes it extremely important that the effects of these drugs are clearly understood.

There is a growing literature, although still in its infancy, documenting the intended effects and side effects of SSRIs, particularly fluoxetine (Prozac®). Although fluoxetine shares some side effects, such as tremor, drowsiness, and sweating, with traditional tricyclic antidepressants (TCAs), these side effects tend to be less serious and are experienced to a lesser extent (Cooper, 1988). While there has been evidence of overdose fatality, alterations in cardiac conduction, and withdrawal

syndrome with TCAs, fluoxetine is typically not fatal, does not have significant effects on cardiac function, and tends to cause less of a withdrawal problem (Guze and Gitlin, 1994). In addition, fluoxetine does not adversely interact with commonly abused drugs or alcohol (Guze and Gitlin, 1994). Fluoxetine's most likely side effects are symptoms not commonly associated with TCAs (i.e., nausea, nervousness, anxiety, insomnia, and diarrhea).

There has been concern regarding a possible association between some SSRIs, such as fluoxetine, and an increased risk of suicide and violence. A review of this issue by the U.S. Food and Drug Administration's Psychopharmacological Drugs Advisory Committee concluded that no valid evidence of a link between SSRIs and either suicide or violent behavior currently exists (Guze and Gitlin, 1994). However, anecdotal reports persist, and this possible side effect needs further study.

Although further study is warranted, there is no clear evidence as yet that these new antidepressant drugs represent significant FFD concerns. On the other hand, depression does have clear work-related impairments. Depression is associated with more functional disability than most chronic medical illnesses (Wells et al., 1989). The impact of depression on reduced occupational productivity and absenteeism is a serious social issue. Perhaps the most significant problem with SSRIs is not the negative effects, but rather that the positive effects are not guaranteed. There is marked interindividual variability in response to these drugs (Guze and Gitlin, 1994). Even when SSRIs do improve psychological well-being, there is a lack of evidence (primarily due to an absence of research) showing that antidepressants have contributed to improved work outcomes (Mintz et al., 1992).

### 4.5.2: Steroids

This section describes what steroids are, the prevalence of their use, perceived benefits, and potential impairment concerns. In general, there is little evidence to suggest that these drugs are widely abused in the workplace. However, due to reports of increased use of these drugs in the general population, discussions regarding including these drugs in federally mandated FFD programs, and because there is evidence that their use can have severe psychological effects, steroids should be regarded as potential risks. In addition, because users don't consider themselves "drug abusers," the effects of these drugs should be communicated to both employees and employers.

Anabolic steroids are synthetic forms of the male sex hormone testosterone. They are used for such legitimate medical purposes as rebuilding body tissue following surgery, infection, or drastic forms of weight loss. They are also beneficial in treating men who are deficient in testosterone. Anabolic steroids are more accurately termed "anabolic-androgenic steroids" for their two primary and, in practical use, inseparable functions: *anabolic* means "growth-promoting" or "tissue-building," while *androgenic* means "masculinizing" (Su et al., 1993, p. 2760; Wadler and Hainline, 1989, as cited in Daigle, 1990, p. 77; Kochakian, 1990). There have been some reports of a recent shift toward the human use of veterinary steroids, which are typically used in race horses. These steroids are more anabolic and less androgenic, thereby promoting a more lean, muscular physique; androgens tend to be more fat-producing (Goldstein, 1990).

While anabolic steroids have legal, beneficial medical uses, these substances are controlled as a result of their potential negative effects. The federal government, through the U.S. Customs Service, Justice Department, and FDA, is attempting to counter the black market for steroids (*MRO Newsletter*, August 1993). "One sports medicine authority . . . has indicated as much as 70 to 80 percent of the steroids produced in the United States are not used for legitimate medical purposes" (Office for Substance Abuse Prevention, 1989, p. 2). Athletes may obtain steroids illegally through coaches, trainers or promoters. Users in the general public may obtain them through dealers of illegal drugs or personal trainers in health gyms (Goldstein, 1990). Illicit steroids tend to be taken at high, "suprapharmacological" doses, far greater than beneficial prescribed doses. As a result, side effects are harmful and more detectable than would be the case with prescribed doses (Cicero and O'Connor, 1990).

The amount of steroids used and the cost are quite variable. According to Goldstein (1990, pp. 84-85):

serious long-term users may spend as much as \$200 to \$400 per week on steroids and the accompanying pharmacopoeia. Since such users may go on cycles of steroid use lasting 12 to 14 weeks, each cycle can cost thousands of dollars . . . . Younger, or less experienced or committed users, will use considerably less, perhaps only one injection per week. Many serious, long-term users are said to fund their use through dealing steroids themselves (pp. 84-85).

Illicit use of steroids typically runs in cycles lasting about 14 weeks beginning with oral (pill) administration and eventually including injections. The amount of use peaks at about nine weeks, after which use decreases until the end of the cycle (Goldstein, 1990).

Illicit steroid use is most often associated with competitive athletes who are interested in improving athletic performance through increased muscle bulk, strength, and endurance obtained by steroid use. Because of the variety of alleged physical benefits, steroid use has surfaced in such various sports as weightlifting, bodybuilding, track and field, football, and swimming (Lombardo, 1990). (For a summary of survey literature on use prevalence among athletes, see Yesalis et al., 1993.)

Steroid use is also spreading beyond competitive sports to a more general market of ambitious exercisers ("look-gooders") typically found in health gyms, who seek alleged benefits of steroids for physique enhancement (Lombardo, 1990, p. 61). It is suspected that many of these users are attempting to improve their physique to compensate for low self-esteem and insecurity (Lombardo, 1990, p. 61; Goldstein, 1990). In other cases, male steroid users are seeking improved physiques for employment as bouncers, nightclub dancers, models, and homosexual prostitutes (Goldstein, 1990, p. 85).

The extent of steroid use among the general public is becoming more difficult to ascertain because of the increasingly severe penalties for illegal distribution and use of these substances (e.g., the federal Anti-Drug Abuse Act of 1988) (Goldstein, 1990, p. 74). Undocumented estimates (see Goldstein, 1990, p. 75, for specific sources) suggest that there are one million regular steroid users in the U.S. However, based on self-report data from the National Household Survey on Drug Abuse (NHSDA), steroid use is very low. The NHSDA figures show that less than 0.2 percent of the surveyed population in 1991 and 0.1 percent in 1992 reported use of steroids in the last year. A negligible number of individuals reported that they had "used steroids in the last month" in 1992 or in 1993 (SAMHSA-NHSDA92, 1993; SAMHSA-NHSDA93, 1994).

Steroid users often report that they view the use of these drugs as legitimate, "non-street" drugs having health-promoting consequences. Regular users report that steroids can produce euphoria, a sense of well-being,

## Legal Drugs

increased libido, aggression, improved self-image, and, most commonly, an enhancement of athletic performance and muscular development (Haupt and Rovere, 1984, Lombardo, 1990, Taylor, 1987, Wilson, 1988, Wilson and Griffin, 1980, all referenced in Cicero and O'Connor, 1990). Other reasons for steroid use include competitive pressure among athletes and peer pressure among adolescents (Fuller and LaFountain, 1987, Johnson et al., 1989, Taylor, 1987, Wilson, 1988, all referenced in Cicero and O'Connor, 1990, p. 11).

Despite common folklore, the actual benefits of steroid use in enhancing the physique and athletic performance are still being debated in the research literature (both animal and human). Lack of definitive findings are in large part due to a paucity of well-controlled studies (see literature summaries in Cicero and O'Connor, 1990; Lombardo, 1990; and Zemper, 1991). Still, after surveying the research literature and the "overwhelmingly consistent anecdotal reports of the athletes [studied]," one author surmised that there was convincing evidence of some benefits, especially under specific circumstances:

... it can be concluded that androgens are associated with increases in strength and lean body mass, especially in experienced weight trainers, when performing high-intensity work in the presence of an adequate diet (Lombardo, 1990, p. 69).

The extent of such benefits, however, can vary greatly across users (Zemper, 1991, p. 125). Also, it is possible that steroid use has a "placebo effect" in which users, expecting to improve their workouts or performance, actually do so due to increased self-confidence or motivation rather than properties of the steroids.

Former users who were surveyed reported two primary reasons for stopping use: 1) "maturing" into a more conventional lifestyle, and 2) deterioration of health due to steroid use (Goldstein, 1990, p. 82).

Testing for steroids originated in competitive sports. The U.S. Olympic Committee, the International Olympic Committee, the National Football League (NFL) and the National Collegiate Athletic Association (NCAA) currently test for steroids (*MRO Newsletter*, August 1993). Urine tests do not look for the anabolic steroids themselves, but rather for all of the metabolites or metabolic breakdown products of the steroid that are excreted in the urine. There are at least 20 different anabolic steroids available, and each steroid compound can have as many as a dozen or more metabolites.

Detecting steroid use through testing is facilitated by the relatively lengthy detection time: three to four weeks for oral administration and six to eight months for injection (Martin, 1991). Steroid tests use gas chromatography and mass spectrometry (GC/MS), and when properly done are essentially 100 percent accurate (Zemper, 1991, p. 127).

Potential impairment effects from the use of steroids include both physiological and psychological damage. There are many documented or alleged potential physiological effects of steroid use and abuse, some of which are life-threatening. Effects may include higher blood pressure, an increased risk of heart attack, enlargement of the prostate, liver cancer, jaundice, testicular atrophy (often leading to infertility and diminished sexual desire), and gynecomastia (abnormal breast development) in males (Office for Substance Abuse Prevention, 1989, p. 2-3). Other studies report the following effects:

- "purple or red-colored spots on the body, unexplained darkening of the skin, bad breath, increased acne" (*MRO Newsletter*, August 1993, p. 1).
- "liver problems . . . kidney problems, hypertension . . . sexual problems in males (testicular atrophy, decreased sperm production, gynecomastia) . . . physical injuries, cholesterol difficulties, cardiovascular problems, stunted growth in adolescents, male pattern baldness . . . gallstones" (Goldstein, 1990, p. 90).
- "endocrine disturbances, particularly in adolescents; cardiac and liver problems; stunted growth; testicular atrophy" (Cicero and O'Connor, 1990, p. 10).

Some effects of steroid use are specific to women; these include easily perceivable, and often irreversible, adoption of male secondary sex characteristics, such as deepened voice, facial hair, increased aggressiveness, menstrual irregularities, hypertrophy of the clitoris, and fetal damage (Strauss, Liggett and Lanese, cited in Office for Substance Abuse Prevention, 1989, p. 3; Goldstein, 1990, p. 90).

In addition to physiological effects, psychological effects have also been reported. As reported by the Office for Substance Abuse Prevention (1989, p. 3), "Steroid-related personality changes may result in inappropriate anger, irrationally destructive behavior, and 'flying off the handle' in response to even minor frustrations." This aggressive behavior has been referred to as "roid rage" (*MRO Newsletter*, August 1993, p. 1). Descriptions of

"roid rages" are prevalent in body-building folklore and, more recently, in the defense pleading of steroid users committing violent crimes.

Although reports of violent, antisocial reactions to prolonged, high-dosage use of steroids has increased concern over the use of these drugs, scientific research on the subject of psychological steroid impairment is inconclusive. One reason is that research on the effects of steroids has tended to be based on laboratory studies that use prescription doses of steroids, rather than the large doses favored by abusers of the drugs (e.g., bodybuilders) (Katz and Pope, 1990). A review of this research literature found mixed results: some studies found significant mood swings toward increased aggressiveness and irritability, while others found no such changes (Katz and Pope, 1990). In a survey of athletes using steroids, Pope and Katz (1988) found that about one-third of the 41 subjects interviewed reported symptoms resembling a manic syndrome (based on criteria in the revised third edition of the *Diagnostic and Statistical Manual of Mental Disorders* [DSM-III-R]). In every case, these symptoms disappeared quickly when steroid use ceased. A more recent cohort study of 88 athletes who were using steroids at very high doses and 68 nonusers of steroids found that "23 percent of the steroid users reported major mood syndromes, including mania, hypomania, or major depression, in association with steroid use" (Pope and Katz, 1994).

It is also unclear whether steroids lead to physical dependence or interact with other drugs. Although steroids are not considered physically "addictive," heavy users may become either physically or psychologically dependent on them (Office for Substance Abuse Prevention, 1989). There is some evidence that steroid use "increases in frequency and amount in chronic users" (Taylor, 1987, as referenced in Cicero and O'Connor, 1990); however, this may be a voluntary increase due to a belief in greater benefits from increased dosages and frequencies rather than caused by physical dependence (Cicero and O'Connor, 1990, p. 13). In a limited number of studies, subjects undergoing immediate withdrawal from steroid use reported withdrawal symptoms such as depression, decreased libido, and fatigue (Cicero and O'Connor, 1990). With respect to drug interactions, research into whether steroid use influences the effects of

other drugs (e.g., mediating the effects of alcohol or opiates) is too preliminary to draw conclusions (Cicero and O'Connor, 1990; Office for Substance Abuse Prevention, 1989). However, at least one recent self-report survey stated that adolescent steroid users are more likely than nonusers to use other, seemingly unrelated drugs such as marijuana, smokeless tobacco, and cocaine, and to share needles for drug use (DuRant et al., 1993). This may imply that use of steroids is associated with other risk-taking behaviors.

In summary, more research needs to be directed at the possible impairment effects of steroid use. Because steroid users often view their drug use as being short term, research and employee awareness programs that stress the immediate side effects of anabolic steroids (such as impotence, male pattern baldness, and possible aggressive tendencies) and the possible long-term consequences of short-term use may be most effective in curbing steroid abuse.

## 4.6: Conclusion

As noted in the introduction to this chapter, although most of the attention in fitness-for-duty programs has been focused on illegal drug abuse, and, to a lesser extent, alcohol abuse, legal drugs may be of equal or greater importance. The risks posed by legal drugs are substantial because these drugs are widely used and their impairing effects are often not foreseen. Individuals who would never consider consuming alcohol before reporting to work may be unaware that taking high doses (or even standard doses) of over-the-counter or prescription medications may pose similar impairment risks.

This chapter has given a general overview of the risks associated with the use, misuse, and abuse of legal drugs. In addition, it has reviewed two types of drugs that have received considerable media and/or regulatory attention in more detail—Prozac® and other new antidepressants and steroids. In general, there is little evidence to suggest that either of these two drug types pose significant fitness-for-duty concerns in the nuclear industry at this time. However, it seems prudent that workers be made aware of the potential impairment effects, and, in the case of steroids, the significant potential health risks, associated with these types of drugs.

## **Endnotes**

1. Carlyn Orians, M.S., Battelle Seattle Research Center  
Charlene Forslund, M.S., Battelle Seattle Research Center

## References

10 CFR Part 26, January 1, 1990. U.S. Nuclear Regulatory Commission. Fitness-for-duty programs. *U.S. Code of Federal Regulations*.

Barnes, V., Fleming, I., Grant, T., Hauth, J., Hendrickson, J., Kono, B., Moore, C., Olson, J., Saari, L., Toquam, J., Wieringa, D., Yost, P., Hendrickson, P., Moon, D. and Scott, W. (1988). *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues* (NUREG/CR-5227). Washington, DC: Nuclear Regulatory Commission.

Breggin, P. and Breggin, G. (1994). *Talking Back to Prozac®: What Doctors Wouldn't Tell You About Today's Most Controversial Drug*. New York, NY: St. Martin's Press.

Cicero, T.J. and O'Connor, L.H. (1990). Abuse liability of anabolic steroids and their possible role in the abuse of alcohol, morphine, and other substances. In G.C. Lin and L.E. Erinoff (Eds.) *Anabolic Steroid Abuse*, NIDA Research Monograph Series #102: 1-28. Rockville, MD: U.S. Department of Health and Human Services.

Cooper, G.L. (1988). The safety of fluoxetine—an update. *British Journal of Psychiatry*, 153(3): 77-86.

Cooper, H. (1993, January 11). Warning: Americans overuse over-the-counter drugs. *Wall Street Journal*.

Daigle, R.D. (1990). Anabolic steroids. *Journal of Psychoactive Drugs*, 22(1): 77-80.

DuRant, R.H., Rickert, V.I., Ashworth, C.S., Newman, C. and Slavens, G. (1993). Use of multiple drugs among adolescents who use anabolic steroids. *The New England Journal of Medicine*, 328(13): 922-926.

Durbin, N., Moore, C., Grant, T., Fleming, T., Hunt, P., Martin, P., Murphy, S., Hauth, J., Wilson, R., Bittner, A., Bramwell, A., Macaulay, J., Olson, J., Terrill, E. and Toquam, J. (1991). *Fitness for Duty in the Nuclear Power Industry: A Review of the First Year of Program Performance and an Update of the Technical Issues* (NUREG/CR-5784). Washington, DC: Nuclear Regulatory Commission.

Finch, J. (1993). Prescription drug abuse. *Primary Care*, 20(1): 231-239.

Goldstein, P.J. (1990). Anabolic steroids: An ethnographic approach. In G.C. Lin and L.E. Erinoff (Eds.) *Anabolic Steroid Abuse*, NIDA Research Monograph Series #102: 74096. Rockville, MD: U.S. Department of Health and Human Services.

Guze, B.H. and Gitlin, M. (1994). New antidepressants and the treatment of depression. *The Journal of Family Practice*, 38(1): 49-57.

Katz, D.L. and Pope, H.G., Jr. (1990). Anabolic-androgenic steroid-induced mental status changes. In G.C. Lin and L.E. Erinoff (Eds.) *Anabolic Steroid Abuse*, NIDA Research Monograph Series #102: 215-223. Rockville, MD: U.S. Department of Health and Human Services.

Kochakian, C.D. (1990). History of anabolic-androgenic steroids. In G.C. Lin and L.E. Erinoff (Eds.) *Anabolic Steroid Abuse*, NIDA Research Monograph Series #102: 29-59. Rockville, MD: U.S. Department of Health and Human Services.

Lombardo, J.A. (1990). Anabolic-androgenic steroids. In G.C. Lin and L.E. Erinoff (Eds.) *Anabolic Steroid Abuse*, NIDA Research Monograph Series #102: 60-74. Rockville, MD: U.S. Department of Health and Human Services.

Martin, D.M. (1991). *Sport Drug Testing in the 1990's*. Paper presented at the Drug-Free Workplace Conference-Exposition, Washington, DC, October 24-25.

Mintz, J., Mintz, L.I., Arruda, M.J. and Hwang, S.S. (1992). Treatments of depression and the functional capacity to work. *Archives of General Psychiatry*, 49: 761-768.

Moore, C., Barnes, V., Hauth, J., Wilson, R., Fawcett-Long, J., Toquam, J., Baker, K., Wieringa, D., Olson, J. and Christensen, J. (1989). *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues* (NUREG/CR-5227, Supplement 1). Washington, DC: Nuclear Regulatory Commission.

MRO Newsletter. (1993, August). Steroids, 2(8): 1.

## Legal Drugs

National Council on Patient Information and Education. (1987). Prescription drug misuse: Many kinds, many costs. *Talk About Prescriptions Month*. Washington, DC: Author.

Office for Substance Abuse Prevention. (1988, September). *The Fact Is...It's Dangerous to Drink Alcohol While Taking Certain Medications*. Rockville, MD: National Clearinghouse for Alcohol and Drug Information.

Office for Substance Abuse Prevention. (1989, September). *The Use of Steroids in Sports Can be Dangerous* (fact sheet MS391). Rockville, MD: National Clearinghouse for Alcohol and Drug Information.

Pope, H.G., Jr. and Katz, D.L. (1988). Affective and psychotic symptoms associated with anabolic steroids. *American Journal of Psychiatry*, 145(2): 487-490.

Pope, H.G., Jr. and Katz, D.L. (1994). Psychiatric and medical effects of anabolic-androgenic steroid use: A controlled study of 160 athletes. *Archives of General Psychiatry*, 51: 375-382.

Robins, L.W., Helzer, J.E., Weissman, M.M., Orvaschel, H., Gruenberg, E., Burke, J.D. and Regier, D.A. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry*, 41: 949-958.

SAMHSA-NHSDA92. (1993, June). *Preliminary Estimates from the 1992 National Household Survey on Drug Abuse: Selected Excerpts*. Washington, DC: SAMHSA, Office of Applied Studies, National Household Survey on Drug Abuse.

SAMHSA-NHSDA93. (1994, July). *Preliminary Estimates from the 1993 National Household Survey on Drug Abuse*. Washington, DC: SAMHSA, Office of Applied Studies, National Household Survey on Drug Abuse.

Su, T.P., Pagliaro, M., Schmidt, P.J., Pickar, D., Wolkowitz, O. and Rubinow, D.R. (1993). Neuropsychiatric effects of anabolic steroids in male normal volunteers. *Journal of the American Medical Association*, 269(21): 2760-2764.

United States Department of Health and Human Services (1990). Technical, Scientific, and Procedural Issues of Employee Drug Testing: Consensus Report (DHHS ADM #90-1684). Washington, DC: United States Department of Health and Human Services.

Weissman, M.M. and Meyers, J.K. (1978). Affective disorders in a U.S. urban community: The use of research diagnostic criteria in an epidemiologic study. *Archives of General Psychiatry*, 35: 1304-1311.

Wells, K.B., Stewart, A., Hays, R.D., Burnam, A., Rogers, W., Daniels, M., Berry, S., Greenfield, S. and Ware, J. (1989). The functioning and well-being of depressed patients: Results from the medical outcomes study. *Journal of the American Medical Association*, 252(7): 914-919.

Yesalis, C.E., Anderson, W.A., Buckley, W.E. and Wright, J.E. (1990). Incidence of the nonmedical use of anabolic-androgenic steroids. In G.C. Lin and L.E. Erinoff (Eds.) *Anabolic Steroid Abuse*, NIDA Research Monograph Series #102: 97-112. Rockville, MD: U.S. Department of Health and Human Services.

Yesalis, C.E., Kennedy, N.J., Kopstein, A.N. and Bahrke, M.S. (1993). Anabolic-androgenic steroid use in the United States. *Journal of the American Medical Association*, 270(10): 1217-1221.

Zemper, E.D. (1991). Drug testing in athletics. In R.H. Coombs and L.J. West (Eds.) *Drug Testing: Issues and Options*: 113-139. New York, NY: Oxford University Press.

# CHAPTER 5: THE DETERRENT EFFECT OF FITNESS-FOR-DUTY PROGRAMS

Joseph Hauth<sup>1</sup>

## 5.1: Introduction

After several years of experience in implementing workforce fitness-for-duty (FFD) programs in military and industrial settings, current attention is focused on how effective these programs have been in deterring illegal drug use and abuse of legal drugs such as alcohol and prescription drugs. Many proponents of workplace drug testing believe that random, unannounced testing has been singularly effective in deterring drug use by making it impossible to predict when testing will occur (Bray et al., 1992; Barnes et al., 1988; Moore et al., 1989; Institute for a Drug-Free Workplace, 1991).

Questions have been raised, however, about the overall effectiveness and costs of FFD programs in dealing with employee drug use (Normand, Lempert and O'Brien, 1994; Walsh, Elinson, and Gost, 1992; DOT, 1993; Harris, 1993; Hartman and Crow, 1993; Newcomb, 1992; West and Ackerman, 1993). In particular, little is known about how effective FFD programs are in deterring—rather than merely detecting and removing—drug users. Further, concerns have been raised about the imposition of random testing on nondrug users as well as on those who use drugs, the high costs of random testing, workforce disruption when employees are called away for testing, and the stringent administrative controls needed to ensure the accuracy and fairness of the testing process (DOT, 1993; AMA, 1991; LeRoy, 1991; LeRoy, 1990).

In response to some of these questions, initiatives have been taken by the Department of Transportation (DOT), the NRC, and other federal agencies to lower their random testing rates in order to reduce the costs of random testing while maintaining an assumed deterrent effect (DOT, 1993; *Drug Detection Report*, 1993). For instance, after reviewing the performance of random testing in the U.S. nuclear power industry, the NRC lowered the annual random testing rate from 100 percent to 50 percent for workers with unescorted access to the protected areas of licensed nuclear facilities (59 FR 502, January 5, 1994). Given continuing questions about optimal levels of testing and the deterrent effect of FFD programs in general, it is helpful to review what is known about the deterrent effect of random drug testing and of the other elements of FFD programs such as testing, sanctions, employee training, employee assistance, and normative or organizational factors that may influence drug use.

First, deterrence theory is reviewed and discussed in the context of FFD programs. Following this discussion, available evidence is reviewed from other federal FFD programs and private industries concerning the deterrent effect of FFD program elements.

## 5.2: Deterrence and Drug Use

In FFD programs, the primary means used to increase the perceived risk of drug use, and thus to deter it, has been to increase both the probability of detection and sanctions. This approach is consistent with the classical notion that deterrence is a function of the perceived risk of being detected for the illicit activity, the severity of the sanctions involved, and the swiftness with which the sanctions are applied in comparison with the gratification derived from the activity in question (Leatherwood and Spector, 1991; Piliavin et al., 1986; Slovic et al., 1988; Howe and Brandau, 1988; Williams, 1985; Evans, Neville, and Graham, 1991; Buikhuisen, 1974). This is referred to as the rational model of deterrence. The risk taker makes an assessment of likely outcomes and modifies his or her behavior accordingly to maximize gratification and minimize undesired outcomes.

A number of industry drug testing programs have policies that are consistent with the rational model. In general, FFD programs are comprised of an integrated set of FFD program elements designed to maximize detection and sanctions for drug use (Barnes et al., 1988; Moore et al., 1989; AMA, 1991; LeRoy, 1991; LeRoy, 1990; Crant and Bateman, 1990; Murphy, Thornton, and Reynolds, 1990). FFD program elements that are believed to contribute to deterrence include pre-employment screening, skilled behavioral observation, random testing, for-cause testing, sanctions for substance use that are well-known and consistently applied, workplace security measures, policy and awareness training, health promotion programs, employee assistance programs (EAPs), and, more recently, measures to prevent subversion of the testing process (see Chapter 6). As discussed in greater detail below, each of these program elements is believed to influence deterrence in a variety of ways. For instance, the effectiveness of random testing may influence whether a person decides to use drugs. Training and education form the decision maker's perception of risk. An EAP may provide support to the occasional user in ceasing drug use altogether and indicates that the organization supports drug use prevention.

## Deterrence

There is support for the notion that perceptions of risk play a critically important role in deterring drug use. Perceived risks include the risk of being tested, the risk of drug use being detected and penalized, the health and safety risks inherent in drug use, and job loss or other employment sanctions. The perceived risks and health effects of drug use are expected to vary by drug type, situational context, and the sanctions involved (Bray et al., 1992; Bachman, Johnston, and O'Malley, 1990; Hall, Flaherty, and Homel, 1992; Schwartz, 1991; U.S. Department of Education, 1990; Gonzalez and Haney, 1990; Lauderback and Waldorf, 1993; Cromwell et al., 1991; Barnes et al., 1988; Moore et al., 1989; Hurst, 1985; Waldo and Chiricos, 1972).

Although risk perceptions are important in deterring drug use, a better understanding of the cognitive processes underlying decisionmaking would improve the rational model of deterrence. For instance, perceived levels of gratification may have a greater influence than the risk of sanctions on decisions to commit illegal acts, especially if the subject has difficulty assessing the probable risk of sanctions (Piliavin et al., 1986; Slovic et al., 1988). One study found that variables that appeared to most inhibit marijuana use were age, fear of the physical consequences of use, and the belief that drug use was immoral; criminal sanctions had little effect on deterring use (Williams, 1985). This finding is supported by others who have found that health risks, peer approval, and moral condemnation, rather than the threat of sanctions, are strong motivations for refraining from marijuana and cocaine use (Bachman, Johnston, and O'Malley, 1990; Burkett and Ward, 1993). The effect of education—important for informing the risk taker's decision making—has had mixed results in reducing repeated offenses for drunk driving and drug use in the military (Little, Robinson, and Burnette, 1990; Kingsnorth, 1991; Stoloff, 1985). Further investigation is needed to better understand the effects of rapid punishment, certainty of punishment, and severity of punishment on deterrence, and how they interact (Howe and Brandau, 1988; Leatherwood and Spector, 1991; Piliavin et al., 1986). Finally, in order to understand deterrence, it is important to understand the kinds of persons involved, and their values regarding the illicit activity (Waldo and Chiricos, 1972). In considering drug use, and deterrence in particular, it is important to understand drug use prevalence in the target population, i.e., which drugs are used, preferences, and habits (Lauderback and Waldorf, 1993).

Research on human decisionmaking and risk assessment further suggests that people's perceptions of the risk of

being tested and having their drug use detected are not necessarily based on rational calculations of probabilities. Human decision making is known to reflect only a "bounded rationality"<sup>2</sup> (Simon, 1956). That is, it differs from the mathematical ideal for a number of reasons that include: information processing limitations, solution heuristic artifacts, and some preference for certainty over maximized utility (Fishburn, 1988; Slovic et al., 1988). These factors lead to several decision phenomena that affect the deterrent value of FFD programs:

*Availability Heuristic.* Probability judgments are based on the ease with which they come to mind (Slovic et al., 1988, p. 678). This would suggest that employees will tend to overestimate the probabilities of well-publicized extreme events (such as termination for drug use).

*Calibration Distortions.* People tend to increasingly and substantially underestimate the probabilities of outcomes as actual probabilities approach one (100 percent certainty); conversely, they increasingly overestimate the probability of outcomes as actual probabilities approach zero (0 percent likelihood) (Slovic et al., 1988, p. 678). This would suggest, for example, that employees would substantially overestimate the likelihood of a low probability event such as being randomly tested in a particular week when only 1/52 of the population is being tested.

*Representativeness Heuristic.* People tend to be led astray in judging the representativeness of events by either attending to irrelevant features or not attending to relevant features (Slovic et al., 1988, p. 681). For instance, a worker who has heard an anecdotal account of a drunk driver recently being tested for drug use may recall that incident foremost in making his decision whether to refrain from drug use on the job, although the two events are unrelated.

Additional factors affect drug use in addition to risk assessments of probable outcomes. Peer pressure, drug culture expertise, family support, organizational support, emotional fatigue, physical or mental stress on the job or at home, addiction, and intrinsic motivations for drug use may affect drug use decisions (Barnes et al., 1988; Williams, 1985; Curry, Wagner, and Grothaus, 1990; Kelly, Zyzanski, and Alemagno, 1991; Eggert and Herting, 1991; Raskin, Novacek, and Hogan, 1992; Harris and Heft, 1992; Connors, 1992; Kaplan and Fukurai, 1992). Consequently, some people use drugs regardless of—or in spite of—immediate risk.

Given what is known about people's risk perceptions regarding the health, safety, and employment risks of drug use, as well as peer, family, and organizational influences, three strategies are suggested by this review to increase deterrence in a FFD program:

1. Increase the perception of risk of detection and employment sanctions for drug use by (a) increasing the probability of being tested through pre-employment testing, for-cause testing, and random testing; (b) increasing the probability of a positive test result by using lower, more stringent testing cutoff levels; (c) decreasing the probability of subversion of the testing process; and (d) increasing the probability that sanctions will be applied by maximizing the technical defensibility of the program, maximizing the accuracy of the results, and ensuring that sanctions are suitable and consistently applied.
2. Increase the perception of health, family, and safety risks of drug use through education, training, and awareness.
3. Increase the risk of peer disapproval by increasing peer pressure against drug use (normative effect) through increased drug awareness training, and encouragement of peer referral programs.

Over whom should be the focus of deterrence efforts in FFD programs? There are three types of workers to consider:

- Workers who do not and will not use illegal drugs or abuse alcohol; these people do not commit any acts that need to be deterred and thus any efforts aimed at deterrence are not applicable to this category of worker; these workers are believed to comprise the majority of workers subject to FFD programs.<sup>3</sup>
- Workers who occasionally use illegal drugs or abuse alcohol; these people may be deterred from occasional use that constitutes a violation of FFD policy.
- Workers who chronically use illegal drugs or abuse alcohol; they cannot be deterred by the threat of sanctions because of their habitual use, dependency on drugs, or disregard altogether for the FFD policies.

Workers in the first category are not deterred *per se* since they do not commit the proscribed actions. Workers in

the second category, occasional users, are truly the focus of efforts aimed toward maximum deterrence. That is, these workers knowingly commit a proscribed action and they are able to refrain from that action, either by rational appeals to personal health, workplace safety and productivity, or because the threat of sanctions is strong enough to curb the illicit activity. However, it is recognized that some occasional users will not be deterred from drug use because the threat of punishment is not compelling, or because of other influencing factors such as stress or peer pressure. Workers in the third category cannot be deterred. These workers should be the focus of efforts aimed at detection rather than deterrence, since they are not amenable to rational appeal or the threat of sanctions. An FFD program should be designed to detect, remove, and/or rehabilitate this third class of worker rather than trying to deter them.

It is unlikely that these three classes of workers are static: a nonuser may become an occasional or chronic user. Likewise, an occasional user may, through circumstance, become a chronic user because of increasing dependency on drugs. An occasional or chronic user, once rehabilitated, may become a nonuser altogether, regardless of the presence of a drug testing program. The former user may relapse. Consequently, there will always be a distribution of the three types of worker in a workforce population.

### 5.3: The Deterrent Effect of FFD Programs

This section reviews what is known about the deterrent effect of FFD programs. Data on FFD program effectiveness in various industries are available but limited. In fact, little is known about the deterrent effect of drug testing programs in the workplace (Normand, Lempert and O'Brien, 1994). Studies have concentrated on examining the rates of drug use detected (usually expressed as the rate of positive test results in a testing population) in industries with comprehensive FFD programs such as the military, transportation, and nuclear power industries (Bray et al., 1992; Bray et al., 1990; Osborn and Sokolov, 1990; Cohen, 1986; Stoloff, 1985). National surveys conducted by the American Management Association (AMA) (1991, 1992, 1993, and 1995) found that positive test results for drug use among surveyed firms continue declining while the number of firms with random drug testing provisions increased. Although the decline in positive drug test rates is attributed in part to the implementation of random testing procedures, it is also attributed to increased education and awareness

policies, as required by federal regulatory agencies. Declines in drug use detection rates are also attributed to larger societal trends (Bray et al., 1992). Similar to results found in the military, FFD programs have shown declining drug use trends in the nuclear power industry (Osborn and Sokolov, 1989), and in the railroad industry (Taggart, 1989).

Anonymous surveys of military personnel have found that drug users believe that urinalysis testing had a deterrent effect on their drug using behavior (Bray et al., 1992; Cohen, 1986). Standardized comparisons of drug and alcohol use among military personnel and civilians support this conclusion (Bray, Marsden, and Peterson, 1991). Military personnel and civilian data sets were matched for sociodemographic characteristics (e.g., age, sex, race, education) and the geographic location of respondents. The study by Bray, Marsden, and Peterson found that military personnel were significantly less likely to use drugs than their civilian counterparts. These findings suggest that the military's FFD programs are effective in reducing drug use (Bray et al., 1992). Their overall effect, however, is not clear. Declining rates of drug use in the military may also be attributed in part to changing military workforce demographics. Compared to the 1970s and 1980s, in 1990 the military population was older, had more officers and more married personnel, and had better educated personnel; all of which are associated with lower levels of drug use (Bray et al., 1990). Finally, observed declines in drug use in the military may be attributed in part to larger societal trends showing declines in illegal drug use. However, similar to earlier results found in the military, the 1992 study by Bray and others also found that military personnel were significantly more likely to abuse alcohol. This may be due to the lack of testing for alcohol in the DOD program and unique cultural factors in the military that encourage drinking.

The following discussion reviews what is known about the deterrent effect of FFD program elements.

### 5.3.1: Random Testing

Random testing is a central component of many FFD programs in safety-sensitive industries. The desired result of random testing is two-fold: (a) a deterrent effect evident in continued abstinence from drugs, and (b) detection of those who continue using drugs despite random testing and the presence of other FFD program elements. Random testing is believed by many to provide a significant deterrent effect because it is unpredictable; people simply do not know when they will be tested and

consequently cannot adjust their drug use to prevent detection. Random testing consists of three key components:

- a sampling method to determine who has been selected for testing out of the population in the testing program. All persons in the testing program should have an equal probability of being selected.
- a rate of selection, or the proportion of the population selected in a given time period. Examples include selecting a number of people for testing equal to all of the targeted population in a one year period (100 percent) or selecting a number equal to half of the population in one year (50 percent).
- a frequency of selection, or the number of times testing occurs within a given period, e.g., daily, weekly, bi-monthly, monthly, once yearly, semi-annually, etc. On a daily selection basis, for instance, the number of workers in the population (e.g., 2,000) would be divided by the number of working days (e.g., 365), or five people a day. On a weekly testing basis, 1/52 of the workforce population would be selected for testing once every week. The frequency of testing does not necessarily correspond with the frequency of selection. Testing frequency often depends on the testing schedule in terms of shift schedules, how many tests can be performed daily, arrangements with the specimen collection and laboratory analysis contractors, and other worksite factors.

Two more attributes of a random testing program impact the effectiveness of the program:

- Assurance that the program cannot be subverted; testing neither detects nor deters substance abuse if the testing process is easily subverted.
- A notification method that assures the test is unannounced and limits the time between notification and testing.

When FFD programs were established, many FFD program developers recognized that random testing should be imposed in such a manner that is unpredictable. They did not know, however, the particular rate of testing that would provide a strong deterrent effect without being overly burdensome. Was a 100 percent testing rate adequate? Should it be higher or lower? The deterrent effect of other testing practices, i.e., the ideal method and frequency of selection, were also not known.

A review of studies of large workforce drug use programs such as those required by DOT and DOD revealed that no research exists that directly addresses the relationship of random testing and deterrence. Hence, there appears to be very little, if any, conclusive scientific evidence regarding the deterrent effect of random testing. Since little is known about the relationship between random testing and deterrence, federal agency FFD program practices vary considerably.

Early studies sponsored by DOD found that increased testing may increase overall *detection* of drug use in the workforce (Stoloff, 1985).<sup>4</sup> However, continued drug use by identified users (recidivism) was shown to be a substantial factor in overall drug use rates. This suggested that a substantial number of those testing positive for drugs were not deterred. Instead, heavy or chronic drug users were merely removed and replaced, presumably with those less inclined to use drugs (Osborn and Sokolov, 1990; Stoloff, 1985). Increasing random testing rates and discharge rates for repeat users lowered the overall prevalence rate of drug use by simply removing repeat offenders. Again, this is a function of detection, not deterrence.

### 5.3.1.1: Random Testing Rates and Detection of Drug Use

In reviewing federal agency random testing practices, the General Accounting Office (GAO) found that the random testing rates vary widely from an annual testing rate of 4 percent to 200 percent (GAO, 1991). Agencies with random testing rates of 5 percent or less include the U.S. Department of Commerce, the U.S. Mint, the Immigration and Naturalization Service (INS), the Federal Bureau of Prisons, the Federal Bureau of Investigation (FBI), and the National Labor Relations Board (NLRB). Agencies with random testing rates of 100 percent or more include the U.S. Navy, the Department of Education, the General Services Administration (GSA), and the National Endowment for the Humanities. Further, some agencies vary the rate of testing. For example, the Selective Service System varies their random testing rate from 28 percent to 100 percent of the sample population. This significant variation indicates a lack of consensus regarding the optimum random testing rate and how random testing should be administered.

In establishing random testing rates, the common practice has been to subjectively weigh the need for detection and deterrence against burdensome requirements, cost-effectiveness considerations, and employee fairness issues. In establishing a 50 percent annual random testing rate

for operators of hazardous pipeline operations, for instance, DOT's Research and Special Programs Administration (RSPA) stated:

RSPA reiterates that a 50% testing rate is necessary to establish a valid confidence level as well as to provide an adequate deterrent to drug use by employees. During the comment period on the proposed rule, RSPA requested specific advice on what the random testing rate should be. Although many commenters suggested rates of 10-20%, none provided any data to support a particular level. RSPA, therefore, chose a random testing rate of 50% in part based on DOT's experience with its own internal drug testing program, as well as the rates used by the military services. Although the military had used higher rates to achieve the deterrent effect referred above, RSPA believed that the 50% rate offered a sufficient balance between a rate high enough to deter use while keeping costs reasonable (54 FR 51842, at 51846).

In some cases, agencies merely noted the assumed deterrent effect of random testing by following precedent and citing DOD's drug testing programs as an example (55 FR 3698, 1990).

Since 1993 several federal agencies have lowered their random testing rates, believing that deterrence can be achieved at lower testing rates. The principal reason for doing so is to reduce program costs in light of declining drug detection rates. Lowering testing rates will also result in administrative savings and reduce the burdens of testing on employers and employees. As noted above, the NRC lowered its annual random testing rate from 100 percent to 50 percent for similar reasons (59 FR 29908, 1994).

The DOT has taken the most significant departure from the traditional balancing approach used to set testing rates. In 1991 several railroads asked the DOT to compare a lower testing rate of 25 percent against its required test rate of 50 percent. The Federal Railroad Administration (FRA) agreed to undertake this study. In July, 1991 FRA began evaluating four experimental railroads using the 25 percent random test rate against a group of control railroads testing at the required 50 percent test rate.

In the first two years of the study, positive test rates did not vary appreciably. According to FRA, through June 30, 1993, the four test railroads conducted 19,958 tests,

## Deterrence

with a positive test rate of .90 percent. The four control railroads conducted 34,121 tests, with a positive rate of .85 percent. The difference in test results between these two groups is not statistically significant.<sup>5</sup>

Altering random test rates illustrates detection capability, not deterrent effect. For instance, Table 5.1 shows the expected detection rate at two different random testing rates for a theoretical testing population with a constant five percent detectable drug prevalence rate.<sup>6</sup> The test rate used is simple random testing, with varying probabilities of being selected for testing. In comparison, the likely detection rate using a modified random testing strategy is also shown. It is assumed that five percent of those tested would test positive for drug use. All things equal, a lower (not equal or higher) number of people should be detected as random test rates are lowered.

In a final rule published on December 2, 1994 (59 FR 62218), the DOT announced that it would set future minimum random drug testing rates according to the transportation industry's overall random drug positive rate, which is determined using annual agency drug program data. Under this system, the DOT lowered the industry random testing rate to 25 percent, as long as the industry-wide positive test rate is less than 1.0 percent for two calendar years while testing at 50 percent. The DOT will return the rate to 50 percent if the industry-wide positive test rate is 1.0 percent or higher in any subsequent calendar year. In 1995 the DOT set the 1996 minimum

random test rate at 25 percent, based on the program data (60 FR 231, 1995).

Table 5.2 shows the FRA industry-wide drug test results from 1990 through 1994. What these results say about deterrence is unclear. Inferring a deterrent effect from random testing rates and consequent detection levels is a tenuous exercise in the absence of controlled studies on people's risk perceptions and consequent drug use behaviors before and after the random test rates were changed. The FRA study assumed that people rationally calculate the probability of testing under different sampling regimes. According to FRA, however, employees were not surveyed regarding their understanding of the changes to the testing rate, nor questioned about whether this influenced their drug use behavior (Rockey, 1993).

The underlying assumption of the FRA study is that people can reliably estimate their risk of detection under different testing rates and adjust their drug use accordingly. This assumption is questionable. For instance, Table 5.3 presents a scenario that calculates the probability of selection according to different testing rates. The scenario assumes that selection is statistically random, and that it can occur at any point in time during each worker's shift. Further, this scenario assumes that each worker has ten paid holidays, ten paid vacation days, and is sick an average of one day a year. The probability of selection is 365 days minus 110 weekend days minus

**Table 5.1: Expected Detection Rates for Random Testing at 25% Test Rate and 50% Test Rate**

Simple Random Testing							
Testing Rate		Testing Population	Not Tested	Tested Once	Tested Twice	Tested Three or More Times	Detection at 5% Prevalence
			36%	37%	19%	8%	
Control Rate	50%	1,000	180	185	95	40	23
Experimental	25%	1,000	90	92.5	47.5	20	12
Modified Random Testing							
Testing Rate		Testing Population	Not Tested	Tested Once	Tested Twice	Tested Three or More Times	Detection at 5% Prevalence
			21%	61%	15%	3%	
Control Rate	50%	1,000	105	305	75	15	24
Experimental	25%	1,000	52.5	152.5	37.5	7.5	12

### Notes:

A 5 percent prevalence rate of drug use is assumed for purposes of illustration. For a statistical analysis of random testing detection capability under different scenarios of drug prevalence in a target population, see Durbin, et al. (1991, NUREG/CR-5784, Appendix C.)

**Table 5.2 FRA Industry-Wide Random Drug Testing Results**

	1990	1991	1992	1993	1994
Total Number of Random Tests	35,228	50,436	42,599	42,199	NA
Number of Positives	365	447	336	303	NA
Percent Positive	1.04%	0.89%	0.79%	0.72%	0.81%

Source: 59 FR 62218, 1994; 60 FR 231, 1995.

10 holidays minus 10 vacation days minus 1 sick day, meaning that each worker is eligible for testing 234 days annually.

Although the probability of testing varies according to the rate of testing, overall there is a very low actual probability of being selected for testing (less than 1 percent at any one time). Since there are no studies on perceived risks of detection according to various testing rates, it is not known what effect, for instance, a constant 0.1 percent probability of testing versus a constant 0.4 percent probability of testing would have on an occasional drug user's decision making process. Do most occasional users (a) conduct a rational assessment of probability of detection given frequency of drug use and likely exposure period, (b) decide that it is simply not worth the risk altogether, or (c) go to the other extreme and figure the odds for not being detected are good no matter what the random testing rate is? The FRA study does not answer these questions. In fact, these questions are impossible to determine by reviewing detection data alone. Too many critical variables are unknown such as:

- particular testing practices and their impact on worker risk perceptions,
- the prevalence of occasional drug users and when and under what circumstances they use drugs: without this information it is impossible to determine changes in behavior, let alone attribute the changes to random testing,

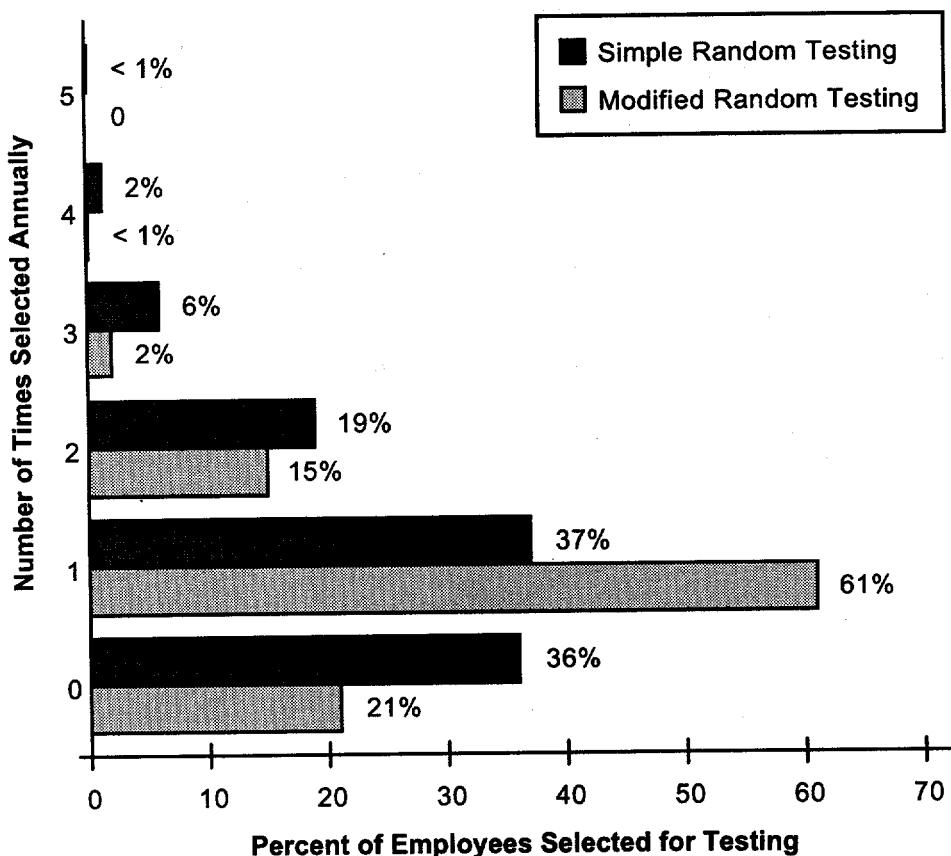
- the prevalence of chronic users in the population who cannot be deterred but should eventually be detected: since these persons cannot be deterred, eliminating these drug users from the workforce population is not a function of deterrence but detection and should be counted as such, and
- the deterrent effect of other FFD program elements: without knowing the contribution of other program elements to deterrence, it is impossible to isolate the singular effect of random testing.

### 5.3.1.2: Selection Methods

In addition to the testing rate, selection methods can also affect the probability, and hence the perceived likelihood, of detection. Under a simple random testing process at a 100 percent annual random testing rate, for instance, a worker may be tested once, twice, perhaps three or four times annually, or not at all, as shown in Figure 5.1. Approximately two-thirds of workers in the testing population will be tested one or more times; the remaining one-third will not be tested. Under a modified random testing strategy, people who are tested are placed into another population that has a smaller probability of being randomly selected than those who have never been tested. This increases the number of people selected for sampling once, while decreasing the probability of repeated selection but retaining some probability of being selected again, as shown in Figure 5.1. In particular, it

**Table 5.3 Probability of Being Selected According to Different Annual Random Testing Rates**

Annual Random Testing Rate	Probability of Selection on Any Given Shift
25%	1 in 936 days on shift or 0.1% probability
50%	1 in 468 days on shift or 0.2% probability
75%	1 in 312 days on shift or 0.3% probability
100%	1 in 234 days on shift or 0.4% probability



Source: Pohlman, 1993.

**Figure 5.1: Simple and Modified Random Testing at 100% Annual Testing Rate**

increases the probability that new hires, who often join the workforce with unknown drug use habits, will be tested within their first year of employment (Pohlman, 1993). It also decreases the adverse effects of repeated testing of people who are not drug users, which may have adverse effects on program acceptance.

It is unclear what modified random testing means for overall deterrence in a FFD program. In particular, it is not known whether a modified random sampling strategy may lead those who have been tested to conclude that the likelihood of repeated testing is sufficiently low to resume occasional drug use. Data from Southern California Edison's program, which adopted a modified random sampling strategy, showed increasingly lower detection rates (Osborn and Sokolov, 1990). Again, these are detection rate data, which do not provide direct evidence of deterrence effects. The data from Southern California Edison were confounded by changes to the sanctions (which became increasingly stringent over time), efforts to detect and deter subversion, and other program changes

over the same period, making it impossible to answer this question.

Statistical models have been used to support and question the deterrent effect of simple random testing. For instance, Burtis, Owens, and Lette (1987) argued that completely random testing better maintains deterrence than approaches that reduce the probability of selection after a first selection. Their argument is based on the view that deterrence is subjectively maximized by the constant uncertainty of being tested, and that only with completely random testing will each employee have a constant probability of selection on every testing occasion.

Burtis et al. also concede, however, that simple random testing is not without statistical artifacts. Indeed, as shown in Figure 5.1, the standard approach could result in about 36 percent of employees not being tested over the course of a year. Also, under a simple random approach, about two percent of workers will be tested four times

annually. It may be difficult to explain to these workers that they are not being harassed or being tested because of suspected drug use. The deterrent effects of under and oversampling are not known.

#### 5.3.1.3: Other Selection Practices

Other selection practices may also affect deterrence. In conducting its random testing program, for example, the military used to conduct "unit sweeps" so that an entire unit was tested at once rather than testing people randomly across units. Further, units were sometimes informed by their commander that testing was to be conducted prior to testing, obviating any supposed benefits from random, unannounced testing (Stoloff, 1985).

#### 5.3.1.4: Summary

Lacking empirical information on the relationship between individual risk perceptions, random testing practices, and behavioral outcomes, the most that can be said about declining drug detection rates is that they are consistent with the argument that drug testing programs in general have a deterrent effect. It is speculative to attribute declining drug detection trends primarily to random testing. Detection data cannot singularly be used to explain the effect of random testing on deterrence, nor to suggest what the most effective random sampling strategy should be.

Despite these limitations in what is known about random testing, proponents believe that random testing is most effective when it creates the perception that there is a strong probability of detection any time, and when the testing process is frequent and visible enough to ensure that the threat of testing is evident. These general assumptions are consistent with what we know about risk perception in theory, but do not shed useful information on optimal random testing rates and other testing practices. Decisions to set and adjust random testing rates and modify other testing practices have been made to date without empirical information on deterrence.

### 5.3.2: Pre-Employment Testing

Pre-employment testing has been one of the most effective FFD program elements in terms of detection (Durbin et al., 1991a). There is some empirical evidence on its deterrent effect as well. The military tested one set of applicants without informing them prior to the test (test results were not linked to applicants). The second sample of pre-employment applicants were informed that they

would be tested. Differences in test results between the two samples showed dramatically lower positive test results for marijuana and cocaine use among the informed applicants. This might have been due to users abstaining from marijuana and/or cocaine after being notified of a drug test, users withdrawing from the application after being notified of the testing requirement, and/or users refraining from applying altogether (MMWR, 1989).

This finding supports the initial deterrent value of pre-employment testing, but the deterrent effect of pre-employment testing on a worker over time is not known. Pre-employment testing is not likely to deter occasional users from applying for employment if they find it easy to abstain prior to the test. They may cease taking drugs until the test is over and until they receive a negative test result to ensure they are not detected. If continued testing is not imminent it is not likely that it will be sufficient to deter future use. Pre-employment testing is likely to detect chronic drug users, either through detection or self-removal from the application process. Pre-employment testing may consequently lead to relatively higher concentration of nonusers and occasional users in the workforce population, resulting in much lower detection rates, and an increased probability of a work climate intolerant of drug use.

### 5.3.3: For-Cause Testing

Another fitness-for-duty program element believed to have a strong detection and deterrent effect is for-cause testing, the referral of employees by supervisors for drug testing and/or employee assistance (Googins and Kurtz, 1980; Hoffman and Roman, 1984) and the identification of possible substance abuse of employees by coworkers. The review of the literature found no studies directly addressing the singular deterrent effects of for-cause testing in FFD programs. However, it is plausible to infer that for-cause testing may have a very large effect on deterrence in the workplace. An effective for-cause program increases the risk of detection and may identify drug users not detected through pre-employment or random testing.

The effectiveness of for-cause testing depends on the inclination of supervisors to refer an employee suspected of drug use (Googins and Kurtz, 1980; Hoffman and Roman, 1984). This program element requires a high level of management commitment to ensure that supervisors and employees do not tolerate drug use, and will not hesitate in acting upon that conviction.

### **5.3.4: Periodic Testing**

Periodic testing is expected to deter drug use among some users, but it may also foster increased drug use among other users. In studying periodic testing of athletes, Coombs and Ryan (1990) found that some athletes actually increased their drug use as a result of periodic drug testing. This was due to the athletes' perception that, once the test was over, it cleared the way for unbridled drug use until the next annual test. This suggests that periodic testing may be counterproductive when used as a stand-alone method to assess fitness for duty, and underscores the importance of unannounced, unpredictable testing to effectively deter drug use. If selection and notification practices are such that random testing occurs only in relatively predictable periods, then, like periodic testing, users who can control their drug use may schedule their use accordingly (Bray et al., 1992; Stoloff, 1985; MMWR, 1989).

### **5.3.5: Sanctions**

Studies of drunk driving and deterrence suggest that sanctions for drug use should be strong, consistent, and well understood in order to have maximum deterrent effect on risk perceptions (Beshai, 1984; Voas, 1986; Peck, Sadler, and Perrine, 1985; Hurst, 1985; Ross, 1992). Although the situational context of driving under the influence of alcohol (DUI) is very different from that of employees in safety-sensitive positions, similar themes are evident. When sanctions are lenient, inconsistent, or not well understood, deterrence is adversely affected (Voas, 1986). Driving under the influence involves a decision willingly to consume an impairing substance, it involves personal and public safety, and it involves the threat of formal sanctions against such behavior. Strong sanctions (license suspensions) appear to be more effective than any known form of alcohol education or rehabilitation, although some rehabilitation programs appear to be effective in reducing DUI recidivism (Peck, Sadler, and Perrine, 1985). Studies on the deterrent effect of random road blocks on drivers have shown that these programs substantially decrease drunken driving incidents (Dunbar, Penttila, and Pikkarainen, 1987). Also, multiple laws designed to increase the certainty of punishment such as sobriety checkpoints and preliminary breath tests are believed to have complementary deterrent effects (Evans, Neville, and Graham, 1991).

Most FFD programs meet the theoretical conditions for effective deterrence: sanctions for drug use are strong, consistent, and well-known industry wide (Durbin et al., 1991b; Bray et al., 1992). In the nuclear power industry,

for instance, the sanctions for first-time offenders require removal from duty for fourteen days, and follow-up assessment and rehabilitation. Repeat offenders are denied access to the protected areas of a nuclear power plant for a minimum of three years, and are usually terminated.

### **5.3.6: Integrity of the Testing Process and Program Administration**

Ensuring the effectiveness of FFD programs is considered important to deterrence: if the program fails to detect drug users effectively, individual perceptions of risk will be lowered and users will be more inclined to continue drug use. In particular, three adverse effects from a poorly run program can erode program credibility (Winkler, 1993):

- a drug user escapes detection and thereafter has a sense of invincibility,
- subversion, exacerbated by newly trained or unskilled collection personnel, may lead to a lack of confidence in the program, and
- simple errors may compromise the overall effectiveness of the program by impeding consistent detection of users, thereby lowering the risk of detection.

Periodic reviews of FFD program effectiveness are designed to ensure the overall quality of the program in a number of areas such as random testing administration; specimen collection and on-site testing procedures and facilities; training; notification, collection, quality assurance, and shipping procedures; program management and administration; worker welfare and rehabilitation; and security. It is plausible to assume that overall program effectiveness can have an effect on individual risk perceptions of the likelihood of testing and detection; poorly administered programs provide opportunities for subversion, whereas well-run programs indicate organizational commitment to a work setting intolerant of drug use.

### **5.3.7: Education, Training, and Awareness of Risks**

Employee perceptions of the health, safety, and employment risks of drug use may be influenced by FFD program training and education about the health and safety dangers of drug use, the effectiveness of various

program elements designed to detect those with known drug use problems, and the effectiveness (i.e., consistency and severity) of sanctions involved for drug use. If deterrence is based on knowledge that informs the individual's risk perception, then it follows that a visible, effective training program is a critical element in deterring illicit drug use. This is consistent with the literature discussed above on risk perception that examines how knowledge informs the decision maker's choices. For instance, if long-term negative effects associated with cocaine or marijuana use are understood, it may be easier to persuade an individual to forsake these drugs.

Although education and awareness is not generally an accepted stand-alone fitness-for-duty strategy, providing credible and useful information for making decisions regarding drug use in relation to personal and public health and safety has been shown to have a strong impact on personal perception of risk and a consequent strong deterrent effect among high school students. For this group, declining marijuana use was attributed to credible reports on the adverse effects of marijuana, and students' first-hand observation of peers who exhibited long-term negative consequences of marijuana use. Declining cocaine use was similarly attributed to some unique factors, namely, the cocaine-related deaths of some well-known athletes (U.S. Department of Education, 1990; Bachman, Johnston, and O'Malley, 1990). Effective education on the health and safety risks of drug use—and first-hand knowledge of the adverse effects of drug use in particular—may be expected to have a similar effect among safety-sensitive workers. As noted earlier, a management association survey found that companies that test and provide other drug prevention services such as education, training, awareness, and employee assistance, have consistently lower positive test rates than companies that only test (AMA, 1995). In the 1995 survey, firms with multiple drug prevention components, such as educational supervisory training, coupled with testing provisions, had a 1.9 percent positive rate, compared to 2.8 percent among companies with testing as the only program component (*Drugs in the Workplace*, 1995).

An additional important component of an effective education strategy is to ensure that the detection capabilities of the drug testing program are well understood for all drugs of use. Some drugs may be perceived as safer to use than other drugs in minimizing the risk of detection through testing or observation. For instance, if a worker understands that the risk exposure of marijuana is much greater than cocaine, he may switch to cocaine, which has a much lower period of detection, in

order to lower the risk of detection. Or, a person who used marijuana and alcohol may switch to alcohol alone since he knows that there are greater risks involved with marijuana. The education, training, and awareness strategy should convey the fact that drug use will be detected not only through random testing, but through other program elements such as for-cause testing, and supervisory observation.

### 5.3.8: Employee Assistance Programs

Employee assistance programs (EAPs) are designed to help individuals with work, family, and health-related problems which may lead to (or be caused by) drug use. Therefore, EAPs are a part of many FFD programs, and are required by the NRC's FFD regulation. EAPs are responsible for providing confidential assistance to employees who are referred or refer themselves. Frequently, the EAP provides the training and employee awareness of the health and safety hazards associated with substance abuse (including misuse and abuse of prescription and OTC drugs), the company's drug policy, program elements, and employee assistance services. The EAP may also be responsible for supervisory training on the effects of drug use, and in establishing a peer referral program (Barnes et al., 1988).

However, the use of EAPs and other rehabilitative or support services may be hindered by employees who believe that getting help is difficult, could adversely affect their career, or could result in disciplinary action (Bray et al., 1992). As discussed above, supervisors may be reluctant to refer employees with known drug use or alcohol-related problems for a number of reasons (Googins and Kurtz, 1980). These are major barriers to the use of EAP and other support services designed to assist drug users in ceasing their drug use altogether. Consequently, providing confidentiality to employees seeking assistance may be a necessary condition in some instances for a person to use EAP services in ceasing their drug use.

### 5.3.9: Organizational Influences on Deterrence

Deterrence can be maximized by assuring the effectiveness of detection, sanctions for drug use, and assistance to employees in ceasing drug use. It can also be supported by encouraging a workforce climate intolerant of drug use. As discussed above, peer pressure, drug culture knowledge, family support, organizational support, emotional fatigue, stress, and intrinsic

motivations for drug use play a role in drug use decisions (Barnes et al., 1988; Meier and Johnson, 1977; Williams, 1985; Curry, Wagner, and Grothaus, 1990; Kelly, Zyzanski, and Alemagno, 1991; Eggert and Herting, 1991; Raskin, Novacek, and Hogan, 1992). These larger factors are critical in promoting a workplace free of drugs. For instance, including the family in drug use treatment has been shown to contribute to effective treatment (Enders and Mercier, 1993). Social networks that promote drug abstention are also important (Eggert and Herting, 1991), although the effectiveness of such programs in the context of FFD programs is unknown.

Motivation is a singularly important predictor of change, including changing drug use habits (Kelly, Zylanski, and Alemagno, 1991; Curry, Wagner, and Grothaus, 1990). However, it is questionable what role organizations should play in motivating people to cease drug use. Intrinsic motivations are ones that are internal to the person such as the need for self-determination or self-competence, concern about negative health consequences, and personal values about safety. Extrinsic motivators are those to which a person responds to receive an extrinsic reward, avoid punishment, or avoid negative reinforcement (such as nagging by a spouse). In a study of smoking cessation, Curry, Wagner, and Grothaus (1990) found that smokers were more successful when they had higher intrinsic motivation to cease smoking relative to extrinsic motivation. This suggests that organizational strategies designed to promote drug prevention should emphasize personal health, safety, and competency over threats of punishment, repeated admonitions to quit, and so on.

If employees believe that drug use poses a legitimate workplace safety and security concern, then it is likely that drug use in any form will be discouraged. The workplace culture should assist in deterring drug use and behaviors that are legitimately perceived as posing a health and safety risk. Surveys of military personnel have shown that general awareness of FFD program policies has a deterrent effect in that the user is aware of FFD policies and modifies drug use accordingly. The military, however, is a unique workplace in that military personnel not only work on base, but also often live and raise families there and have stronger social ties with other personnel than would be expected in a civilian workforce. Drug use may be more strongly influenced by the values and norms shared in a military community setting than in civilian communities (Bray et al., 1992; Bray et al., 1990).

In *Operation Redblock*, peer pressure has been used in the railroad industry to dissuade coworkers from working under the influence of drugs. The program operates without supervisory or management oversight and was initiated by union staff. Employees are responsible for removing workers under the influence of drugs or alcohol. Offenses are initially handled by peers; repeat offenders are subjected to management discipline. The program has been noted for encouraging use of the railroad's EAP (Pack, 1993). However, there is little evidence about the efficacy of this and other constructive confrontation approaches used in FFD programs (Harris and Heft, 1992).

### 5.4: Conclusions

Several conclusions may be drawn from this review of the literature. Deterrence is affected by multiple factors: individual risk perceptions, FFD program elements such as sanctions, types of testing and employee assistance, and peer, family, and organizational attitudes toward drug use. This supports a multifaceted approach to maximize deterrence in FFD programs:

- effective detection capabilities coupled with strong, consistent and well-understood sanctions and efficient program administration,
- education or training on and awareness of the negative consequences of drug use, and the effectiveness of the program in detecting drug use, and
- an organizational culture that discourages drug use, supported by an EAP that is perceived as being effective, confidential, and non-punitive.

There are limitations to a deterrence strategy. Some occasional users will continue to use drugs regardless of risk. Chronic users cannot be deterred *per se*. Consequently, it is imperative to detect these employees through an effective testing program.

Deterrence is difficult to measure directly. In the absence of direct information on rates of drug use under different testing rates, indirect methods such as attitudinal studies and longitudinal studies of drug use trends may be used to infer the deterrent effect of FFD programs, i.e., by assessing the role of risk perceptions, peer, family, and organizational influences, and examining consequent behaviors and outcomes.

Several of the studies reviewed here have concluded that declining drug use detection rates indicate a deterrent effect from random testing, but these inferences are speculative. Detection rate data, often used to document a deterrent effect, provide little meaningful information on deterrence. In fact, there are no empirical data to justify the random testing rates currently used in FFD programs. The lack of empirical research on the deterrent effect of FFD programs in general and random testing in particular is unfortunate, given widespread claims on the importance of random testing's singular deterrent effect. Ongoing efforts to lower random testing rates are also being undertaken without any empirical basis. These efforts to lower random testing rates are not based on deterrence as much as a desire to reduce testing costs and administrative burdens. What effect these changes in

testing rates will have on deterrence in FFD programs is unknown.

Although there is some indication that other FFD program elements serve to deter drug use, little is known about their independent and synergistic effects in many workplace settings where these programs are carried out. It is clear from this review that additional information on attitudes toward drug testing programs in general and random testing in particular should be gathered and reviewed in order to ascertain the role of risk perception behaviors and other influences in the context of FFD programs. This would permit the design of programs that maximize deterrence of drug use. The deterrent value of FFD programs in general and random testing in particular will remain questionable until these issues are addressed.

## Endnotes

1. Joseph Hauth, M.P.A., Battelle Seattle Research Center
2. "Rational decisions" are usually defined as those that are consistent with a mathematical maximization of expected utility (von Neumann and Morgenstern, 1947).
3. Accurate estimates of the prevalence of drug use in testing populations are scarce. Data can be obtained from the National Institute on Drug Abuse Household Survey, which conducts a national survey of respondents to determine drug use prevalence, type of drug use, and the frequency of use. From this information, prevalence estimates for a workforce population can be developed. Nonusers comprise an estimated 85-90 percent of the workforce (for a further discussion of these estimates, see Chapter 3, Prevalence of Drug Use in the United States). Chronic and occasional users are believed to comprise the remaining 10-15 percent of the workforce population. Chronic users are persons who use drugs so frequently that their use may be continuously detectable. Occasional users are those persons who use drugs infrequently throughout the year and in such a manner that their drug use is not always detectable.
4. The effects of various random testing sampling strategies and overall detection rates are well-documented and are not repeated here. For a discussion of random testing rates and associated probabilities of detection for different types of drug abuse, see Durbin et al., 1991 (NUREG/CR-5784).
5. At a 95 percent confidence interval, the difference between the experimental test rate (.90 percent) and the control rate (.85 percent) is not significant (*t* value equals .60). That is, the difference between the experimental and control group positive test rates is not meaningful statistically.
6. Note: prevalence rates vary by drug type, dose, individual physiology, and other factors. For a review of expected drug prevalence rates and detection rates for occasional and chronic drug use, see Durbin et al., 1991 (NUREG/CR-5784).

## References

54 FR 24468. June 7, 1989. "Nuclear Regulatory Commission. Fitness for Duty Programs: Final Rule and Statement of Policy." *Federal Register*.

54 FR 51842. December 18, 1989. "Control of Drug Use in Natural Gas, Liquefied Natural Gas, and Hazardous Liquid Pipeline Operations." *Federal Register*.

55 FR 3698. February 2, 1990. "Anti-Drug Program for Personnel Engaged in Specific Aviation, Activities." *Federal Register*.

56 FR 22905. May 17, 1991. "Proposed Test Program to Evaluate Random Drug Testing Rate." *Federal Register*.

58 FR 15810. March 24, 1993. "Modifications to Fitness-for-Duty Program Requirements." *Federal Register*.

59 FR 502. (January 5, 1994). Nuclear Regulatory Commission. "Modifications to Fitness-for-Duty Program Requirements." *Federal Register*.

59 FR 29908. (June 9, 1994). Department of Health and Human Services. "Mandatory guidelines for federal workplace drug testing programs." *Federal Register*.

59 FR 62218. December 2, 1994. "Random Drug Testing Program; Final Rule." *Federal Register*.

60 FR 231. December 1, 1995. "Random Drug Testing: Announcement of 1996 Minimum Testing Rate; Designation of New Laboratory for Post-Accident Toxicology Testing." *Federal Register*.

American Management Association [AMA] (1991). "1991 AMA Survey: Workplace Drug Testing and Drug Abuse Policies," New York: Author.

American Management Association [AMA] (1992). "1992 AMA Survey: Workplace Drug Testing and Drug Abuse Policies," New York: Author.

American Management Association [AMA] (1993). "1993 AMA Survey: Workplace Drug Testing and Drug Abuse Policies," New York: Author.

American Management Association [AMA] (1995). "1995 AMA Survey: Workplace Drug Testing and Drug Abuse Policies," New York: Author.

Bachman, J.G., Johnston, L.D., and O'Malley, P.M. (1990). "Explaining the Recent Decline in Cocaine Use Among Young Adults: Further Evidence That Perceived Risks and Disapproval Lead to Reduced Drug Use," *Journal of Health and Social Behavior*, 31(June): 173-184.

Barnes, V., Fleming, I., Grant, T., Hauth, J., Hendrickson, J., Kono, B., Moore, C., Olson, J., Saari, L., Toquam, J., Wieringa, D., Yost, P., Hendrickson, P., Moon, D., and Scott, W. (1988). "Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues (NUREG/CR-5227)." Washington, DC: Nuclear Regulatory Commission.

Beshai, N. (1984). "California DUI Law: One Year Implementation," *Abstracts and Reviews in Alcohol and Driving*, 5:3, 11-20.

Bray, R.M., Kroutil, L.A., Luckey, J.W., Whelless, S.C., Iannacchione, V.G., Anderson, D.W., Marsden, M.E., and Dunteman, G.H. (1992). *1992 Worldwide Survey of Substance Abuse and Health Behaviors Among Military Personnel*. Research Triangle Park, NC: Research Triangle Institute.

Bray, R.M., Marsden, M.E., and Peterson, M.R. (1991). "Standardized Comparisons of Use of Alcohol, Drugs, and Cigarettes Among Military Personnel and Civilians," *American Journal of Public Health*, 81:7, 865-869.

Bray, R.M., Marsden, M.E., Rachal, J.V., and Peterson, M.R. (1990). "Drug and Alcohol Use in the Military Workplace: Findings from the 1988 Worldwide Survey." In: S.W. Gust, J.M. Walsh, L.B. Thomas, and D.J. Crouch, (Eds.), *Drugs in the Workplace: Research and Evaluation Data, Volume II*. NIDA Research Monograph No. 100. Rockville, MD: National Institute on Drug Abuse, 25-43.

Bray, R.M., Marsden, M.E., Guess, L.L., Wheless, S.C., Iannacchione, V.G., and Keesling, S.R. (December 1988). *1988 Worldwide Survey of Substance Abuse and Health Behaviors Among Military Personnel*

## Deterrence

(RTI/4000/06-02FR), Research Triangle Institute: Research Triangle, NC.

Buikhuisen, W. (1974). "General Deterrence: Research and Theory," *Abstracts on Criminology and Penology*, 14 (3), 285-298.

Burkett, S.R., and Ward, D.A. (1993). "A Note on Perceptual Deterrence, Religiously Based Moral Condemnation, and Social Control," *Criminology*, 31 (1), 119-133.

Burtis, C.A., Owings, J.H., and Lette, R.S. 1987. Statistical considerations of the random selection process in a drug-testing program. *Clinical Chemistry*, 33 (11), 46B-48B.

Cohen, S. (November 1986). "The Military Worldwide Surveys: Deterrent Effects of Urine Testing on Drug Use," *Drug Abuse and Alcoholism Newsletter*, Vista Hill Foundation, 15:9.

Coombs, R.H., and Ryan, F.J. (1990). "Drug Testing Effectiveness in Identifying and Preventing Drug Use," *American Journal of Drug and Alcohol Abuse*, 16:3 and 4, 173-184.

Crant, J.M., and Bateman, T.S. (1990). "An Experimental Test of the Impact of Drug-Testing Programs on Potential Job Applicants' Attitudes and Intentions," *Journal of Applied Psychology*, 75:2, 1-5.

Cromwell, P.F., Olson, J.N., Avary, D.W., and Marks, A. (1991). "How Drugs Affect Decisions by Burglars," *International Journal of Offender Therapy and Comparative Criminology*, 35 (4), 310-320.

Connors, M.M. (1992) "Risk Perception, Risk Taking and Risk Management Among Intravenous Drug Users: Implications for AIDS Prevention," *Society, Science and Medicine*, 34 (6), 591-601.

Curry, S., Wagner, E.H., and Grothaus, L.C. (1990). "Intrinsic and Extrinsic Motivation for Smoking Cessation," *Journal of Consulting and Clinical Psychology*, 58 (3), 310-316.

DOT. (1993). *Meeting on Random Drug Testing*, (proceedings, Holiday Inn Capital Columbia Room, 550 C Street, Southwest, Washington, D.C., Monday, February 1, 1993). Washington, D.C.: U.S. Department of Transportation.

Drug Detection Report "DOT Considers Testing Rates Set by Company or by Laboratory." February 1993.

Drug Detection Report *Drugs in the Workplace*, "AMA Survey: Positive Drug Tests Down, MROs and Random Testing Up." August 1995, 6.

Dunbar, J.A., Penttila, A., and Pikkarainen, J. (1987). "Random Breath Tests in Finland Have Halved Drinking and Driving Rate," *British Medical Journal*, 295:101-103.

Durbin, N., Fleming, T., Murphy, S., Macaulay, J., Westra, C., Olson, J., and Christensen, J. (1991a). "Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports CY1990 (NUREG/CR-5758)." Washington, DC: Nuclear Regulatory Commission.

Durbin, N., Moore, C., Grant, T., Fleming, T., Hunt, P., Martin, R., Murphy, S., Hauth, J., Wilson, R., Bittner, A., Bramwell, A., Macaulay, J., Olson, J., Terrill, E., and Toquam, J. (1991b). "Fitness for Duty in the Nuclear Power Industry: A Review of the First Year of Program Performance and an Update of the Technical Issues (NUREG/CR-5784)." Washington, DC: Nuclear Regulatory Commission.

Eggert, L.L., and Herting, J.R. (1991). "Preventing Teenage Drug Abuse: Exploratory Effects of Network Social Support," *Youth and Society*, 22 (4), 482-524.

Enders, L.E., and Mercier, J.M. (1993) "Treating Chemical Dependency: The Need for Including the Family," *The International Journal of the Addictions*, 28 (6), 507-519.

Evans, W.N., Neville, D., and Graham, J.D. (1991). "General Deterrence of Drunk Driving: Evaluation of Recent American Policies," *Risk Analysis*, 11 (2), 279-289.

Fishburn, P.C. (1988) Utility theory (pp. 445-452). In: S. Kotz et al. (Eds.) *Encyclopedia of Statistical Sciences*. Vol. 9. New York, NY: John Wiley and Sons.

GAO (1991). *Employee Drug Testing: Status of Federal Agencies' Programs: Report to the Chairman, Committee on Governmental Affairs, U.S. Senate*, (May 1991). Washington, DC: U.S. General Accounting Office, GAO/GGD-91-70, 14-19.

Googins, B., and Kurtz, N.R. (1980). "Factors Inhibiting Supervisory Referrals to Occupational Alcoholism Intervention Programs," *Journal of Studies on Alcohol*, 41:11, 1196-1208.

Gonzalez, G.M., and Haney, M.L. (1990). "Perceptions of Risk as Predictors of Alcohol, Marijuana, and Cocaine Use Among College Students," *Journal of College Student Development*, 31, 313-318.

Hall, W., Flaherty, B., and Homel, P. (1992). "The Public Perception of the Risks and Benefits of Alcohol Consumption," *Australian Journal of Public Health*, 16 (1), 38-42.

Harris, M.M. (1993). "Drugs in the Workplace: Setting the Record Straight," *The Journal of Drug Issues*, 23 (4), 727-732.

Harris, M.M., and Heft, L.L. (1992). "Alcohol and Drug Use in the Workplace: Issues, Controversies, and Directions for Future Research," *Journal of Management*, 18 (2), 239-266.

Hartman, S.J., and Crow, S.M. (1993). "Drugs in the Workplace: Setting Harris Straight," *The Journal of Drug Issues*, 23 (4), 733-738.

Hoffman, E., and Roman, P.M. (1984). "Effects of Supervisory Style and Experientially Based Frames of Reference on Organizational Alcoholism Programs," *Journal of Studies on Alcohol*, 45:3, 260-267.

Howe, E.S. and Brandau, C.J. 1988. "Additive Effects of Certainty, Severity, and Celerity of Punishment on Judgments of Crime Deterrence Scale Value." *Journal of Applied Social Psychology*. 18(9), 796-812.

Hurst, P.M. (1985). "Blood Alcohol Limits and Deterrence: Is There a Rational Basis for Choice?," *Alcohol, Drugs, and Driving*, 1:1 and 2, 122-133.

Institute for a Drug-Free Workplace (1991). "What Employees Think About Drug Abuse,"

Kelly, R.B., Zyzanski, S.J., and Alemagno, S.A. (1991). "Prediction of Motivation and Behavior Change Following Health Promotion: Role of Health Beliefs, Social Support, and Self-Efficacy," *Society, Science and Medicine*, 32 (3), 311-320.

Kaplan, H.B., and Fukurai, H. (1992). "Negative Social Sanctions: Self-Rejection, and Drug Use," *Youth and Society*, 23 (3), 275-298.

Kingsnorth, R.F. (1991). "The Gunther Special: Deterrence and the DUI Offender," *Criminal Justice and Behavior*, 18 (3), 251-266.

Lauderback D., and Waldorf, D. (1993). "Whatever Happened to Ice?: The Latest Drug Scare," *The Journal of Drug Issues*, 23 (4), 597-613.

Leatherwood, M.L., and Spector, L.C. (1991). "Enforcements, Inducements, Expected Utility and Employee Misconduct," *Journal of Management*, 17 (3), 553-569.

LeRoy, M.H. (1991). "Discriminating Characteristics of Union Members' Attitudes Toward Drug Testing in the Workplace," *Journal of Labor Research*, XII:4, 453-466.

LeRoy, M.H. (1990). "Drug Testing in the Public Sector: Union Member Attitudes," *Journal of Collective Negotiations*. 19:3, 165-173.

Little, G.L., and Robinson, K.D. (1990). "Treating Drunk Drivers with Moral Reconation Therapy: A Two-Year Recidivism Study," *Psychological Reports*, 66, 1379-1387.

"Lower Random Testing Rate Urged by Commission," *Drug Detection Report*, July 20, 1993.

*Morbidity and Mortality Weekly Report [MMWR]* (1989). "Prevalence of Drug Use among Applicants for Military Service—United States, June-December 1988," 38:3, 580-583.

Moore, C., Barnes, V., Hauth, J., Wilson, R., Fawcett-Long, J., Toquam, J., Baker, K., Wieringa, D., Olson, J., and Christensen, J. (1989). "Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues (NUREG/CR-5227; PNL-6652)." Washington, DC: Nuclear Regulatory Commission.

Murphy, K.R., Thornton III, G.C., and Reynolds, D.H. (1990). "College Students' Attitudes Toward Employee Drug Testing Programs," *Personnel Psychology*, 43:615-631.

Normand, J., Lempert, R.O. and O'Brien, C.P., Eds. (1994). *Under the Influence? Drugs and the American*

## Deterrence

*Work Force.* Washington, D.C.: National Academy Press.

Newcomb, M.D. (1992). "Substance Abuse and Control in the United States: Ethical and Legal Issues," *Society, Science and Medicine*, 35 (4), 471-479.

NRC. (1992). *Modification to the Random Testing Rate*. SECY-92-271. Washington, D.C.: U.S. Nuclear Regulatory Commission.

Osborn, C.E., and Sokolov, J.J. (1990). "Drug Use Trends in a Nuclear Power Facility: Data From a Random Screening Program." In: S. W. Gust, J. M. Walsh, L. B. Thomas, and D. J. Crouch, (Eds.), *Drugs in the Workplace: Research and Evaluation Data, Volume II*. NIDA Research Monograph No. 100. Rockville, MD: National Institute on Drug Abuse, 25-43.

Osborn, C.E., and Sokolov, J.J. (1989). "Drug Use Trends in a Nuclear Power Company: Cumulative Data From an Ongoing Testing Program." In: S.W. Gust, and J.M. Walsh (Eds.), *Drugs in the Workplace: Research and Evaluation Data*. NIDA Research Monograph No. 91. Rockville, MD: National Institute on Drug Abuse, 69-80.

Pack, E. In: DOT. (1993). *Meeting on Random Drug Testing* (proceedings, Holiday Inn Capital Columbia Room, 550 C Street, Southwest, Washington, D.C., Monday, February 1, 1993). Washington, D.C.: U.S. Department of Transportation, 283-288.

Peck, R.C., Sadler D.D., and Perrine, M.W. (1985). "The Comparative Effectiveness of Alcohol Rehabilitation and Licensing Control Actions for Drunk Driving Offenders: A Review of Literature," *Alcohol, Drugs, and Driving*, 1:4, 15-39.

Pohlman, B. In: DOT. (1993). *Meeting on Random Drug Testing*, (proceedings, Holiday Inn Capital Columbia Room, 550 C Street, Southwest, Washington, D.C., Monday, February 1, 1993). Washington, D.C.: U.S. Department of Transportation, 56-68.

Piliavin, I., Gartner, R., Thornton, C., and Matsueda, R. L. (1986). "Crime, Deterrence, and Rational Choice," *American Sociological Review*, 51 (1), 101-119.

"Railroad Random Testing Study To Continue Into Third Year," *Drug Detection Report*, June 21, 1993, 5.

Raskin, R., Novacek, J., and Hogan, R. (1992). "Drug Culture Expertise and Substance Use," *Journal of Youth and Adolescence*, 21 (5), 625-637.

Rockey, W. "FRA Experimental Random Test Rate Program," in: DOT. (1993). *Meeting on Random Drug Testing*, (proceedings, Holiday Inn Capital Columbia Room, 550 C Street, Southwest, Washington, D.C., Monday, February 1, 1993). Washington, D.C.: U.S. Department of Transportation, 119-130.

Ross, H. L. (1992). *The Deterrent Capability of Sobriety Checkpoints: Summary of the American Literature*. DOT HS 807 862. Washington, D.C.: National Highway Traffic Safety Administration.

Schwartz, S. (1991). "Decision Factors and Program Preferences of Drug-Using and Non-Using Students," *The Journal of Drug Issues*, 21 (3), 527-541.

Simon, H.A. (1956). "Rational Choice and the Structure of the Environment," *Psychological Review*, 63, 129-138.

Slovic, P., Lichenstein, S., and Fishhoff, B. (1988). "Decision Making." In: R.C. Atkinson et al. (Eds.), *Stevens' Handbook of Experimental Psychology* (2nd Ed.), Vol. 2. New York, NY: John Wiley and Sons.

Stoloff, P.H. (1985). *The Effectiveness of Urinalysis as a Deterrent to Drug Use*, p. 11, Washington, DC: Department of the Navy.

Taggart, R.W. (1989). "Results of the Drug Testing Program at Southern Pacific Railroad." In: S. W. Gust, and J.M. Walsh (Eds.), *Drugs in the Workplace: Research and Evaluation Data*. NIDA Research Monograph No. 91. Rockville, MD: National Institute on Drug Abuse, 97-108.

U.S. Department of Education. (1990). "Survey Notes Continuing Decline In Drug Use By High School Seniors," *Schools Without Drugs: The Challenge*, 4:3.

Voas, R.B. (1986). "Evaluation of Jail as a Penalty for Drunk Driving," *Alcohol, Drugs, and Driving*, 1:4, 15-39.

von Neumann, J., and Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Waldo, G.P., and Chiricos, T.G. (1972). "Perceived Penal Sanction and Self-Reported Criminality: A

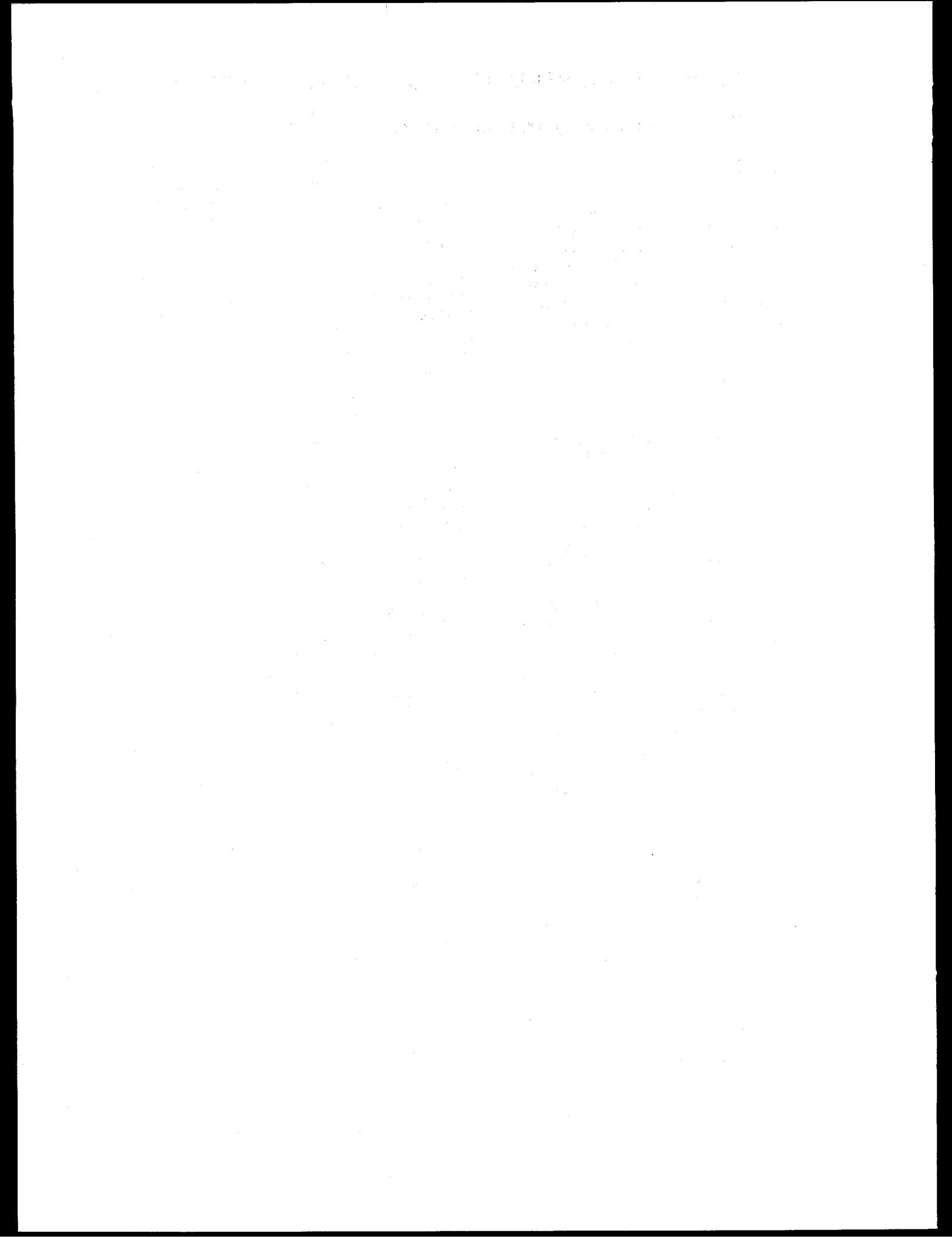
Neglected Approach to Deterrence Research," *Social Problems*, 19:4,522-540.

Walsh, D.C., Elinson, L., and Gost, L. (1992). "Worksite Drug Testing," *Annual Review of Public Health*, 13, 197-221.

West, L.J., and Ackerman, D.L. (1993). "The Drug Testing Controversy," *The Journal of Drug Issues*, 23 (4), 579-595.

Williams, F.P., III. (1985) "Deterrence and Social Control: Rethinking the Relationship," *Journal of Criminal Justice*, 13, 141-151.

Winkler, H. "Georgia Power Company Presentation." In: DOT. (1993). *Meeting on Random Drug Testing*, (proceedings, Holiday Inn Capital Columbia Room, 550 C Street, Southwest, Washington, D.C., Monday, February 1, 1993). Washington, D.C.: U.S. Department of Transportation.



## CHAPTER 6: SUBVERSION AND COUNTERSUBVERSION

Chris Moore, Nancy Durbin, and Jennifer Macaulay<sup>1</sup>

### 6.1: Introduction

Subversion, which refers here to the deliberate attempt to evade the detection of substance abuse, can substantially decrease the effectiveness of drug testing programs. Program effectiveness hinges on valid and accurate results and fairness. If some drug users are permitted to evade detection because of vulnerabilities in the program, then the integrity of the FFD program is compromised and the program will neither detect nor deter substance abuse.

If individuals in the testing program believe that subversion is not being prevented, it not only reduces the deterrent effect of the testing program, it can also erode support for the testing program among nondrug-abusing employees. Asking employees to participate in a testing program is justified by the need for a working environment free from substance abuse and its effects. If the program is seen as ineffective at preventing subversion, then the justification for the program is suspect—that is, if individuals continue to abuse drugs and are not identified by the program, what good is it?

In addition to damaging the effectiveness of and support for a particular program, subversion of the testing process can also compromise the evaluation of program performance industry-wide. The NRC's FFD program performance data and data from other drug testing programs such as the military (Hildebrand, 1993) and transportation industry (Edgell, 1993) have shown continuing declines in drug detection rates for several years. (In 1992 the declining drug detection trends leveled off, see Chapter 3 of this report for a discussion.) Proponents of drug testing suggest that declining drug detection rates over the last several years indicate that program policies and procedures are effective. However, declining drug detection rates may be due in part to drug users successfully eluding detection. Unfortunately, it is not possible to know how many drug users in these testing populations are escaping detection by subverting the testing process; what is known is that there have been attempts to subvert the testing process that have failed. Because successful subversion attempts are defined as those that are not identified, the nature and extent of subversion can only be estimated by examining unsuccessful subversion attempts. While this does not provide hard evidence about the total extent of the problem, it does provide some basis for understanding the scope of the problem.

Further evidence for the extent of the subversion problem is the proliferation of products and sources of advice on

how to subvert the drug testing process. Products and advice can be found on the Internet at web sites provided by the National Organization for the Reform of Marijuana Laws (NORML) ([www.norml.org](http://www.norml.org)) and High Times Magazine ([www.hightimes.com](http://www.hightimes.com)). High Times Magazine and others advertise for many products (books, home drug tests, adulteration and flushing related products), as well as advertising a 900 phone line with recorded information about how to pass a drug test. NORML provides a similar service via a 900 number with live operators to provide advice. The impact of widely available subversion advice and products can be nicely summed up by a statement made on the High Times Hotline recorded advice, "Most informed users can pass a drug test."

This chapter provides an overview of subversion techniques (6.2); reviews evidence of and experience with subversion, focusing on evidence of subversion in the nuclear power industry (6.3); discusses currently available countersubversion strategies (6.4); reviews two important issues regarding subversion, the identification of subversion in drug testing programs and the protection of employee rights (6.5); and provides a summary and conclusions (6.6). The focus of this chapter is on subversion that occurs prior to and during the process of collecting specimens for urinalysis. Although subversion can also occur at other points in the process—such as during shipment and at the laboratory—notification for testing and specimen collection are the focal points in this chapter because available evidence suggests that these are the most common targets for subversion attempts and because they represent the segments of the testing process over which FFD program managers have the greatest control.

### 6.2: Subversion Techniques

This section reviews the various ways in which the drug testing process can be subverted. Subversion can occur at any stage in the drug testing process—from "hacking" the computerized programs that select people for random testing (to ensure that some names never come up) to altering Department of Health and Human Services (DHHS) laboratory analysis reports. While it is important to review these aspects of the program with regard to vulnerability to subversion, this section focuses on techniques that affect collection of the urine specimen for testing.

The collection process includes notifying the individual of the test, obtaining a urine specimen from the individual,

## Subversion and Countersubversion

and assuring that chain of custody is maintained. For an individual who is abusing drugs, one effective means of evading detection of substance abuse is to avoid submitting any specimen. This can be accomplished, for example, by never being available to take a test. Another strategy is to delay a test until drug metabolites are no longer detectable in urine. Other effective means include using techniques that will cause a urine specimen to test negative even though it does, in fact, contain drug metabolites. This type of subversion is accomplished by three principal techniques: substitution, dilution, and adulteration.

- Substitution: substituting drug-free urine or another substance for the specimen provider's own urine.
- Dilution: diluting the urine so that the quantity of drugs or metabolites is below the cutoff<sup>2</sup> level used in testing.
- Adulteration: adding some substance to the specimen that will interfere with the tests and cause false-negative results.

Each of these techniques can be accomplished either *in vivo* or *in vitro* (Cody, 1990). *In vivo* subversion is any specimen tampering that occurs within the donor's body (endogenous substitution, dilution, or adulteration); *in vitro* subversion refers to tampering that occurs outside the body and after the specimen enters the collection

container (exogenous substitution, dilution, or adulteration). Examples of each of these types of subversion are provided in Table 6.1. The primary objective of each of these techniques is to cause a false-negative test result.

### 6.2.1: Specimen Substitution

Substitution is the submission of a surrogate specimen in lieu of the donor's specimen, usually during specimen collection procedures. Using *in vivo* substitution, a donor attempts to provide a clean urine specimen with substitute urine that is concealed within his or her own body. Examples of this technique include using a catheter to fill the bladder with "clean" urine or a woman placing a container of urine (such as a urine-filled condom) within her vagina. Both of these techniques may allow the donor to provide a substitute specimen in a manner that renders observation or detection of the substitution difficult, though not impossible. Both methods permit providing a specimen with a credible temperature if the substitute urine has been in the body for a period of time. Both methods require preparation prior to the test as donors must engage in the difficult and possibly painful process of placing drug-free urine inside their bodies.

*In vitro* substitution also requires preparation prior to the test. Although it doesn't require placing drugs in a body, it may be more difficult than *in vivo* substitution because 1) the substitute urine is not concealed in the body and

Table 6.1: Examples of the Six Types of Subversion

	<i>In Vivo</i>	<i>In Vitro</i>
Substitution	Filling one's bladder, via a catheter, with substitute clean urine.	Filling the specimen container with substitute clean urine or another substance that is smuggled into the collection site.
Dilution	Drinking large quantities of fluids and/or consuming diuretics prior to donating a specimen to dilute the specimen or flush drugs or metabolites from one's body (hydration).	Adding water or other liquid to the specimen to dilute the specimen.
Adulteration	A variety of substances may be ingested. It has been shown that aspirin may result in lower rates of change in absorbency than the rate for a drug-free EMIT d.a.u. calibrator, possibly resulting in false-negative results (Wagener, Linder, and Valdes, 1994).	Adding to the specimen any of various substances, ranging from salt to products specifically designed for subverting urine drug tests.

must be hidden in some other way, and 2) if urine is tested for temperature, the substitute urine must be within the appropriate temperature range when delivered to collection personnel (this usually requires warming the specimen and using a thermometer to check the temperature of the specimen prior to delivering it to the collection personnel).

Substitution requires that the individual have immediate access to a surrogate specimen or have the time and opportunity to obtain such a specimen prior to reporting to the collection site for testing.

*In vitro* substitution can be quite effective, however. One example of highly effective *in vitro* substitution was cited by an operator on the NORML Hotline (personal communication, March 2, 1996). In this instance the donor had submitted several suspicious samples in the past, and was notified he would have to provide an observed specimen. This individual bought a realistic looking plastic penis, hollowed it out such that when he squeezed it fluid would run out the tip, and then warmed a substitute sample just before testing. He was able to give an "observed" sample. Interestingly, although the sample was found to be slightly too cold, the observer swore it was acceptable.

### 6.2.2: Specimen Dilution

Dilution refers to the practice of thinning the urine specimen with water or other fluid to lower the concentration of drugs or drug metabolites below test cutoff levels. The *in vivo* practices are referred to as hydration, while thinning the urine *in vitro* is referred to as dilution.

Hydration has more potential than *in vitro* dilution as a subversion technique. It is achieved by consuming large quantities of fluids to increase the volume of the urine, thereby lowering its concentrations of drugs or their metabolites. The donor may also use diuretics which increase the amount of fluid excreted and thus reduce the percentage of solids or metabolites in the urine. To counter visual checks of the specimen for dilution, consumption of B-complex vitamins are used to make the diluted urine more yellow and normal in appearance.

Hydration is often recommended by subversion "experts" who provide information on beating drug tests. NORML, for instance, maintains a drug testing information telephone service. This service advises people that the best ways to avoid having marijuana use detected through urine screening are: 1) abstinence for 40 to 50 days prior

to the test, and 2) flushing the system prior to testing by drinking large quantities of fluids (representative of NORML, personal phone communication, March 2, 1996). "Pope Weasel," who provides tape recorded information on beating drug tests through High Times Magazine's drug urinalysis "hotline," also recommends hydration over other methods (Pope Weasel, High Times Magazine, toll-call recording, March 2, 1996). Weasel offers the following recommendations:

- Drug users should not rely on adulterants, as many do not work, are detectable, or the donor might be observed attempting to add them to the urine.
- Drug users should do everything they can to reduce the concentration of drugs in their systems prior to testing. They should never give up the first urine of the day, as this specimen will be especially concentrated.
- Donors should drink as much fluid as they can, staying up all night prior to the test to consume liquids if necessary. Weasel also states that the effectiveness of this purging can be enhanced by taking diuretics (specific products are recommended) and by taking saunas and recommends the use of vitamin B to make dilute urine appear a more normal yellow color.

While these recommendations may not be practical for the user subject to random testing with a short time span between notification and testing, they are indicative of the popularity of hydration among those who would subvert drug tests.

A booklet produced by Byrd Laboratories also recommends hydration while disparaging a number of potentially effective adulterants (Nightbyrd, 1991).

Dilution of urine *in vitro* has less potential as a subversion technique than hydration. Minimal counter-subversion practices usually include either having no water sources at the collection site or coloring water sources at the collection site with blue dye. If these practices are in place, the donor would need to smuggle water or other appropriate fluid into the collection site to use for the dilution. Hence, this option poses many of the same difficulties as smuggling a drug-free urine specimen (substitution) or an adulterant (adulteration) into the collection site. The liquid must be obtained before entering the collection site; it must be smuggled into the collection facility and added to the specimen without alerting collection site personnel; and, if temperature is

checked at the collection site, it must be warmed (either before or after it is added to the specimen) so that the temperature of the specimen will be within an acceptable range.

### 6.2.3: Specimen Adulteration

Adulteration involves adding some substance to the urine that will interfere with tests and cause false negative results. Although products are advertised as *in vivo* adulterants, it is generally believed that *in vivo* adulteration is ineffective since there is little evidence that any substance can be consumed to alter the results of a urine test. An exception is aspirin in therapeutic doses, which may cause false-negative results for drug screens using EMIT® d.a.u. assays. Aspirin (salicylates) were found to result in lower rates of change in absorbency than the rate for a drug-free calibrator, potentially leading to false-negative results (Wagener, Linder, and Valdes, 1994). It is not known whether aspirin has been used to mask drug use. Further empirical research on this issue is needed. Aspirin use is pervasive, and the issue does not seem to be warranted based on drug testing experience to date (Arthur Zebelman, Laboratory of Pathology of Seattle, Inc., personal communication, June 12, 1994). Furthermore, the drug of interest in the experiment by Wagnener, et al. (1994) was cocaine metabolites; it is not known whether aspirin ingestion affects test results for other drugs of abuse. Such results should therefore be considered preliminary.

There has been some debate in the scientific community on the use of golden seal root, used *in vivo* as an herbal tea, to mask drug use. Controlled studies should be performed before concerns about the adulterating effects of golden seal tea are dismissed (Cody, 1990).

Finally, although vitamin B may be taken to provide a yellower specimen, it has no effect on the laboratory analysis.

In contrast to the questionable effectiveness of *in vivo* adulteration, there are a number of known substances that can be added to urine *in vitro* to interfere with screening tests. There is no single "best" *in vitro* adulterant, because the adulterants vary with the drug consumed and the testing technology. The most common *in vitro* adulterants are salt, oxidizing agents such as bleach, some soaps, some liquid drain cleaners, sodium bicarbonate, vinegar, Visine® eye drops, golden seal tea, and products manufactured specifically for the purpose of adulterating urine tests, such as Byrd Laboratories' UrinAid®, Mary Jane's Super Clean 13, and Wizzies. (Many of these

adulterants are discussed in more detail in NUREG/CR-5227 [Barnes et al., 1988] and NUREG/CR-5784 [Durbin et al., 1991].)

Continued study has provided substantial evidence that a small quantity of Visine® eye drops can cause false-negative test results for marijuana and benzodiazepines on some screening assays (Mikkelsen and Ash, 1988; Cody and Schwarzhoff, 1989; Pearson, Ash, and Urry, 1989). Another product, Mary Jane's Super Clean 13, is marketed as a cleanser for fine instruments, jewelry, and "excretions such as urine," with emphasis on preventing "false positives for any reason." (It is advertised as "the ultimate purifier designed for emergency situations when instruments or liquid need to be perfectly clean to pass the closest scrutiny," and, "the ideal product for those worried about dirty testing instruments and false positives.")

Byrd Laboratories, which had previously marketed "UrinAid®" as an adulterant causing false-negative marijuana test results on EMIT® and GC/MS assays, issued a "Consumer Notice" warning of a new method to detect UrinAid®. (This method is discussed in Section 6.4.3.3.) However, its primary ingredient, glutaraldehyde, a substance that does not normally occur in urine, is still obtainable through noncommercial vendors (*MRO Alert*, 1993). Although Byrd advertised this product specifically for "avoiding false-positive tests" for marijuana, it can produce false-negative test results on virtually all immunoassay screens for all of the major drug and drug classes included in workplace drug testing programs (*MRO Alert*, 1993).

### 6.3: Evidence of and Experience with Subversion

The review of the limited information on evidence of and experience with subversion indicates that the subversion techniques reviewed in Section 6.2 are the most significant concerns in urine test subversion. Some new evidence casts light on the relative popularity of subversion strategies, and some refinements in methods have been identified. The most significant of these new findings are:

- Refined techniques for *in vitro* substitution (i.e., substituting surrogate samples) have been detected at several nuclear facilities.
- Subversion by hydration is a growing concern in drug testing programs.

- Increasing sophistication practiced by workers attempting to subvert the drug testing process is suggested by the pattern of results seen in various drug testing programs.

The current available information on the evidence of and experience with subversion is presented below. Examples of subversion in the nuclear power industry were obtained from licensee event reports to the NRC<sup>3</sup> and were provided in summary form to the authors by Mr. Loren Bush of the NRC (see Appendix F).

### 6.3.1: *In Vitro* Specimen Substitution

This type of subversion has been the most often identified by nuclear power plant licensees. This may reflect the relative ease of identifying this type of subversion, and not be evidence that it is the most common type of subversion that is attempted. In a number of cases, surrogate samples were identified due to the need to produce a specimen at the appropriate temperature. At several nuclear facilities workers were found using heat sources (microwave ovens, commercial hand warmers<sup>4</sup> and warm water) to warm surrogate urine samples, intending to substitute these specimens at body temperature for their own urine at the collection site (*Drug Detection Report*, 1992, NUREG/CR 5758, Vol. 3, [Fleming et al., 1993]). In at least one of these cases, the individual admitted that he had previously submitted surrogate samples. In another case reported to the NRC, a supervisor at a nuclear power plant, who was arrested for growing marijuana at a nuclear power plant site, admitted during the subsequent investigation that he had been submitting surrogate samples obtained from his son. At another nuclear power plant a specimen provided for pre-access testing was identified as having a low specific gravity. Specific gravity and creatinine tests, discussed in Section 6.4.3.3, can detect dilute samples or samples adulterated or substituted with specific gravity or creatinine substantially different from human urine. The licensee determined that the individual had a fake bladder strapped inside his thigh with a surrogate sample.

In 1993 a nuclear utility reported that a licensed operator attempted to provide a substitute sample during FFD testing. Later in 1993, the utility reported that another licensed operator submitted a cold test sample. The employee was subsequently required to provide a witnessed sample, but resigned before receiving the results of the witnessed sample (NUREG/CR-5758, Vol. 4, [Westra et al., 1994 p. C-12]).

### 6.3.2: Specimen Hydration and Dilution

The general consensus among organizations providing advice on subverting drug testing is that specimen hydration and dilution are the preferred strategies (see section 6.2.2, above). Evidence from laboratories provides some support to the popularity of hydration and dilution as techniques for subversion. As reported by Gary Hemphill of MEDTOX Laboratories, the number of suspect specimens identified by his laboratory has risen from a rate of about 3 percent to a rate of about 5 percent. Although MEDTOX also detects adulteration, hydration is the technique that is most often used and is the most successful at producing false negatives (G. Hemphill, personal communication to Loren Bush, NRC, 1994).

Findings from two studies suggest that some specimen donors are hydrating or diluting specimens to avoid drug use detection. The first study, by the National Laboratory Certification Program (NLCP), examined the effects of using dilution as an indicator of attempted subversion (*MRO Newsletter*, 1993). This study examined a sample of 10,000 Department of Transportation (DOT) specimens that were selected at random from DOT-regulated companies in a 6-day period. Of these samples, 758 (7.6%) were found to be "dilute" by either creatinine or specific gravity tests, using NIDA cutoff levels.

The 758 dilute DOT specimens were compared to a second control group of 758 DOT specimens of normal concentration. Of the dilute specimens, 75 (9.9%) tested positive for traces of any of five drugs, and nine of these (1.2%) were positive using the DHHS cutoff levels. The comparable control specimen positive rates were 2.8% (n=21) and 0.3% (n=2), respectively. That is, the dilute samples were three times more likely than the control samples to test positive for trace amounts of drugs, and four times more likely to test positive by DHHS cutoff levels. However, 90 percent of dilute samples contained no trace of drugs.

This study also examined the rate of dilution among samples obtained from the DOT drug testing program and among a control group of 200 samples from volunteers not subject to drug testing. The rate of dilution (7.6%) among those included in the testing program was 2.5 times higher than that of the control group (3%; N=6).

Finding that 90 percent of the specimens with low creatinine contained no detectable quantities of drugs indicates that many donors who did not attempt to subvert the testing process were nevertheless out of range in a

## Subversion and Countersubversion

creatinine test for dilution. It appears that these findings are attributable largely to the natural variation in dilution/concentration levels in "normal" urine. The natural variation also accommodates consumption of large quantities of fluids due to heat or exercise. Thus, it is likely that any other dilution analysis will also have this disadvantage. Based on these findings, the NLCP concluded that creatinine and specific gravity were not efficient indicators of specimens that contained drugs (*MRO Newsletter*, 1993).

The second study, conducted by Mayer and Hemphill,<sup>5</sup> used 7,705 urine specimens collected in pre-employment and random workplace drug testing programs. These specimens were tested for cocaine, marijuana, and creatinine content. Creatinine levels below 20 mg/dL (milligrams per deciliter) were considered sufficiently low to warrant suspicion. Specimens initially screened as negative for drugs (at standard NIDA cutoff levels) and as positive for low creatinine were screened a second time using lower cutoff levels to determine if detectable traces of drugs or drug metabolites were present.

The study results indicate that low creatinine levels are associated with increased frequency of trace drugs and their metabolites in urine (i.e., levels indicating recent use but below the standard cutoff levels). However, the results also suggest that creatinine analysis will indicate suspect urine in many cases where no subversion has been attempted. About three percent of all specimens examined (238 specimens) had creatinine levels below the cutoff level, while only 0.3 percent of all specimens had both low creatinine levels and trace amounts of drugs or drug metabolites present. Thus, in 90 percent of the cases with low creatinine, no traces of drugs or metabolites were found in the urine. This means that dilute specimens are not necessarily indicative of subversion.

These studies suggest that dilute specimens are more common among populations being tested for drugs than among populations that are not subject to drug testing. In addition, they indicate that dilute specimens are substantially more likely than non-dilute specimens to show evidence of drug metabolites when level of detection (LOD) testing is used. However, they also indicate that approximately 9 out of 10 specimens that are identified as dilute show no evidence of drug metabolites. Caution should be used in generalizing these findings, since rates of dilution and of positive drug tests will vary across industries, types of drug test (e.g., random vs. pre-employment tests), and other drug testing program characteristics. In addition, the second study was limited to two drugs, cocaine and marijuana, and the results

might have differed if tests had been conducted for a more extensive panel of drugs.

In a few cases NRC licensees identified specimens that had low specific gravity and/or creatinine and that subsequently were found to be positive for drug metabolites. Licensees attributed these outcomes to hydration or flushing. Although it is difficult to demonstrate that hydration is the outcome of an intentional subversion attempt (as noted earlier, hydration may be the result of a number of legitimate causes), these provide possible evidence of the use of hydration in the nuclear industry.

### 6.3.3: Other Evidence of Subversion

NRC licensees have identified a number of cases in which adulterants were used to subvert the testing process. Instances of samples submitted with antifreeze, soap, Mountain Dew, and other adulterants have been reported to the NRC (see Appendix F).

Although precise estimates are lacking, there is anecdotal and empirical evidence that subversion exists in drug testing programs in a number of other settings: programs in rehabilitation settings, professional sports drug testing programs, and other workplace drug testing programs (e.g., Person and Ehrenkranz, 1987).

There is some evidence that workers may favor specific subversion techniques that have proven useful to their colleagues. Laboratories testing for the specific adulterant UrinAid®, for example, have found that positive test results for this adulterant occur in batches: there are no positives at some sites and relatively large numbers at others (Jack Smith, Chimera Research and Chemical Inc., personal communication, 1993). This suggests that drug users in these workplaces are sharing strategies or drawing on the same information to subvert the tests.

It is likely that the testing process will foster interest in ways to subvert it by some people required to undergo testing. Coombs and Ryan (1990) found that college athletes who were subjected to drug testing programs were much more knowledgeable about subversion techniques than were those not tested. For example, 18.8 percent of those athletes who were subject to testing "strongly agreed" with the statement, "I know ways to 'beat the system' by timing my drug use," while only 3.7 percent of those not subject to testing agreed with this statement. These findings suggest that those athletes subject to testing had made the effort to learn about ways of escaping detection. Consequently, in any testing

population it is reasonable to suspect that there will be some drug users who seek to subvert the testing process.

#### 6.3.4: Summary

Licensees are required to report all instances of subversion by supervisors and licensed operators, but are not required to report all instances of subversion in the general workforce. Hence, evidence on subversion in NRC-regulated FFD programs is limited to subversion by licensed operators and supervisors. Although available evidence suggests that subversion is possible and that it does happen, the full extent of subversion is unknown in NRC-mandated FFD programs.

In order to better understand the extent of subversion, information is needed on the type and frequency of subversion activities that drug users engage in, rather than relying solely on anecdotal reports of subversion or inferring subversion from large-scale analyses of urine specimens. Despite a lack of understanding on the extent of subversion, there are a number of available preventive and analytic methods to deter and detect subversion in FFD programs. These countersubversion methods are discussed in the following section.

### 6.4: Countersubversion Strategies

Countersubversion measures have been developed to prevent evasion of testing and to prevent or detect specimen substitution, dilution, or adulteration. This section provides a review of countersubversion strategies to combat known and suspected subversion techniques. This includes a summary and, where appropriate, an update of previous examinations of countersubversion techniques for workplace drug testing in NUREG/CR-5227 (Barnes et al., 1988), NUREG/CR-5227, Supplement 1 (Moore et al., 1989) and NUREG/CR-5784 (Durbin et al., 1991). Two new analytical countersubversion techniques are discussed: the use of a specific test for adulteration and tests for the ionic strength of specimens.

#### 6.4.1: Overview of Countersubversion Measures

There are two types of countersubversion measures:

- preventive measures, intended primarily to deter subversion and/or render it ineffective, and

- detection measures, intended to identify subversion attempts after the fact.

This distinction is not absolute because the use of a detection measure can be expected to deter, and thus prevent, some drug users from attempting the subversion.

Table 6.2 lists the most common techniques used in attempts to subvert the urine testing process, and the countermeasures that can be used to mitigate each technique. This table is based on information previously reported in NUREG/CR-5784 (Durbin et al., 1991) and augmented with information from the *MRO Newsletter* (1995). The table is not intended to provide a comprehensive listing of subversion techniques, their potential effectiveness, or every countersubversion measure that might be used. Important limitations of this summary include the following:

- The adulterants listed in the table represent only a subset of those that may be used. The ones listed generally represent those that are well known.
- Studies examining the effects of all the adulterants on all testing technologies have not been identified. Thus, some of the adulterants may cause false-negative results for drugs or testing technologies other than those listed here.
- Countersubversion measures such as "security of the collection site" and "minimizing notification" have only been listed in the table where it is believed they are especially relevant; measures such as these provide some protection against virtually all of the subversion techniques listed here.

#### 6.4.2: Preventive Countermeasures

Ideally, countersubversion measures should be preventive. Measures should deter subversion attempts by making them impossible or extraordinarily difficult. Preventing subversion requires implementing testing procedures that preclude obvious ways to conceal drug use. Effective and

reliable preventive countersubversion measures include the following:

- Random and unannounced testing.
- Short times between notification and arrival at the collection site that make hydration or obtaining a

## Subversion and Countersubversion

**Table 6.2: Subversion Techniques and Countersubversion Measures**

Subversion Technique	Drug Test That Can Be Affected	Possible Countersubversion Measures	
		Preventive	Detective
Avoid or delay testing	All	<ul style="list-style-type: none"> <li>No prior warning and minimized time from notification to collection</li> </ul>	<ul style="list-style-type: none"> <li>Check out cases where testing has been delayed or not been conducted</li> </ul>
Substitution of drug-free urine specimen	All	<ul style="list-style-type: none"> <li>Security of collection site</li> <li>No prior warning and minimized time from notification to collection</li> </ul>	<ul style="list-style-type: none"> <li>Specimen temperature</li> </ul>
Substitution of something other than urine for specimen	All	<ul style="list-style-type: none"> <li>Security of collection site</li> <li>No prior warning and minimized time from notification to collection</li> </ul>	<ul style="list-style-type: none"> <li>Specimen temperature, pH, specific gravity</li> <li>Nonanalytic inspection (color, carbonation)</li> </ul>
Hydration/ Dilution	All	<ul style="list-style-type: none"> <li>No prior warning and minimized time from notification to collection</li> </ul>	<ul style="list-style-type: none"> <li>Specimen specific gravity, creatinine content</li> </ul>
Salt adulteration	EIA <sup>a</sup> for all drugs <sup>b</sup>		<ul style="list-style-type: none"> <li>Specimen specific gravity</li> <li>Nonanalytic inspection (visual residue)</li> </ul>
Adulteration with oxidizing agents (bleach)	EIA for all drugs RIA <sup>c</sup> for Amphetamine, Opiates, and PCP		<ul style="list-style-type: none"> <li>Specimen pH, peak temperature (or temperature taken very shortly after the adulterant is added)</li> <li>Nonanalytic inspection (smell)</li> </ul>
Adulteration with soap	EIA for Barbiturates, Benzodiazepines, Marijuana, PCP		<ul style="list-style-type: none"> <li>Specimen pH, ionic strength, specific gravity (not all measures necessarily effective for all soaps)</li> <li>Nonanalytical inspection (cloudiness, bubbles)</li> </ul>
Sodium Bicarbonate adulteration	EIA for Opiates, PCP		<ul style="list-style-type: none"> <li>pH</li> </ul>
Vinegar adulteration	EIA for Marijuana		<ul style="list-style-type: none"> <li>pH</li> </ul>
Visine <sup>®</sup> adulteration	EIA for Marijuana		<ul style="list-style-type: none"> <li>Positive control assay</li> </ul>
Golden Seal Tea adulteration	EIA for Marijuana		<ul style="list-style-type: none"> <li>Visual inspection</li> </ul>
UrinAid <sup>®</sup> adulteration	EIA for Marijuana, possibly EIA for all other drugs		<ul style="list-style-type: none"> <li>Positive control assay, aD Perfect<sup>®</sup> (marketed by Chimera)</li> </ul>

a. EIA = enzyme immunoassay.

b. In this table "all drugs" refers to amphetamines, barbiturates, benzodiazepines, cocaine, marijuana, opiates, and PCP.

c. RIA = radio immunoassay.

substitute specimen or adulterant in preparation for a drug test very difficult.

- Tight security of information about testing to ensure that there are no "leaks" alerting workers as to when testing might be expected.
- Collection site security which is intended to assure that no adulterants or substitutions can be brought in.
- Control of handling and storage of specimens and chain-of-custody procedures to remove opportunities to tamper with specimens after they are collected and sealed.

#### **6.4.3: Detection Measures**

Countersubversion techniques must be able to detect subversion that is not prevented or deterred. Detection countermeasures consist of general observation techniques, nonanalytical examination of specimens, specific analytic techniques, and lower cutoff levels.

##### **6.4.3.1: Observation Techniques**

Observation can play two roles in an effective countersubversion program. First, collection site personnel can be trained to watch for indications of attempted subversion. Any unusual behavior can be noted on the chain-of-custody form. Second, urine specimens can be collected under direct observation when there is reason to believe that the donor may have attempted to subvert the process when providing a previous specimen. In addition, program managers can note whether individuals selected for testing have not been tested as scheduled. Observation may yield indications of attempted substitution, such as indications of attempts to conceal a sample or providing the specimen quickly. To detect dilution through observation, collection site personnel can be on the lookout for circumstances that would be suggestive of specimen dilution, such as the availability of water. Adulteration may be detected through observation of foreign materials at the collection site or circumstances which might mask the presence of adulterants (such as baggy clothing in which adulterants could be concealed).

##### **6.4.3.2: Nonanalytical Examination of Specimens**

Nonanalytic techniques include examining specimens for any abnormalities. While these techniques can be an important part of countersubversion efforts, it would be inappropriate to impute subversion to individuals based

solely on these indicators alone. A collection site person can note whether specimens appeared clear or cloudy, contain abnormal bubbles or have unusual odor, and alert the supervisor or the Medical Review Officer. The NRC FFD rule requires licensees to visually inspect urine specimens for color and clarity to identify suspect specimens. However, visual inspection is of limited efficacy in detecting subversion. Normal urine will range in color from very light yellow to dark amber or even slightly pink, and from clear to turbid depending on concentration of the urine and the health of the donor. A visual examination of a specimen may reveal indications (such as unusual color or carbonation) of substitution of another substance. Thus, while any visual anomalies can be noted on the chain-of-custody form, they are not normally considered strong evidence of subversion. Other means of detecting suspect specimens include shaking specimens to look for soap-type bubbles or noting the presence of unusual smells such as bleach or soap which may reveal tampering. Solid matter in normal urine is highly atypical. Although a collection of kidney or bladder stones is possible, observed solid matter is much more likely to be undissolved adulterant material.

##### **6.4.3.3: Analytical Examination of Specimens**

A number of analytic techniques are available including examining the temperature, pH, specific gravity, and creatinine of the specimen. In addition, specific tests are available for adulterants. The following sections provide a summary of analytical techniques to identify subversion attempts.

##### **Temperature**

Checking the temperature of urine specimens at the collection site can detect various subversion methods. The sensitivity of this test for subversion will depend on both the acceptable temperature range and the precision of the measurement technique. The broader the temperature range, the easier it will be for the drug user to provide a substitute specimen that falls within the accepted interval. Current testing standards of the DHHS set a lower temperature limit of 90°F, which may not permit detection of many subversive substitute samples (Person and Ehrenkranz 1989). However, such broad ranges may be necessary if measurement techniques are relatively imprecise and have large margins of error, or if collection techniques allow passage of several minutes between sample delivery and measurement of temperature, requiring an adjustment for possible differential cooling.

## Subversion and Countersubversion

Temperature measurement with more precise instruments or more efficient procedures can support the use of a narrower temperature range while minimizing the risk of false attributions of subversion. Precision and efficiency reduce the amount of variability in measures, which the range must expand to accommodate. A narrower temperature range would reduce the chance that a substitute or adulterated specimen was judged to be of acceptable temperature.

Peak urine temperature provides a more useful measurement than a temperature taken with a normal thermometer after the specimen has been given to the specimen collector (NUREG/CR-5784 [Durbin et al., 1991]). Peak temperature measurement uses a thermometer that locks onto and reports the highest temperature achieved by the specimen. Use of a peak temperature thermometer may permit a narrower temperature range, because the acceptable range need not include a large adjustment to accommodate possible cooling of the urine between the times of voiding and measurement.

Some thermometers have created certain difficulties. Collectors in one NRC utility noted that use of collection bottles with temperature strips gave specimen providers the chance to check whether a subverted specimen was of acceptable temperature. The collectors changed the procedure to require that the specimen be delivered in a plastic cup, which was immediately emptied into the collection bottle for the temperature check (*Drug Detection Report*, 1993). This reduced the chance of subversion by substitution, but might dictate a slightly broader temperature range to adjust for differences in the speed of specimen transfer.

*Temperature for detecting substitution.* Measurement of peak urine temperature may detect substitution. If a pre-warmed substitute specimen is provided, the peak temperature thermometer will report the temperature of the specimen when it entered the container, rather than its temperature when tested a few minutes later, after the donor has allowed the specimen to cool to the normal temperature.

*Temperature for detecting dilution.* Dilution of a specimen may be detected by checking specimen temperature if the substance used to dilute the specimen was of a temperature significantly higher or lower than the normal range for urine.

*Temperature for detecting adulteration.* Peak temperature measures are effective in detecting adulterants, such as

bleach, that cause chemical reactions with urine that raise the temperature of the specimen.

### pH

Checking the pH of a fluid measures its relative acidity or alkalinity and can be used to determine whether a specimen is within the normal range of alkalinity for urine. pH is measured on a scale ranging from 1 to 14; a reading of 1 is highly acid, 14 is highly alkaline and fluid with a pH of 7 is neutral. Normal urine is slightly acidic. The normal pH range for urine is 4.8 to 7.8. Measuring the pH of a collected urine sample can be accomplished with the use of an inexpensive disposable test strip. Chimera Labs also offers testing kits for pH that can be used in some current immunoassay testing instruments (Jesse Carter, Chimera Labs, personal communication, March 20, 1996).

Natural body changes can produce an altered urine pH that may differ significantly from the normal range (Free and Free, 1975). For example, conditions that may cause pH to be out of range include diabetic ketosis, respiratory acidosis, emphysema, phenylketonuria, potassium depletion, GI tract infection, hyperventilation, and prolonged vomiting. pH testing only indicates that the specimen is not within the normal range for urine—not that subversion was attempted.

*pH for detecting substitution.* pH testing may detect the possible substitution of a substance with a pH significantly different from that of urine.

*pH for detecting dilution.* pH testing can detect possible *in vivo* adulteration with substances intended to alter urine pH by increasing the rate of urine excretion.

*pH for detecting adulteration.* The effectiveness of pH in detecting adulteration depends on both the pH and the quantity of the adulterant used. While pH is most effective in detecting adulterants that are extremely acidic or alkaline, it is possible to detect adulterants with a pH closer to that of urine if a relatively large quantity of the adulterant is added.

### Specific Gravity

Specific gravity (SG) is the ratio of dissolved solids to water in liquid. Normal urine specific gravity is between 1.010 and 1.030. Although SG may indicate attempts to subvert, there are legitimate reasons for variation in SG (Free and Free, 1975). The following conditions can cause specific gravity to be out of range: salt-poor diet,

low protein diet, diabetes insipidus, potassium deficiency, bone disease, hyperparathyroidism, and dehydration.

Disposable test strips can be used for specific gravity testing, much like the simple disposable test strips for testing pH. Products for testing SG may place the specific gravity indicator on a strip that measures a number of other factors as well. The use of these multi-purpose strips is relatively expensive, and creates the risk that the collection site person will obtain additional (confidential) medical information.

Other specific gravity tests available are not overly complex or expensive for on-site testing, and can avoid the collection of confidential information. Collection site personnel at one NRC licensee site modified an Ames Clinitek machine to read only specific gravity and pH from a multipad test strip. However, they found its validity in determining dilution was unacceptably low due to insufficient precision. The licensee now tests for specific gravity on-site with a refractometer, which is relatively inexpensive (approximately \$300), reliable, easy to operate, and adequately precise (Michael Calloway, San Onofre, personal communication, 1993). Chimera Labs also offers testing kits for SG that can be used in some current immunoassay testing instruments (Jesse Carter, Chimera Labs, personal communication, March 20, 1996).

Care must be taken to avoid cross-contamination (e.g., instruments must be cleaned between tests). Decanting a portion of the specimen for specific gravity analysis may be an appropriate practice to ensure that the specimen to be forwarded to the lab does not become contaminated. Disposable pipettes may be used to obtain a drop of urine for analysis (Michael Calloway, San Onofre, personal communication, 1993).

*Specific gravity testing for substitution.* Tests can indicate the presence of a substitute substance that has a specific gravity significantly different from that of urine.

*Specific gravity testing for dilution.* Hydration or dilution can be detected by specific gravity tests. Specific gravity below 1.003 may suggest hydration or dilution.

*Specific gravity testing for adulteration.* Adulteration with dissolved solids, such as salt, will yield a high specific gravity.

### Creatinine Testing

Creatinine, an anhydride of creatine, is a normal alkaline existing in urine and blood. Creatinine measures in urine will drop when a person consumes large quantities of fluids (or when urine is diluted with a substance that does not contain creatinine, such as tap water). Creatinine levels below 20 mg/dL are considered below the normal range.

Creatinine testing is relatively complex and typically is performed by a laboratory. Chimera Labs now offers testing kits for creatinine that can be used in some current immunoassay testing instruments (Jesse Carter, Chimera Labs, personal communication, March 20, 1996).

*Creatinine testing for detecting substitution.* Creatinine testing that yields a measurement of creatinine below 20 mg/dL may indicate substitution of a substance that has a low creatinine level.

*Creatinine testing for detecting dilution.* Hydration or dilution can result in a measurement of creatinine below 20 mg/dL.

*Creatinine testing for detecting adulteration.* No known adulterants are detected by creatinine tests.

### Tests to Detect UrinAid®

The Department of Health and Human Services is preparing to fund research to evaluate the impact of UrinAid® on the testing of urine specimens. UrinAid® is an adulterant which was sold specifically by Byrd Laboratories to cause false-negative results on EIA marijuana tests. As discussed in Section 6.2, this adulterant may also cause false-negative test results on immunoassay tests for most drugs included in workplace drug testing programs. Its active component, glutaraldehyde, is also thought to interfere with the recovery of the drug through gas chromatography/mass spectrometry analysis (GC/MS) and possibly destroy THC, a marijuana metabolite.<sup>6</sup>

Currently there are few effective countersubversion mechanisms identified to detect the adulterant UrinAid®. Labs can use aD Perfect®, a product of Chimera Laboratories specifically designed for detecting UrinAid® and other adulterants such as Visine®. According to the

## Subversion and Countersubversion

manufacturer, the aD Perfect® test is relatively inexpensive and suitable for on-site testing (Jack Smith, Chimera Research and Chemical Inc., personal communication, 1993). A study found that this product identified UrinAid® consistently—with no false positives or false negatives (Goldberger, 1994).

### Ionic Strength

Tests for ionic strength will detect the presence of certain soaps that are effective urine adulterants and may raise the ionic strength of the specimen. Specific tests for ionic strength are not necessary if the specimen is tested for specific gravity. Changes in ionic strength will also change the specific gravity of the specimen. (Jesse Carter, Chimera Chemicals and Research Inc., personal communication, March 20, 1996).

### Positive Control Assay

A positive control assay is a general test for adulteration. This test is relatively expensive and is typically performed by a certified laboratory. This test is effective at detecting the use of adulterants to mask drugs and their metabolites. The analyst adds a known quantity of drug metabolite to an aliquot of the specimen. If the sample tests negative, something is interfering with detection (Dr. Arthur Zebelman, Laboratory of Pathology of Seattle, Inc., personal communication, 1994).

The costs of performing this test routinely may be prohibitive. Although this type of test can detect all effective adulterants, the nature of this test precludes it from identifying the specific adulterant being used. However, if positive control assays indicate subversion, they provide a nonspecific warning of some type of adulteration.

#### 6.4.3.4: Lower Cutoff Levels

Hydration and dilution are effective because they decrease the concentration of drugs in the specimen—often below the cutoff levels specified by DHHS. Although DHHS provides cutoff levels for drug testing, DHHS certified laboratories can accurately perform tests at lower levels. The lowest level at which a lab can reliably determine the presence of a drug or its metabolite is referred to as the level of detection (LOD) for the lab for that drug. Using the LOD as a cutoff level for drug testing decreases the effectiveness of hydration or dilution as a subversion technique.

DHHS guidelines have lowered the cutoff level for marijuana screening tests from 100 nanograms per milliliter (ng/mL) to 50 ng/mL (see Chapter 2 for a discussion of changes to DHHS guidelines). This change would hinder but not stop the practice of *in vivo* hydration. Because the NRC FFD rule specifies only the maximum cutoff levels that licensees may use, and allows licensees to use lower cutoff levels at their discretion, lower cutoff levels are available to licensees under the NRC's FFD regulation.

### 6.4.4: Discussion of Analytic Countersubversion Measures

Analytic procedures can be an effective means of detecting indications of subversion. Some of these techniques can be appropriately performed on site, others are more appropriately performed at a DHHS-certified laboratory.

Using analytic techniques to test for subversion at the collection site has both advantages and disadvantages. The primary advantage of testing onsite is that possible indications of subversion are obtained immediately. In some cases, analytic methods may be more effective when they provide immediate evidence of subversion because follow-up actions can be taken before the donor leaves the collection site. Analytic procedures that may be appropriate for on-site testing are temperature, pH, and SG. They are not time consuming, do not require unduly expensive equipment, and do not require that the person performing the test have training beyond that of a certified laboratory technician.

The disadvantages of testing onsite are: 1) a possible increase in opportunities to compromise specimen integrity and the need for stringent quality assurance procedures; 2) the potential for inappropriate actions to be taken by collection site personnel; and 3) on-site resource limitations on testing options. Operational procedures can be used to ensure that specimen handling does not compromise specimen integrity. If not performed properly, actions such as decanting, test striping, or pipetting of a specimen could contaminate a specimen. Potential contamination can be minimized by the use of new, sealed, disposable materials (e.g., disposable test strips and disposable thermometers). Persons performing the tests should have credentials demonstrating that they know how to care for specimens in a manner that ensures against contamination. While training specifically designed for collection site persons may be adequate, it

may be advantageous to have a trained laboratory technician perform some tests onsite, especially if multiple tests are being performed. It is important to make collection site personnel aware that not all specimens that are outside of normal specifications are indications of subversion.

Other tests such as creatinine tests and positive control assays will generally need to be performed in a clinical laboratory. Performing these procedures may require resources beyond those typically available at a collection site. If specimens are analyzed at the drug testing laboratory, there may be a considerable time delay. The Medical Review Officer may not be notified of suspected subversion for five to seven days after the specimen is provided. This may allow sufficient time for the drugs to be metabolized and excreted, eliminating any possibility of detection during recollection of a specimen.

Regardless of where the testing is conducted, analytic tests for subversion can in most cases identify grounds for concern, not definitive evidence of subversion itself. Determining subversion requires additional information and expertise in interpreting findings. If there are grounds to suspect subversion, the following actions may be appropriate:

- Notify an MRO or other appropriate professional.
- Collect an additional specimen under direct observation (e.g., if adulteration or substitution is suspected).
- Use a different analysis methodology to render the subversion ineffective (e.g., if hydration is suspected, the specimen could be tested using lower-than-standard cutoff levels or LOD). Alternatively, test all specimens at lower cutoff levels or LOD.

## 6.5: Subversion Issues

Two issues regarding subversion are determining the prevalence of subversion in the workplace and protecting workers' rights when countersubversion strategies are employed.

### 6.5.1: Measuring Prevalence

There is little direct information on the extent of subversion (Carter, 1993). Measuring the extent of subversion directly is difficult; it relies on detecting people in the act or identifying specimens that show

evidence of dilution or adulteration. Indirect methods (such as anonymous self-report surveys of drug users and their subversion habits) have not been conducted.

In the absence of concrete knowledge, some beliefs about subversion effectiveness have acquired the status of accepted truths despite questions about their accuracy (Cody, 1990). The few data available on countersubversion are often of doubtful quality. As one researcher stated, "In some cases the studies (of countersubversion testing methods) ... are either very limited in scope, or the data is inconclusive, misinterpreted, or open to numerous interpretations" (Carter, 1993, p. 1). This lack of reliable data limits the ability to reach informed decisions about the efficacy of countersubversion measures.

Evidence on the existence and prevalence of subversion generally derives from cases of detected or *unsuccessful* subversion. A low incidence of observed subversion suggests three possibilities: 1) there is little attempted subversion present, 2) the attempted subversion that is present is successful and therefore not being detected, or 3) awareness and measures to detect or prevent subversion are nonexistent.

Individuals whose subversion attempts are detected often state that they had successfully subverted the system on previous occasions prior to being caught, or testify that they know of others who are currently successful in subverting the system (Coombs and Ryan, 1990). This is also consistent with information provided to the NRC by nuclear facility licensees (Mr. Loren Bush, NRC, personal communication, 1993). Consequently, it is likely that the actual rates of subversion are higher than the limited data might suggest. For instance, whether worksite-level knowledge about successful subversion techniques may increase subversion's effectiveness and thus reduce its detection and apparent rate of occurrence is unknown. If a workplace drug testing program tests for adulteration of urine specimens but does nothing to test for dilution of specimens, it is possible that alternative techniques such as dilution will be preferred in order to escape detection. Knowing that adulteration is uncommon is no guarantee that other types of subversion are not underway.

Finally, indications of subversion are not necessarily evidence of subversion. Subversion is the intent to evade detection of drug use. Except in the case of direct observation of subversive actions, most analytical findings are not able to discern the intent behind a suspicious finding. It is necessary that the judgment that subversion has taken place be made by the Medical Review Officer

or other qualified professional who reviews all available evidence, such as: observations made by collection personnel (unusual behavior), analysis of the specimen for validity, and indications of possible subversion, if any.

In particular, limited empirical evidence available supports the conclusion that any sensitive test for hydration/dilution will result in a significant number of specimens out of the range of values considered normal when no subversion has been attempted. As discussed previously in Section 6.3, while studies showed that specimens considered dilute were more likely to test positive for drug metabolites than those specimens that were not dilute, approximately 90 percent of dilute specimens showed no indications of drug metabolites. Hence, as with pH, tests for dilution only provide an indication of potential subversion. The amount of fluid that a drug user would need to consume to lower the concentration of drugs and/or drug metabolites in the urine is a quantity well within the range some people consume regularly or on certain occasions. For example, forcing fluids is medically recommended for fighting the common cold and other viral infections, and is an integral component of many weight loss programs. Consuming large quantities of fluids to replace those lost during intense exercise or heat are also common occurrences. There are many legitimate reasons for people to use both prescription and over-the-counter diuretic medications, which will result in dilute urine. Consequently, any test for hydration will also detect cases in which the donor has consumed relatively large quantities of fluid for legitimate purposes.

### 6.5.2: Workers' Rights

The actions taken to control subversion must be balanced with the need to maintain worker support of the program and protect workers' rights. Countersubversion practices that are unnecessarily intrusive or result in false accusations of subversion may do more harm than good.

In cases where specimens are out of the prescribed range of temperature, specific gravity, pH, or other specimen quality measure, the finding does not necessarily indicate a confirmed case of subversion, but only that a specimen is unsuitable for the intended analysis. It is important to pursue unsuitable specimens with full respect for the donor's rights and with the intention of obtaining a reliable analysis. Leaving workers to feel accused of any wrongdoing because of these preliminary test results may harm employee support of the testing program. It is appropriate to inform a worker that, because the urine is not within the parameters used on the given test, some

specific action needs to be taken in order to ensure a reliable test result (e.g., an additional sample collected at a later time or testing the original sample at lower cutoff levels or LOD). Alternative reasons why a specimen is dilute or otherwise abnormal should be fully explored by an MRO or other qualified physician when subversion is suspected.

## 6.6: Summary and Conclusions

Subversion can undermine the effectiveness of FFD programs. Subversion may be more common where it is relatively easy and promises to be successful. If drug users know that the tests can be easily beaten, they will probably be more likely to try subverting the testing process rather than abstaining from unacceptable drug use. Conversely, effective and well-publicized countersubversion measures may contribute to overall program effectiveness and deter drug users from using drugs altogether if they know they cannot "beat the system."

Based on this review of subversion, it is apparent that there are a number of opportunities to subvert the testing process, and that attempts to subvert do occur.

Appropriate actions to guard against subversion include:

- reviewing the program from the perspective of possible subversion opportunities,
- reviewing the current countersubversion measures with respect to possible subversion opportunities,
- incorporating additional preventive measures as needed, and
- ensuring that specimen collection personnel and other personnel in the FFD program are trained to detect subversion, conduct nonanalytic examinations of specimens, and employ analytic techniques to detect subversion (both onsite and at contract laboratories).

Dilution (either *in vivo* or *in vitro*) appears to be a well-known and effective subversion technique. Procedures that minimize the time available to consume fluids after the notification of a test, and analytic procedures that test for creatinine and specific gravity, can be used to deter and detect dilution. However, most instances of dilution are probably due to factors other than subversion. Care must be taken to investigate whether dilution is due to legitimate causes such as rehydrating after strenuous exercise or drinking lots of fluids because of a cold.

Analyzing suspicious specimens or all specimens at lower cutoff levels or LOD may be a more appropriate response to dilute specimens than responses such as an observed collection of a second specimen. If this practice is followed, all appropriate quality assurance measures should be taken to ensure that the analyses are reliable at the lower levels.

Adulteration and substitution are other subversion techniques of concern. Countering adulteration and substitution is possible by ensuring security of the collection site and conducting the following on-site examinations:

- Minimization of time between notification and collection.
- Visual inspection of the specimen for unusual color, undissolved solids, soap bubbles or film, and other highly suspicious uncharacteristic properties (but not for cloudiness).
- Notation of any unusual smell, such as from bleach or other chemicals.
- Test for specimen temperature using as narrow a band as technically feasible.
- Test for specimen specific gravity.
- Test for specimen pH.

It is important to protect employee rights when subversion is suspected. Care should be taken not to undermine the integrity of a program by misinterpreting findings from these tests. All abnormal test results or suspected cases of subversion should be reviewed by a Medical Review Officer. Follow-up actions that are

unnecessarily inconvenient or intrusive are undesirable. Direct observation is regarded by many workers as highly intrusive and can undermine support for the program (and therefore its effectiveness). Hence, direct observation would be appropriate if a specimen's unsuitability is consistent with adulteration, tampering, or substitution, but not as appropriate if the specimen's unsuitability is consistent with dilution.

Given the difficulty of detecting subversion, it is useful to confer with other FFD program managers regarding their experiences with subversion, and sharing experiences on effective countersubversion methods. Toward this end, results of all detection analyses could be recorded and periodically analyzed to gain better information on subversion rates, patterns, and the efficacy of the detection techniques themselves. By sharing program experiences on subversion with FFD program managers, the NRC and other regulatory agencies can provide a better industry-wide understanding of subversion techniques, their extent, and how to counter them.

For instance, in order to facilitate uniform interpretation of laboratory reports of abnormal test results and analysis for the presence of adulterants, the U.S. Department of Transportation (DOT) provides guidance on the language to be used on laboratory reports or in the remarks section of the drug testing custody and control form for DOT-mandated drug testing specimens (DOT, 1993). Similar guidance from the NRC for licensee program performance reports may be desirable.

Increasing the understanding of subversion should improve the effectiveness of drug testing in the workplace. In fact, until the extent of subversion and its effects on drug test results are better understood, knowledge of the true extent of drug abuse will remain unknown.

## Endnotes

1. Chris Moore, M.S., Battelle Seattle Research Center  
Nancy Durbin, Ph.D., Battelle Seattle Research Center  
Jennifer Macaulay, Ph.D., Battelle Seattle Research Center
2. The maximum screening and confirmation levels required by 10 CFR Part 26 are:

Drug	Screening Level	Confirmation Level
Marijuana	50 ng/mL	15 ng/mL
Cocaine	300 ng/mL	150 ng/mL
Opiates	300 ng/mL	300 ng/mL
Phencyclidine	25 ng/mL	25 ng/mL
Amphetamine	1000 ng/mL	500 ng/mL
Alcohol	.04% BAC	.04% BAC

3. Licensee event reports are formal reports to the NRC and are publicly available.
4. This product is a hand-wrap that uses a chemical reaction to create heat so it need not be conspicuously plugged into an electric socket.
5. Data from this study were originally presented at the poster session on July 22, 1993, at the American Association for Clinical Chemistry's annual meeting. Since that time, the study has stirred considerable controversy and some of the data has been published in the following sources, which have been drawn upon in this report: *Drug Detection Report*, Sept. 8, 1992 and Carter, 1993.
6. For a discussion of testing techniques for drugs of abuse, see NUREG/CR-5227 (Barnes et al., 1988).

## References

Barnes, V., Fleming, I., Grant, T., Hauth, J., Hendrickson, B., Kono, B., Moore, C., Olson, J., Saari, L., Toquam, J., Wieringa, D., Yost, P. (1988) NUREG/CR-5227 *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues*.

Carter, J. (1993). Adulteration testing: Fact and fiction. *MRO Newsletter*, 2(1), 1-2.

Cody, J. (1990). Specimen adulteration in drug urinalysis. *Forensic Science Review*, 2(1) 64-75.

Cody, J. and Schwarzhoff, R. (1989). Impact of Adulterants on RIA analysis of urine for drugs of abuse. *Journal of Analytical Toxicology*, 13, 277-284.

Coombs, R. and Ryan, F. (1990). Drug testing effectiveness in identifying and preventing drug abuse. *American Journal of Drug and Alcohol Abuse*, 16, 173-184.

Department of Transportation. (1993, December 7). Reporting of Drug Test Results: Abnormal test results and Analysis for presence of adulterants. Washington D.C.: Author.

*Drug Detection Report*. (1992). "Nuke workers caught warming urine samples in microwave." 2(6), 1 and 4.

*Drug Detection Report*. (1992). "Study suggests changing standard for dilute samples." 2(11), 4-5.

*Drug Detection Report*. (1993). "Nuclear plants take measures to tighten testing programs." 3(14), 1 and 6.

Durbin, N., Moore, C., Grant, T., Fleming, T., Hunt, P., Martin, R., Murphy, S., Hauth, J., Wilson, R., Bittner, A., Bramwell, A., Macaulay, J., Olson, J., Terrill, E., Toquam, J. (1991). NUREG/CR-5784 *Fitness for Duty in the Nuclear Power Industry: A Review of the First Year of Program Performance and an Update of Technical Issues*.

Edgell, K. (1993). DOT random experience. DOT Meeting on Random Drug Testing, Washington D.C., February.

Fleming, T., Westra C., Field I., Forslund C., Durbin N., Grant, T. (1993). NUREG/CR-5758, Vol 3, *Fitness for duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports, CY 1992*.

Free, A. and Free, H. (1975). *Urinalysis in clinical laboratory practice*. Cleveland, OH: CRC Press.

Goldberger, B. (1994). Effects of glutaraldehyde (UrinAid) on detection of abused drugs in urine by immunoassay. *Clinical Chemistry*, 40(8), 1605.

Hildebrand, R. (1993). Drug abuse reduction in the military service. DOT Meeting on Random Drug Testing, Washington D.C., February.

Mikkelsen, S. and Ash, K. (1988). Adulterants causing false negatives in illicit drug testing. *Clinical Chemistry*, 34, 2333-2336.

Moore, C., Barnes, V., Hauth, J., Wilson, R., Fawcett-Long, J., Toquam, J., Baker, K., Wieringa, D., Olson, J. (1989). NUREG/CR-5227, Supplement 1, *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues*.

*MRO Alert*. (1993). "Adulteration update: UrinAid! Product recalled and discontinued." 4(9), 10.

MRO Newsletter. (1993). NCLP dictates perception of adulteration testing. 2(5), 6-8.

MRO Newsletter (1995). Urine Adulterants. 4(6), 5.

Nightbyrd, J. (1991). *Conquering the Urine Tests: A Complete Guide to Success in Urine Drug Testing*. Austin, TX: Byrd Laboratories.

Pearson, S., Ash, K., and Urry, F. (1989). Mechanism of false-negative urine cannabinoid screens by Visine eyedrops. *Clinical Chemistry*, 35, 636-638.

Person, N. and Ehrenkranz, J. (1987) Detecting urine sample adulteration in drug testing programs. American Association of Clinical Chemistry Annual Meeting, Albany, NY.

Person, N. and Ehrenkranz, J. (1989). Evaluation of urine temperature measurement methods to screen urine specimens for drug testing. American Association of Clinical Chemistry Annual Meeting, Atlanta, GA.

## Subversion and Countersubversion

Wagener, R. E., Linder, M.W., and Valdes, R. (1994). Decreased Signal in Emit Assays of Drugs of Abuse in Urine After Ingestion of Aspirin: Potential for False-Negative Test Results. *Journal of Clinical Chemistry*. 40 (4), 608-612.

Westra, C., Forslund, C., Field, I., Gutierrez, J., Durbin, N., Grant, T., Moffitt, R. (1994). NUREG/CR-5758, Vol. 4, *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports, CY 1993*.

## CHAPTER 7: FITNESS FOR DUTY: RETURN TO WORK

### The Implications of Drug Type, Types of Users, Treatment Methods, Treatment Outcomes, and Return-to-Work Policies

Carlyn Orians and Susan Kinne<sup>1</sup>

#### 7.1: Introduction

Concern about substance abuse and its effects on the workplace has prompted many employers to adopt substance abuse policies and programs. Programs were also adopted because of Department of Transportation (DOT) regulations and the Drug Free Workplace Act. Others instituted programs to take advantage of the initiatives that several states provided. The principal objective of these programs is to achieve a drug-free work environment. The development of stringent policies has been especially evident in industries where employees are performing tasks that directly affect public safety. The U.S. Nuclear Regulatory Commission (NRC), for example, through its fitness-for-duty (FFD) rule (10 CFR Part 26), requires all licensees authorized to construct or operate a nuclear power reactor and all licensees authorized to possess or transport Category I nuclear materials to establish and maintain FFD programs. These programs essentially cover all employees who have unescorted access to protected areas or to Category I material for which concerns about safety are paramount.

Following a confirmed violation of the worksite policy on substance abuse, several organizational responses are possible. Some employers choose to terminate the offending employee. Others provide the opportunity for reinstatement. Employers providing the opportunity for reinstatement are faced with major decisions regarding the treatment, rehabilitation, and return to work of the affected employee. The need for continual, long-term monitoring of reinstated employees presents additional issues. These reinstatement-related decisions require that the needs and rights of employees are balanced against the need of management to ensure that employees are fit to perform the job duties required of them when and if they are permitted to return to work.

There is little in the way of professional consensus to guide those who must make these decisions.

- The success rates for methods of treating drug and alcohol abuse are not well established.
- There is no widely accepted definition of abuse or recovery from addiction and abuse.
- There are no widely accepted criteria for measuring employee fitness for return to duty.

These issues are further complicated because most research on treatment and treatment outcomes focuses on drug or alcohol dependent individuals who are not in the workforce and gives little attention to the users who are identified only through a positive test for drugs or alcohol at work. Since definitions of dependency are based in large part on behavioral and social symptoms of drug and alcohol abuse, symptoms that may include difficulty maintaining steady employment, there is limited application of these research findings to the population of currently employed workers who test positive. In a sort of Catch-22, a client who is able to return to steady employment may be viewed by the treatment program personnel as having a successful treatment outcome. This does not provide any guidance to the employer faced with the decision of whether or not to permit an employee to return to work in the first place.

This chapter reviews the literature on types of drug and alcohol users, treatment methods, and treatment outcomes, as well as the scant literature on return-to-work issues, recognizing the limitations discussed above. Despite questions about the applicability of this literature to the workforce population of interest here, it is the only research on return-to-work issues available. The literature does provide evidence indicating that treatment outcomes and subsequent ability to reliably perform in the workplace differ by characteristics of the user and the context of use (including use patterns, employment, and social stability). These findings provide insights about how types of drug and alcohol users, treatment methods, and outcomes and return-to-work policies might apply to the population of interest.

The chapter concludes with a discussion of the important elements of an assessment of fitness to return to work. In light of the difficulty in defining suitable outcomes, many employers are choosing to focus efforts on developing procedures that must be followed in treatment, rehabilitation, and the decision to permit reentry into the workplace. The procedures typically specify the decisionmakers and address some of the criteria, but appropriately leave the ultimate determination of fitness to professional judgment.

This discussion of return-to-work issues is intended to assist the reader in understanding the knowledge base upon which these decisions are made. It is not intended to suggest that professional judgment in selecting

treatment and determining fitness to return to work be replaced.

### 7.2: Categories of Users

The distinction between various categories of users is particularly important in a workplace setting. Workplaces with pre-employment screening will have eliminated most physiologically dependent users prior to employment. Therefore, the majority of employees identified through drug and alcohol testing programs are expected to be users who do not fit the classic diagnosis of dependence. Many of them might be considered controlled users because they exhibit little or no evidence of physiological dependence, compulsive use patterns, or social dysfunction (i.e., they are holding down a responsible job). However, some contend that any use of an illicit substance in the face of workplace sanctions constitutes evidence of dependence. There is disagreement in the professional literature regarding these definitions (Murphy, 1989).

The lack of attention paid to the controlled user in the research literature reflects a widely-held assumption that controlled use of illegal drugs is not possible. Some observers hold that all use inexorably leads to abuse and addiction. This view is in accord with the classic disease model of drug and alcohol abuse. In contrast, others believe that controlled use is possible, with long-term or moderate use patterns capable of being sustained for years without leading ultimately to heavy or compulsive use (Murphy, 1989).

Different patterns of use are well-differentiated for some drugs but not others. For example, the distinction between social drinking and alcoholism is readily made. Controlled use of marijuana is also believed to exist but not controlled use of psychedelics or opiates (Zinberg 1984, p. 241). Assumptions about the addictive qualities of psychedelics or opiates militate against the idea that control of these drugs is possible. Even when different use patterns are acknowledged, distinctions have been vague, imprecise, and impossible to compare across substances (Zinberg, 1984).

Empirically it has been difficult to dispute these beliefs because controlled drug users are hard to identify and study. There is some evidence that occasional or nonaddicted opiate users, in and out of treatment, equal or outnumber opiate addicts (Zinberg, 1984). A few

longitudinal studies of small populations have documented the stability of controlled cocaine use over periods as long as eleven years (Murphy, 1989). A follow-up study of Army-enlisted men upon their return from Vietnam revealed a widespread ability of former drug users (many of them addicted while in Vietnam) to use narcotics occasionally upon their return without becoming addicted (Robins, 1974). (Also see Chapter 3 for a discussion of drug use prevalence.)

Given the variety of terms (e.g., controlled, casual, nonaddicted, recreational, and noncompulsive users) currently used to describe users that do not fit the classic dependency model, and none of which accurately describe the variety of users that are detected through workplace testing, we have coined a new term "functional abusers" to describe the population of interest here. Functional abusers refers to the population of workers who, although able to hold down a responsible job, have violated company policy by abusing drugs or alcohol in the workplace. It should be noted that the term does not imply that such abusers function properly or to the best of their ability, nor does it indicate whether the person is impaired on the job.

### 7.3: Treatment

#### 7.3.1: Treatment Approaches

The term "treatment" must be used carefully, for it takes different meanings when applied to users of different substances and to differing intensities of substance use. Drug or alcohol abuse treatment is a term that is normally used to refer to services aimed at the drug or alcohol dependent individual, not necessarily the user identified through workplace testing. The relevant literature contains little professional discussion of appropriate treatment for those people who do not fit the classic diagnosis of dependence (e.g., those who we are referring to as functional abusers). Therefore, the following discussion is necessarily focused on treatment for dependent users.

Approaches to drug and alcohol abuse treatment for the dependent user are varied, depending in part upon the different models or theories of drug dependency that guide their development. These theories fall generally into three classes: medical (or physical), psychological, and social. Most treatment programs are based on one of

these models or on a hybrid of them and most of them share a common conceptualization of drug and alcohol dependency as an illness. In contrast to a "moral" model which views abuse as a moral failing, an "illness" model suggests that there is a condition that can be fixed or ameliorated (if viewed as a chronic illness) through treatment. The models vary as to which disease process is emphasized.

One outcome of the dominance of the illness model is that the question of the trustworthiness and reliability of a person who has a substance abuse problem is never addressed in the treatment literature. Trustworthiness and reliability are not concepts that fit within an illness model. Even those practitioners who question the illness model (Cocozzelli, 1989; Fingarette, 1988) have not entered into a discussion of trustworthiness and reliability. Yet, establishing reliability is a real concern as an outcome of treatment for the group we are calling functional abusers. From the employer's point of view, establishing the reliability of the employee who has violated the workplace sanctions on drug and alcohol use may in fact be the overriding issue in the rehabilitation and return-to-work decision. In practice, reliability is sometimes approached as an issue of compliance with a treatment plan and a psychological assessment of openness and honesty (D. Lowe, Hanford Environmental Health Foundation, personal communication, 1993). This idea is taken up further later in the chapter (see Sections 7.5.1 and 7.6).

The following discussion of the different models of drug dependency applies to both drugs and alcohol unless specified otherwise.

### 7.3.1.1: Medical or Physical Models

According to a medical model of dependency, people who are physically dependent are presumed to have a genetically transmitted biochemical abnormality that predisposes them to abuse of a drug if they use it (genetic theories). Alternatively, people are presumed to experience a metabolic alteration in response to exposure to narcotics (metabolic theories) (Wesson, 1986). Because not everyone has this biochemical abnormality, the genetic theory explains why not everyone who drinks heavily develops alcoholism. Consequently, genetic theories are used to bolster the illness model of drug and alcohol dependence by explaining why some drug and alcohol abusers cannot return to "controlled" use. Metabolic theories have been used primarily to explain opiate (heroin) addiction, supplying support for opiate dependency as a disease or illness. As a result, metabolic

theories provide a medical foundation for methadone maintenance therapy.

### 7.3.1.2: Psychological Models

According to some psychological models of drug dependency, the cause of drug abuse lies in the person's psyche, not in his or her genetic heritage or metabolic functioning. Therefore, treatment that ignores the underlying psychopathology that leads to drug abuse cannot succeed in the long term. Furthermore, according to some theories, proper treatment of the underlying psychopathology should eventually result in the cessation of the drug abusing behavior, even in the absence of any treatment directly targeting drug use. Drug abuse behavior is viewed as a symptom and not a cause. Treatment approaches vary from those that address only the underlying psychopathology to those that first treat the addiction and then treat the psychopathology once abstinence has been achieved (Wesson, 1986).

Psychological models are also used to explain why formerly addicted persons who are no longer physically dependent upon a drug often return to abuse patterns. The theory is that they may physically respond to the presence of social stimuli formerly associated with drug seeking and withdrawal behavior. Aversion treatment is designed to counteract this conditioned response through the administration of nausea causing substances while subjecting the person to the relevant physical or social stimuli—smell, taste, visual settings, etc.—thereby linking these stimuli in their minds to the sensation of nausea. Repeated administrations of particular stimuli are intended to result in an aversion to (as opposed to a desire for) the drug (Wesson, 1986).

Social learning theories acknowledge the role of classical psychological models, but focus on those cognitive processes that occur between exposure to the stimulus and the drug using behavior. Treatment based on social learning theories focuses on teaching coping skills, skills that will help avoid using drugs when faced with a high risk situation (Wesson, 1986).

### 7.3.1.3: Social Models

There is a growing focus in the research literature on social and environmental factors related to use and recovery (Cocozzelli, 1989). This relationship is best developed in alcohol abuse research, where epidemiology studies have found that a person's environment may either facilitate or obstruct alcohol consumption. Post-treatment studies have similarly found that stable and

## Return to Work

unstressful environments are associated with reductions in both alcohol consumption and alcohol-related problems, and that poor outcomes are associated with environments having few family and work resources and little support. Another line of research has focused on those people who undergo recovery in the absence of any formal treatment program. This natural recovery process has been related to positive changes in a person's social milieu, such as a marriage or birth; positive vocational satisfaction; changes in health (for the worse); and previous success at self-control (Cocozzelli, 1989). Social networks are also being examined to explain both initiation of drug and alcohol use and the recovery process.

Social models, with their emphasis on environmental factors, present a challenge to the illness model conception of alcoholism and drug dependence. Many treatment programs currently accept uncritically the veracity of the illness model, despite growing empirical evidence that cannot be explained using this model (Cocozzelli, 1989, p. 546; Fingarette, 1988). The growing evidence suggests that the "traditional intake-treatment-discharge conception needs to be expanded to take into consideration extra treatment factors" relating to the patient's environment (Cocozzelli, 1989, p. 546).

The rare studies of controlled or nondependent drug users indicate that user norms and informal social controls support stable nonaddictive drug use (Murphy, 1989). Contextual factors such as friendships with other users and structured use with friends do better than demographic or psychological variables at explaining stability of use (Kaplan, 1992; Carlin, 1977). Even among those who are drug-dependent, there is clear evidence that social norms set boundaries on amount and conditions of drug use (Zinberg, 1984).

As the social environment is an important influence on substance use, it also becomes a focus of treatment efforts. More specifically, social models emphasize the importance of the post-treatment environment in achieving treatment goals. This approach suggests that treatment plans should incorporate assessments of a person's family and work environments. Strategies to reduce stress and develop coping skills can become part of the treatment plan.

### 7.3.2: Treatment Goals

Treatment programs vary in the goals that are set for their patients. In a nationwide survey of outpatient drug and alcohol treatment programs, complete abstinence, stable

social relationships, positive physical health, positive emotional well-being, and steady employment were identified as the most important treatment goals (Price, 1991). These goals target the physical, psychological, and social manifestations of drug use and dependence. Not all treatment programs, however, stress abstinence as a goal. Some focus on controlled use in the belief that abstinence is not realistic given the client population that is served and that improvement over the pre-treatment condition is preferable to treatment failure. Controlled use as a goal is typically found in programs that treat clients with high levels of drug use and dependency such as methadone programs.

Here, as in all the treatment literature, appropriate goals for functional abusers are not addressed. They are simply not part of the mainstream treatment population. From the perspective of the employer, however, it is imperative that the employee be unimpaired at work. To achieve this goal, abstinence from alcohol and legal drugs that could impair performance during work and specified pre-work periods is typically prescribed in a substance abuse policy. This is equivalent to the construction and observance of a new norm governing appropriate use. In a workplace setting where the employer has a zero tolerance for use of illegal drugs, as in the nuclear industry, appropriate behavior with regard to these drugs is complete abstinence.

### 7.3.3: Treatment Selection Process

Selection of a treatment program results from a diagnostic phase which confirms the employee's substance abuse and identifies the level and type of substance use, and the psychological, medical, and environmental factors that are deemed important by the treatment model. No single treatment program is appropriate for all substance users or abusers. Diagnosis is performed by substance abuse treatment professionals on a case-by-case basis and results in a "personalized" treatment plan where the best match between the prospective patient and program is sought. The "best match" takes into account both the characteristics of the employee and the treatment program (intensity, philosophy, goals, type of patient it accepts). In addition to the development of a treatment program, the diagnostic phase often includes the development of plans for follow-up and future employment. Both of these elements are required by the NRC's fitness-for-duty rule.

Selecting a treatment program also demands consideration of the needs of the employer. For example, treatment goals must be compatible with the employer's expectation

regarding fitness to return to work. In the treatment of safety-sensitive personnel, it is likely that a goal of reducing a drug abuse problem would not be sufficient; a goal of abstinence from illicit drugs should be sought in these circumstances. In the case of the abuser of alcohol or of legal drugs, identification of a compatible treatment goal is not as readily apparent. In these cases, the appropriate goal may depend upon the pattern of use and professional judgment about likely future use patterns and the effect on the workplace.

#### 7.3.3.1: Diagnosis

A variety of diagnostic techniques are available to assess levels and patterns of drug and alcohol use. Diagnostic tools include a battery of psychological and medical tests that can be administered through interviews, pen and pencil exercises, and/or physical examinations. For alcohol abuse and alcohol dependence alone there are at least 15 recognized methods, techniques, and approaches for assessment and diagnosis (Herrington, 1987). However, establishing the use of alcohol or illegal drugs does not signify dependency. For this reason, the American Medical Association (AMA) recommends specific medical diagnosis for each patient. Furthermore, the AMA's 1972 treatment guidelines state that "mere" use of an opiate "even if periodic or intermittent . . . cannot be equated with drug dependency" (reported in Pickens, 1991). Thus, for opiates, as well as for other drugs and alcohol, a complex, individualized diagnosis is necessary to establish dependence. One widely-used systematic method for diagnosing abuse was introduced by the American Psychiatric Association (APA). The APA's Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R) provides a consistent and uniform meaning of abuse and dependence that is often used to distinguish between the dependent and casual user of drugs or alcohol. According to DSM-III-R criteria, dependence is a syndrome that consists of multiple symptoms, including an inability to control drug use, impairment of normal functioning, and physiological evidence of chronic drug use (reported in Pickens, 1991). While substance abuse is considered a disease according to this view, most of the symptoms are behavioral rather than physical.

A variety of screening tests have been developed to assess the severity of drug and alcohol abuse based on these criteria or its precursors (e.g., DSM-III), including the Schedule for Affective Disorders and Schizophrenia (SADS), the Diagnostic Interview Schedule (DIS), the Structured Clinical Interview for DSM-III-R (SCID), the Drug Abuse Screening Test (DAST), and the Minnesota

Multiphasic Personality Inventory (MMPI) (Hasin, 1991). (See Sweetland, 1991 for a catalog of available tests.) Some of the diagnostic instruments rely on interviews, but many of them are paper and pencil self-assessments.

Common methods for diagnosing alcoholism include the Alcohol Use Disorders and Associated Disabilities Interview Schedule (AUDADIS), the Michigan Alcoholism Screening Test (MAST), which is a 25-item questionnaire to which respondents answer yes or no, and an approach developed by the Criteria Committee of the National Council on Alcoholism (CRIT), and its modified version (MODCRIT), that delineates major and minor signs of alcoholism that are readily identifiable through interviews and observations. MAST, CRIT, MODCRIT, and DSM-III-R are consistent in viewing a core set of symptomatic behaviors as representative of alcohol abuse syndrome (Herrington, 1987).

However, with respect to alcoholism in particular, there is a growing movement towards a conception of alcohol dependence as a condition with a great deal of variation in degree of dependence and disability (Cocozzelli, 1989). According to the DSM-III-R classification, a person is diagnosed as either having, or not having, one of several categories of alcohol-use disorders. Under the new conception of alcohol dependence as a state that varies along some continuum, a person is seen as exhibiting some level of drinking behavior along multiple dimensions: salience of drink-seeking behavior, increased tolerance to alcohol, repeated withdrawal symptoms, repeated relief or avoidance of withdrawal symptoms by further drinking (i.e., development of a stereotypical drinking pattern to avoid withdrawal symptoms), subjective awareness of a compulsion to drink, and reinstatement of the syndrome after abstinence (Cocozzelli, 1989). Diagnostic procedures, therefore, are beginning to address the range of physical, psychological, and social manifestations of alcohol use that can occur rather than simply diagnosing the presence or absence of an alcohol abuse syndrome.

Personal interviews, both structured and unstructured, can provide some of the information needed to make a diagnosis and select appropriate treatment. Interviews can produce useful information about characteristics such as level and type of substance use, history of use, patterns of use (e.g., drinking alone), preoccupation with drugs and alcohol, and related problems (e.g., stress, fatigue, financial concerns, divorce). They can also be used to assess the motivation of the employee to comply with treatment plans. This information can in turn be used to shape appropriate treatment and determine subsequent

## Return to Work

treatment outcome. Interviews can yield important information that a physical (medical) examination or a paper and pencil self-assessment may fail to uncover. Drawbacks to personal interviews concern the accuracy of information that is provided and hence the validity of the subsequent diagnoses. Employees may feel that it is in their best interest to conceal information or to minimize their substance abuse problem, or they may themselves be in denial about the extent of the problem (Cohen, 1986).

Paper and pencil self-assessments have the advantage of being less dependent on the skill of the interviewer and therefore are considered more reliable in making a diagnosis (D. Lowe, Hanford Environmental Health Foundation, personal communication, 1993). Information obtained through written instruments can also assess the level and type of substance use, history of use, patterns of use, and preoccupation with drugs and alcohol. They can also assess the subject's honesty and openness as an indicator of reliability.

Medical examinations can provide physical evidence of drug and alcohol abuse such as skin tracks, infection, or a shooting tattoo. Indirect signs may include jaundice, contact dermatitis, or dental disorders (Cohen, 1986). A physical examination can also help to determine whether a worker is experiencing other medical problems, resulting from or independent of the substance use, that could interfere with treatment or fitness for duty. In view of the various advantages and disadvantages associated with the different diagnostic methods, utilizing a combination of techniques may be the most accurate way to assess the severity of drug and alcohol abuse.

### 7.3.3.2: Limitations of Patient Matching

Selecting the best match between the prospective patient and program necessarily takes into account the characteristics of the work setting, the employee, and the treatment program. Political and financial pressures have contributed to an increase in efforts to match patients with treatments once a diagnosis has been completed in the belief that these efforts will contribute to better or more cost efficient outcomes. While conceptually attractive, the ability to effectively match patient to program has been methodologically elusive (McLellan 1991). To date there have been very few studies within the drug abuse field that have attempted to evaluate the success of these efforts. However, there is a growing body of empirical work that attempts to assess the success of matching patients to treatments in the field of alcohol abuse (Mattson et al., 1994a). These studies indicate that much work remains to be done to recommend appropriate

matching strategies. However, a few preliminary findings are worthy of mention as they apply to the context of workplace substance abuse policies.

The first conclusion from the few studies that have been completed is that patient factors have been more predictive of outcome from treatments than have treatment process factors. Mattson et al. (1994a) conducted a chronological review of empirical studies matching alcoholic clients to treatment (1970-1993) and found that although 31 studies provided positive empirical evidence for patient treatment matching, there was considerable variation in the types of treatment studied. In other words, because treatment programs vary in their intensity (inpatient/outpatient), philosophy, goals, and type of patient they accept, outcome studies to date have failed to clearly identify which factors lead to better outcomes for which patients. In spite of this variability in treatment process, it is generally acknowledged that the majority of programs offer few choices and relatively standard procedures (Mattson et al. 1994b). However, often treatment interventions are not clearly operationalized and monitoring of treatment is not systematic (Mattson et al. 1994a). This makes comparison across treatments difficult. In general, better measurement techniques are available for patient characteristics than for the treatment process. Other methodological measurement issues that hinder the evaluation of the treatment processes include questions as to how to reliably measure such treatment program characteristics as staff turnover, staff qualifications, and morale, and how to compare treatment programs in diverse environments (Ball, 1994).

Of those patient characteristics that have been studied, the ones that have been most useful in accurately predicting successful outcomes have been social, economic, and psychiatric factors. People with relatively good social and economic supports and fewer psychiatric problems do well in most treatments, independent of treatment method. Severity of drug dependence, family history of substance abuse, and presence of antisocial personality disorder have been generally predictive of poorer outcomes from all treatments. Those with more serious psychiatric problems also do less well generally, but have done most poorly in outpatient settings (McLellan, 1991). These results provide support for both psychological and social models of drug dependence.

Thus, it has been far easier to identify patient factors that predict successful outcomes. There is no discussion in the literature of the best treatment for functional abusers. In fact, there is still widespread disagreement among both practitioners and research investigators about specific

types of treatment which are effective for various drug abuse patients (Ball, 1994). However, those who study controlled users point to differences in the setting of use as a partial explanation of differences in ability to control use, and emphasize the importance of norms governing substance use in explaining why controlled users maintain their control over time. This suggests that rather than focusing on deterrence methods, "treatment" for controlled users might involve eliminating settings the user regards as appropriate for use. It may also mean that controlled use of those substances for which zero tolerance is not a workplace requirement (i.e., alcohol and legal drugs) could be a reasonable goal for these users.

At this time there are no clear predictors of differential outcomes from different treatment methods for different patients, and thus there are no consistent guidelines regarding the best match of patient with treatment program. However, this is an area of expanding research. Developments in patient treatment matching warrant attention over the next few years as researchers and practitioners seek a better understanding of factors that lead to successful outcomes.

## 7.4: Treatment Outcomes

Studies of drug and alcohol treatment outcomes show mixed results in achieving abstinence for sustained periods of time. This remains true even when treatment programs are disaggregated into their component parts and outcomes are examined against treatment philosophies, length of treatment, treatment intensities, and the content of treatment programs. Patient factors, including motivation, chronicity and severity of drug or alcohol abuse, age, gender, and personal and social resources, appear to be better predictors of successful outcomes (Collins, 1993). This suggests that functional abusers, who are by definition employed, would have better outcomes than are typically reported in the treatment literature provided that they are sufficiently motivated to abstain. Evidence from the nuclear industry suggests that a high-paying job does provide a strong motivation to comply with treatment (D. Lowe, Hanford Environmental Health Foundation, personal communication, 1993).

The absence of generally accepted criteria for success and failure have added to the difficulty of comparing the results of different studies. Definitions of rehabilitation, relapse, and recovery vary widely as do measures of treatment outcomes. This has contributed to the wide variations in reported recovery and relapse rates.

Consistent with the treatment literature, these definitions apply to the dependent user, not to the controlled user.

This section presents some common definitions of rehabilitation, recovery, and relapse (7.4.1), and common outcome measures (7.4.2), followed by a summary of treatment outcomes from a variety of studies (7.4.3). The section concludes with a discussion of the relevance of these studies to the decision to permit workplace reentry following treatment (7.4.4).

### 7.4.1: Definitions of Rehabilitation, Recovery, and Relapse

Rehabilitation is sometimes defined as the process of returning a person to a prior level of functioning (Smith, 1984). Used in this way, it generally refers to return to employment, although it can also include return to health or to economic or psychosocial functioning. As such, it may be a goal of medical treatment, family therapy, physical therapy, and education (Smith, 1984).

In normal usage, recovery is used in this same way, to mean "the act of regaining or returning toward a normal or usual state" (Webster, 1971, reported in Maddux, 1986, p. 60). By this definition, recovery is also a process, not a state, as in the use of the term "recovering alcoholics" by Alcoholics Anonymous (AA). In contrast, others define the term recovery to refer to a person's state of being: acceptance of the inability to control the use of alcohol, acceptance of alcoholism as a disease, and the maintenance of sobriety through active avoidance of alcohol by using peer support and adherence to the 12-step program (Smith, 1984). The 12-step program which is the foundation of AA and other self-help approaches, is a philosophy and behavior code that stresses spiritual revitalization (Collins, 1993; Chappel, 1993). According to this definition, recovery is a necessary step for rehabilitation but is not synonymous with rehabilitation.

The American Medical Association (AMA) in 1970 developed criteria for recovery from drug dependence that sought to establish what would normally be thought of as a "cure." These criteria included three years of abstinence from the primary drug and no abuse of other substances (Maddux, 1986, p. 60). Another researcher (Bejerot) recommended 5 years of abstinence as a working definition of recovery (reported in Maddux, 1986, p. 60). This points out the lack of consensus on what recovery is and how it can be measured. It also highlights the fact that there is disagreement about whether drug and alcohol dependence can be "cured."

## Return to Work

Relapse is defined as a resumption of substance abuse following a period of abstinence. While conceptually straightforward, in practical use it is measured or defined in many different ways. A single use after a period of abstinence may qualify as a relapse according to some working definitions, while only a return to baseline use would qualify under other definitions. There is also no consensus on how long one has to be drug free in order for relapse to occur. Some treatment professionals view relapse as a continuum, consistent with a theoretical view of amelioration rather than cure as the goal of drug treatment for a chronic disease.

### 7.4.2: Outcome Measures

Treatment outcomes are measured differently according to the goals established by the treatment program. A commonly used outcome measure is the percent of clients who meet the particular treatment goals. While these goals typically include abstinence from drug or alcohol use or limited use (e.g., average number of days using narcotics, or complete abstinence both upon completion of the program and at specified intervals following treatment), other measures of social functioning, psychological functioning, or physiological functioning are also frequently used to evaluate outcomes. Additional outcome measures (both positive and negative) include physical illnesses, comparisons of pre- and post-treatment employment levels, family disruptions, incarceration rates, and psychiatric symptomatology (Cocozzelli, 1989; Tims, 1991).

As mentioned previously, however, discussions of treatment and treatment outcomes focus largely on the drug or alcohol dependent individual. Since definitions of dependency are based largely on behavioral and social symptoms of drug and alcohol use, including difficulty in maintaining steady employment, the application of these research findings to the population of functional abusers is limited.

### 7.4.3: Treatment Outcome Studies

Short-term methods for separating substance abusers from their abused substances and controlling withdrawal symptoms are highly developed. Nonetheless, over the longer term, relapse continues to be the rule and not the exception among substance abusers entering or completing treatment (Rounsaville, 1986). Research suggests that a majority of those leaving a treatment program will relapse at least once, with the odds highest during the first three months following treatment. A comparison of relapse rates for three types of

addiction—alcohol, heroin, and cigarettes—showed the relapse curves to be markedly similar, with about a 60 percent relapse rate during the first three months. Another 15-20 percent relapsed during the next three months, with the relapse rate levelling off in months six to twelve. About 25-30 percent remained abstinent for the full twelve months (Hunt, 1971, reported in Foote, 1988).

There have been several major treatment follow-up studies over the past two decades. One of the earlier widely published studies was from the Drug Abuse Reporting Program (DARP). The DARP research was based on a population of 44,000 clients admitted to drug abuse treatment among 52 programs during 1969 to 1974. The sample consisted largely of opioid addicts who were interviewed at admission, during treatment, and subsequently followed up in the community. Drug use for the follow-up sample declined dramatically pre-treatment to post-treatment and continued to diminish over the next two and three years of post-treatment (Tims, 1991). The DARP studies compared three major modalities—methadone maintenance, therapeutic community, and drug-free outpatient—and each was found to be effective in comparison with pre-treatment conditions. Relapse remained high, however. Two-thirds of the population followed refrained from daily opioid use by year three, but only one-quarter used no drugs. These figures were maintained or slightly improved between year three and year twelve post-admission (Tims, 1991).

The Treatment Outcome Prospective Study (TOPS) used personal interviews to obtain detailed drug use histories from 11,750 clients entering 41 separate treatment programs in 10 cities from 1979 to 1981. A sample (totalling 5,000) of these clients were reinterviewed three months, twelve months, two years, and four years after treatment. Frequency and type of drug use (heroin, other narcotics, cocaine, and other nonnarcotics) were assessed both before and after treatment. In the TOPS study, drug use in the first year after treatment was common. Post-treatment relapse rates differed by drug types. Relapse to daily use (post-treatment) was much greater among heroin users than among users of other narcotics, cocaine, and nonnarcotics (54% to 25% or less, respectively). Frequency of use pre-treatment was also found to be associated with post-treatment relapse. Monthly pre-treatment users of drugs in all drug categories were less likely to relapse than pre-treatment weekly users. This suggests a better outcome for those patients with less frequent or more controlled use. For those who did use drugs in the first year after treatment, the average number

of days to relapse was 55 days among heroin users and 72 days among cocaine users (Hubbard, 1986).

Another study called the Drug Abuse Treatment Outcome Study (DATOS) is currently underway. In this study, follow-up of patients from 50 programs will be undertaken at three months and twelve months post-treatment. This study will emphasize the conceptualization and measurement of impairment, a composite of substance dependence, psychiatric severity, and deficits in social functioning (Tims, 1991). With a focus on the common treatment methods for cocaine, opiates, and polydrug abuse among different socioeconomic groups and impairment levels, the DATOS study will provide a contrast to the earlier studies that focused on opioid (DARP) and multiple drug abuse (TOPS), the common problems of their respective eras. Each of these studies examined the major treatment modalities in place at the time.

Alcohol treatment programs boast success rates as high as 70-80 percent but success varies with the population treated and the treatment and care provided (Walsh, 1987; Herrington, 1987). As with illegal drug abuse, relapse is most common in the first year following treatment. Thus, many practitioners advise that ongoing treatment (aftercare) continue for at least one to two years following initial treatment (Herrington, 1987).

Based on the evidence that patient background (e.g., motivation, good social and economic supports, and fewer psychiatric problems) is more important than treatment factors in predicting outcomes, some researchers have questioned whether treatment for alcoholism is effective at all. They believe that whatever value the treatment does provide resides in the practical advice and personal support it gives, rather than in the program's particular medical or nonmedical techniques (Fingarette, 1988). Although some drinkers do improve during the course of a program, some researchers have questioned whether the improvement is any greater than that observed in the comparable population of drinkers who do not enter treatment. Following an eight-year clinical study, Vaillant found that "there is compelling evidence that the results of our treatment were no better than the natural history of the disease" (reported in Fingarette, 1988, p. 77). Five comparable studies that Vaillant examined confirmed this finding (reported in Fingarette, 1988, p. 77).

In summary, although drug and alcohol treatment has been shown to be effective in lowering immediate use in the post-treatment population, relapse remains high.

Patient characteristics have been shown to have more effect on outcomes than treatment characteristics but an adequate understanding of those factors that predict long-term recovery has remained elusive. There is also evidence that the post-treatment environment and availability of aftercare is important to avoiding abuse and remaining abstinent (Tims, 1991).

#### 7.4.4: Limitations of Outcome Studies

The utility of the outcome studies discussed above to the workplace environment and the decision to permit workplace reentry following treatment is extremely limited. There are a wide range of reported outcomes that are difficult to compare because of variation in outcome measures. But perhaps most importantly, it is difficult to apply the results of studies where the study population is likely to differ markedly from that encountered in a professional workplace environment like that of the commercial nuclear power industry.

Most drug treatment evaluation studies have dealt with the opioid "street addict." To the degree that patient characteristics predict outcomes, the prognosis for those patients who are detected through workplace testing programs may be considerably better than those who form the treatment outcome study population (Walsh, 1987).

Evidence from the limited studies of nontreatment populations provides support for the contention that efforts to curtail substance abuse may be more successful in populations that resemble the workplace population than among stereotypical "street" users. In a study of 470 returning Vietnam enlisted men, Robins et al. (1974) found that almost half had tried either heroin or opiates or both while in Vietnam and one fifth of them used narcotics there with significant regularity to develop some signs of physical or psychological dependence. However, most of the men who used narcotics heavily in Vietnam stopped when they left and had not begun again 8 to 12 months later. Of those who continued use, most reported that they had not become addicted or readdicted. These recidivism rates among Vietnam veterans who were addicted during their service are much lower than those rates reported in studies evaluating recidivism among the general treatment outcome study populations. Furthermore, the majority of Vietnam veterans who used narcotics in Vietnam did not receive treatment. Most long-term studies of civilians have been confined to groups of people who have received treatment. People who seek treatment or who are forced to seek treatment for addiction generally have comparatively higher recidivism rates than were found among the returning

Vietnam veterans. This could be a function of the type of study population (e.g., the typically "general" population as represented by the Vietnam draftees versus the typical "street" user usually studied in treatment outcome evaluations). The Robins et al. (1974) study suggests that the application of treatment policies for narcotic users should not be generally based on the treatment outcome findings of those addicts who are typically studied in the treatment context.

Yet, while this provides added support for the notion that the long-term treatment outcomes of the alcohol and drug abusers who are identified in workplace drug testing programs may be much better, it is still not clear what type of treatment is the most effective for this group. The results suggest that treatment might focus on modifying exposure to social contexts where use has occurred, identifying typical settings for use, and stressing the importance of compliance in retaining the valued job. Results also suggest that providing worksite support to the employee following the positive test may be critical to his or her long-term success in modifying use patterns.

### 7.5: Assessment of Fitness to Return to Work

Despite a lack of knowledge of the efficacy of various treatment methods, and a lack of consensus on whether recovery is a realistic goal and how to measure it, it is still necessary to obtain assurance of a worker's fitness to adequately perform assigned activities before readmitting that employee to the workplace. Reliance on professional judgment along with appropriate provisions for follow-up and continued monitoring of fitness during the months and years following reinstatement is the typical recourse.

#### 7.5.1: Declaration of Fitness

There are no widely accepted indicators of fitness to return to work. Substance abuse professionals who are called upon to make a determination base their judgment on such factors as:

- detoxification from the substance in question,
- the absence of side effects resulting from withdrawal that could affect performance, and
- the satisfactory resolution of any additional problems noted during initial diagnosis (Segal, 1992).

Because of concerns about the potential for bias, it has been recommended that certification of medical assurance be made by a substance abuse professional who is not the employee's family physician (Segal, 1992).

One workplace drug testing program, that of the Department of Transportation, requires that the Medical Review Officer (MRO) obtain an assessment from the rehabilitation counselor that covers the patient's substance abuse, progress in any rehabilitation effort, and the prognosis and recommendations for aftercare services. Upon reviewing this report, the MRO must feel satisfied that the employee complied with the conditions and requirements of the rehabilitation program in which he or she participated. At that point the MRO is asked to make a recommendation regarding return to duty (Clark, 1990).

Other practitioners stress the importance of psychological assessment in determining the reliability of employees. Tests and interview formats (i.e., Minnesota Multiphasic Personality Inventory [MMPI], motivational interviewing) are available to assess the honesty, openness, and motivation of the employee to comply with treatment plans and goals (D. Lowe, Hanford Environmental Health Foundation, personal communication, 1993; Miller, 1991).

At a minimum, assurance of a worker's fitness to return to work should rest on satisfactory compliance with the chosen treatment program. While the selection of the best treatment program for the employee is based on professional judgment rather than professional consensus, the importance of completing the agreed upon treatment is well established in the outcome literature. Those patients who complete primary treatment and who follow the aftercare plan are consistently those who do best in long-term follow-up studies (Collins, 1993). Compliance with a treatment program and an aftercare plan may provide evidence of employee reliability, although the causal relationship may not be clear (i.e., do patients do better because they complete treatment and remain in aftercare, or are patients who have strong personal motivation and commitment more likely to remain in treatment and aftercare?).

#### 7.5.2: Continued Monitoring of Fitness

Assuring compliance with aftercare programs is an important part of continued monitoring of fitness. For the employer to receive reports on the employee's progress, a return-to-work agreement must be negotiated in conformity with federal regulations that protect employee

privacy in drug and alcohol treatment programs. The regulations apply to all treatment programs which receive any form of federal financial assistance.<sup>2</sup> Patient consent for release of information about treatment must be given in writing, must cover a number of specific issues, and is revocable at any time (Segal, 1990). Only with this consent, and the information gathering it permits, can the employer be confident that treatment and aftercare are proceeding as intended.

Once an employee is certified to return to duty, most employers implement standard follow-up procedures aimed at ensuring that an employee maintains the agreed upon aftercare, and that he or she remains free of drugs as shown by more frequent drug testing. In the nuclear power industry, for example, the fitness-for-duty rule (10 CFR Part 26) requires that workers reinstated to unescorted access following a positive drug test be tested on a regular basis for the next three years at a rate of at least once every three months, and that more frequent testing be conducted during the first four months (at least once a month). Since recovering drug abusers sometimes substitute another drug for the one that they received treatment for (substituting alcohol for cocaine, for example), it has been recommended that follow-up testing include the full range of drugs originally included in the testing program (Segal, 1992). Some advocate an expanded panel of drugs for follow-up and for-cause testing.

Despite a poor ability to successfully predict recovery from drug or alcohol abuse, the importance of environmental conditions and aftercare to maintaining abstinence has been shown. According to social models of drug dependency, these factors are central to successful long-term recovery. Employee Assistance Programs (EAPs) can play an important role in developing a return-to-duty agreement at the time of initial diagnosis, and assisting in the workplace reentry process. They can provide assistance in selecting a treatment program, locating supplemental community resources, monitoring program completion, and facilitating and encouraging aftercare. They can also help to watch for signs of relapse.

Although no studies have examined the effectiveness of EAP follow-up, and thus offer no direct guidance on the role of EAPs in working with employees returning to work, EAPs are in a better position to monitor performance, have better access to feedback from supervisor monitoring, and are more able to provide routine support than are treatment programs. Foote and Erfurt (1988) make a persuasive case for the valuable role

that EAPs can play in aftercare, although they report that, by and large, EAPs have placed their energies on the tasks of case finding, intake, and referral, with little attention being devoted to relapse prevention and aftercare. Research indicates that left to themselves, only a minority of alcoholics have been found to seek aftercare.

Evidence of the importance of worksite support to positive treatment outcomes has come from several studies. Billings and Moos (reported in Foote, 1988) found that work support was especially important for men. Their research on the effect of work settings on outcomes indicates that work settings high in work pressure and supervisor control and that lack autonomy and clarity about role expectations have negative health effects (e.g., increased alcohol consumption). However, settings characterized by high job involvement and supportive relationships with coworkers and supervisors have beneficial effects and diminish the stressful aspects of the employee's job setting. In general, Billings and Rudolf (1983) have found that life stresses (e.g., negative life change events like a death of a friend or economic problems and chronic strains in work and family environments such as time pressure and interpersonal conflict), coping responses (e.g., active, cognitive or positive behavioral responses as opposed to avoidance), and social resources (availability) had influence on recovery as did patients' treatment experiences. Moberg et al. (reported in Foote, 1988) reported that patients from companies with a full alcoholism program did better than those from companies without alcoholism programs, suggesting the possible impact of worksite support. However, the particular aspects of worksite programs that produce the improved effect is not known (Foote, 1988). Research does indicate that coercion to treatment by the employer (e.g., threat of termination) does not appear to have a negative effect on outcome and thus ought to be continued, with protection of individual employee rights (Moberg et al., 1982).

## 7.6: Conclusion

Following a confirmed violation of the worksite policy on substance abuse, several organizational responses are possible. Some employers choose to terminate the offending employee. Others provide the opportunity for reinstatement. Employers providing the opportunity for reinstatement are faced with major decisions regarding treatment, rehabilitation, and return to work of the affected employee. There is little in the way of professional consensus to guide those who must make these decisions.

## Return to Work

The results of this investigation suggest that:

- The minimal level of assurance of a worker's fitness to return to work is denoted by satisfactory compliance with the chosen treatment program. While the selection of the best treatment program for the employee is based on professional judgment rather than professional consensus, the importance of completing the agreed upon treatment is well established in the outcome literature. Those patients who complete primary treatment and who follow the aftercare plan are consistently those who do best in long-term follow-up studies. Compliance with an agreed upon treatment program and aftercare plan is an important element of fitness to return to duty and might be regarded as a measure of employee reliability and commitment to modify behavior.
- Most drug and alcohol treatment outcome studies focus on the highly dependent drug or alcohol abuser, typically the "street addicts." Therefore, the relevance of these findings to the workplace population and the decision to permit workplace reentry following treatment is limited.
- Patient factors, including motivation, chronicity and severity of drug or alcohol abuse, age and gender, and personal and social resources, appear to be better predictors of successful outcomes than treatment program factors. This suggests that employed workers who test positive (functional abusers) would have better outcomes than are typically reported in the treatment literature provided that they are sufficiently motivated to abstain in order to keep a good job. One expert familiar with the nuclear industry has found that a high-paying job does provide a strong motivation to comply with treatment.
- Appropriate treatment for functional abusers is not well studied although, as noted above, research on factors that predict successful treatment outcomes suggests that functional abusers would have better outcomes than are typically reported in the treatment literature, regardless of the type of program.

Developments in patient treatment matching warrant attention over the next few years as researchers and practitioners seek a better understanding of the relationship between patients, treatment methods, and successful outcomes.

- In light of the difficulty in defining suitable treatment and appropriate treatment outcomes for the violator of the worksite policy on substance abuse, many employers are instead choosing to focus efforts on developing procedures that must be followed in treatment, rehabilitation, and the decision to permit reentry into the workplace. The procedures specify the decisionmakers and address some of the criteria, but ultimately leave the determination of fitness to professional judgment.
- Treatment and aftercare plans should be developed with explicit goals that correspond to workplace expectations: abstinence from illegal drugs and abstinence or controlled use of alcohol or legal drugs in a manner that does not interfere with productivity and safety.
- The EAP is in a position to provide ongoing monitoring of fitness upon return to work—necessary to assure compliance with treatment and aftercare goals—and to provide worksite support and encouragement for long-term recovery.
- Discussions of employee reliability and trustworthiness do not appear in the treatment literature because they are not concepts that fit within the prevailing illness model of drug and alcohol dependency.

The discussion of return-to-work issues in this chapter is intended to assist the reader in understanding the knowledge base upon which treatment and return to work decisions are made. It is intended to assist the reader in understanding the currently available literature in this area and to reinforce the need for professional judgment in selecting treatment and determining fitness to return to work.

## Endnotes

1. Carlyn Orians, M.S., Battelle Seattle Research Center  
Susan Kinne, Ph.D, Center for Disability Policy and Research, University of Washington
2. According to Segal, the definition of federal financial assistance is so broad that virtually every reputable drug and/or alcohol treatment program is covered by the regulations (Segal, 1990).

## References

10 CFR Part 26. (January 1, 1990). U.S. Nuclear Regulatory Commission. "Fitness for duty programs." *U.S. Code of Federal Regulations*.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders: DSM-III-R*. Washington, DC.

Ball, J.C. (1994) Why Has it Proved so Difficult to Match Drug Abuse Patients to Appropriate Treatment? *Addiction*, 89, 263-265.

Billings, A.G., Rudolf, H.M. (1983) Psychosocial Processes of Recovery Among Alcoholics and Their Families: Implications for Clinicians and Program Evaluators. *Addictive Behaviors*, 8, 205-218.

Carlin, A.S., Stauss, F.F. (1977). Descriptive and Functional Classifications of Drug Abusers. *Journal of Consulting and Clinical Psychology*, 45(2), 222-227.

Chappel, J.N. (1993). Long-Term Recovery From Alcoholism. *Recent Advances in Addictive Disorders*, 16(1), 177-187

Clark, W.H. (1990). The Medical Review Officer and Workplace Drug Testing. *Journal of Psychoactive Drugs*, 22(4), 435-445.

Cocozzelli, C., Hudson, C.G. (1989). Recent Advances in Alcoholism Diagnosis and Treatment Assessment Research: Implications for Practice. *Social Service Review*, 63(4), 553-552.

Cohen, S., Callahan, J. (1986). *The Diagnosis and Treatment of Drug and Alcohol Abuse*. NY: The Hawthorne Press, Inc.

Collins, G.B. (1993). Contemporary Issues in the Treatment of Alcohol Dependence. *Recent Advances in Addictive Disorders*, 16(1), 33-48.

Fingarette, H., (1988). *Heavy Drinking: The Myth of Alcoholism as a Disease*. CA: University of California Press.

Foote, A., Erfurt, J.C. (1988). Posttreatment Follow-up, Aftercare, and Worksite Reentry of the Recovering Alcoholic Employee. In Galanter, M. (Ed.), *Recent Developments in Alcoholism, Vol.6*. NY: Plenum Press: 193-203.

Hasin, D.S., (1991). Diagnostic Interviews for Assessment. *Alcohol Health and Research World*, 15(4), 293-302.

Herrington, R. (1987). Alcohol Abuse and Alcohol Dependence: Treatment and Rehabilitation. In Herrington, R., Jacobson, G, Benzer, D (Eds.) *Alcohol and Drug Abuse Handbook*. St. Louis, MO: Warren H. Green, Inc. 180-217.

Hoffman, N.G., Halikas, J.A., Mee-Lee, D., Weedman, M.S. (1991). *Patient Placement Criteria for the Treatment of Psychoactive Substance Use Disorders*. Washington, DC: American Society of Addiction Medicine.

Hubbard, R.L., Marsden, M.E. (1986) Relapse to Use of Heroin, Cocaine, and Other Drugs in the First Year After Treatment (NIDA Research Monograph 72). In Tims, F.M., Leukefeld, C.G., (Eds.) *Relapse and Recovery in Drug Abuse*. Rockville, MD: Department of Health and Human Services, National Institute on Drug Abuse, 157-166.

Hunt, W.A., General, W.R. (1973) Relapse Rates After Treatment for Alcoholism. *Journal of Community Psychology*, 1(1), 66-68.

Kaplan, C.D., Bieleman, B., TenHouten, W.D. (1992). Are there 'casual users' of cocaine? *Intraval Ciba Foundation Symposium*, 166: 57-80.

Maddux, J.F., Desmond, D.P. (1986). Relapse and Recovery in Substance Abuse Careers (NIDA Research Monograph 72). In Tims, F.M., Leukefeld, C.G., (Eds.) *Relapse and Recovery in Drug Abuse*. Rockville, MD: Department of Health and Human Services, National Institute on Drug Abuse, 49-71.

Mattson, M.E., Allen, J.P. Longabaugh, R., Nickless, C.J., Connors, G.J., Kadden, R.M. (1994a) A Chronological Review of Empirical Studies Matching Alcoholic Clients to Treatment. *Journal of Studies on Alcohol*, Suppl. No. 12, 16-29.

Mattson, M.E., Donovan, D.M. (1994b) Clinical Applications. The Transition from Research to Practice. *Journal of Studies on Alcohol, Suppl. No. 12*, 163-166.

Miller, W.R., Rollnick, S. (1991). *Motivational Interviewing: Preparing People to Change Addictive Behavior*. NY: The Guilford Press.

McLellan, A.T., Alterman, A.I. (1991). "Patient Treatment Matching: A Conceptual and Methodological Review With Suggestions for Future Research" (NIDA Research Monograph 106). In Pickens, R.W., Leukefeld, D.S.W., Schuster, C.R., (Eds.) *Improving Drug Abuse Treatment*. Rockville, MD: Department of Health and Human Services, National Institute on Drug Abuse, 114-135.

Moberg, P.D., Krause, W.K., Klein, P.E. (1982) Post-treatment Drinking Behavior among Inpatients from an Industrial Alcoholism Program. *The International Journal of Addiction*, 17(3), 549-569.

Murphy, S.B., Reinerman, C., Waldorf, D. (1989). An 11-year Follow-up of a Network of Cocaine Users. *British Journal of Addiction*, 84, 427-436.

Pickens, R.W., Fletcher, B.W. (1991). Overview of Treatment Issues (NIDA Research Monograph 106). In Pickens, R.W., Leukefeld, D.S.W., Schuster, C.R., (Eds.) *Improving Drug Abuse Treatment*. Rockville, MD: Department of Health and Human Services, National Institute on Drug Abuse, 1-19.

Price, R.H., Burke, A.C., D'Aunno, T.A., Klingel, D.M., McCaughrin, W.C., Rafferty, J.A., Vaughn, T.E. (1991). Outpatient Drug Abuse Treatment Services, 1988: Results of a National Survey (NIDA Research Monograph 106). In Pickens, R.W., Leukefeld, D.S.W., Schuster, C.R., (Eds.) *Improving Drug Abuse Treatment*. Rockville, MD: Department of Health and Human Services, National Institute on Drug Abuse, 63-93.

Robins, L.N., Davis, D.H., Goodwin, D.W. (1974). Drug Use by U.S. Army Enlisted Men In Vietnam; a Follow-up on Their Return Home. *American Journal of Epidemiology*, 99(4) 235-49.

Rounsaville, B.J. (1986). Clinical Implications of Relapse Research (NIDA Research Monograph 72). In Tims, F.M., Leukefeld, C.G., (Eds.) *Relapse and Recovery in Drug Abuse*. Rockville, MD: Department of Health and Human Services, National Institute on Drug Abuse, 172-184.

Segal, J.A. (1990). Developing Return-to-Work Agreements. *HR Magazine*, December 1990, 86-89.

Segal, J.A. (1992). Drugs, Alcohol and the ADA. *HR Magazine*, December 1992, Vol. 37, 73-76.

Smith, D.E., Wesson, D.R. (1984). Substance Abuse in Industry: Identification, Intervention, Treatment and Prevention. In Smith, D.C., Wesson, D.R., Zarkin, E.L., Novey, J.H.

*Substance Abuse in the Workplace*. American Society for Industrial Security. Haight Ashbury Publications.

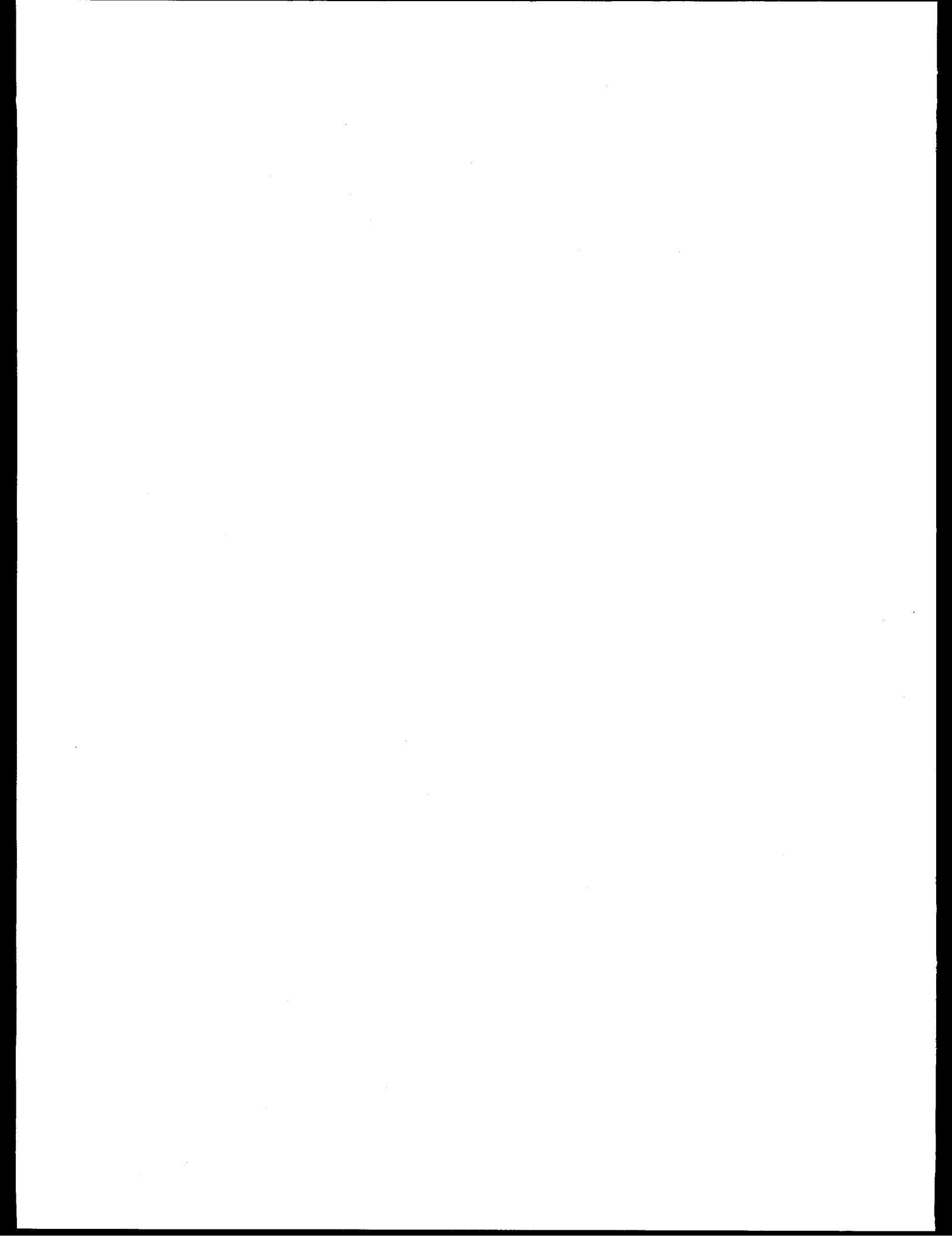
Sweetland, R.C., Keyser, D.J. (Eds.) (1991). *Tests: A Comprehensive Reference for Assessments in Psychology, Education, and Business*. Austin, TX: Pro-ed, Inc.

Tims, F.M., Fletcher, B.W., Hubbard, R.L. (1991). Treatment Outcomes for Drug Abuse Clients (NIDA Research Monograph 106). In Pickens, R.W., Leukefeld, D.S.W., Schuster, C.R., (Eds.) *Improving Drug Abuse Treatment*. Rockville, MD: Department of Health and Human Services, National Institute on Drug Abuse, 93-113.

Walsh, J.M., Yohay, S.C. (1987). *Drug and Alcohol Abuse in the Workplace: A Guide to the Issues*. Washington, DC: National Foundation for the Study of Equal Employment Policy.

Wesson, D.R., Havassy, B.E., Smith, D.E. (1986). Theories of Relapse and Recovery and Their Implications for Drug Abuse Treatment (NIDA Research Monograph 72). In Tims, F.M., Leukefeld, C.G., (Eds.) *Relapse and Recovery in Drug Abuse*. Rockville, MD: Department of Health and Human Services, National Institute on Drug Abuse, 5-19.

Zinberg, N. (1984). *Drug, Set and Setting: The Basis for Controlled Drug and Alcohol Use*. New Haven, CT: Yale University Press.



# CHAPTER 8: DRUG RETENTION OVER TIME

Ronald Wilson<sup>1</sup>

## 8.1: Introduction

Although testing for illicit drug use is clearly a central element of fitness-for-duty (FFD) programs, it is important to understand the limitations of such testing. This chapter discusses limitations regarding testing for marijuana, opiates, amphetamines, cocaine, phencyclidine (PCP), and alcohol. These are the substances for which the Nuclear Regulatory Commission's (NRC's) fitness-for-duty (FFD) regulation requires testing. In particular, this chapter addresses one major limitation: the relatively short time that most drugs remain in the body and, thus, are detectable at or above the standard testing cutoff levels.

## 8.2: General Drug Metabolism Issues

Normal metabolism breaks down drugs and eliminates them relatively quickly. Whether using urine, blood, serum, or breath as the testing medium, drug analysis can identify drug use only if the person has retained enough drug or metabolite in the body at the time the specimen was collected that would be detectable by the particular analytical method being used. Therefore, time is an important consideration in the laboratory detectability of drug use. Cody and Schwarzhoff (1989) state that "The most common cause for a negative result for a drug user is seen when they abstain from drug use for a relatively short period of time before providing a sample" (p. 277). As discussed in detail in this chapter, most drugs, including the substances for which the NRC rule requires testing, are eliminated from the body in three days or less. Marijuana is an exception in that, depending on the frequency of its use, it may be retained for relatively long periods.<sup>2</sup> Frequent use causes marijuana to be stored in fat cells and released slowly over a period of weeks or even months following its last consumption. In contrast, alcohol is the shortest acting drug that is included in the NRC's testing requirements. It is metabolized at a rate of about one drink per hour (DuPont, 1991). The metabolism rates of the other four drugs included in the NRC's testing panel—cocaine, opiates, amphetamines, and PCP—fall somewhere between marijuana and alcohol.

There are numerous physical conditions and behavioral factors that affect how long a drug or metabolite may be retained in the body and, therefore, be potentially detectable through urinalysis. The chemical nature of the drug, i.e., how amenable it is to being metabolized by the body, and the amount and purity of drug taken are important factors. The body type and health of the user,

the amount of fluids consumed, the pH of the urine, and the efficiency of the liver and kidneys in breaking drugs down into their metabolites and eliminating them are additional significant considerations. The frequency and nature of drug use are also important determinants. Frequent, chronic drug use can overcome normal physiological cleansing, thereby causing some drugs or their metabolites to be retained for considerable periods. All these factors complicate attempts at determining definitive, tight time intervals during which particular drugs or their metabolites can be reliably detected by urinalysis.

Polydrug use, i.e., the taking of multiple drugs, further compounds the difficulty of determining how long a particular drug may remain detectable after use. Based upon the results of drug-use screens performed during hospital emergency room visits, polydrug use appears to be fairly common across the U.S. The various combinations of chemicals created by polydrug use cause metabolic and physiological changes in bodily processes. These changes, in turn, may affect the duration of the period that each drug will remain detectable through urinalysis. Because the possible chemical combinations caused by polydrug use are so great, the impact of polydrug use on a given drug's period of detectability is beyond the scope of this chapter. Instead, the action of only one drug on the body at a time is examined. (A separate bibliography on polydrug use has been included at the end of this chapter.)

Because of the relatively short time periods during which marijuana, opiates, cocaine, amphetamines, and PCP are detectable by conventional urine drug testing, and other factors such as the low probability of collecting a specimen during these periods, test results in testing programs across the U.S., including those in NRC licensee FFD programs, do not accurately reflect the drug use by those employees being tested. Experts are of the opinion that current random testing programs identify only a small percentage of the workforce that engages in infrequent drug use. The positive test rate for confirmed drug use in the commercial nuclear power industry has been quite low during the first four years of the FFD program operation. In 1993, for example, this industry's overall confirmed positive test rate was only 0.62 percent (1,512 of the 242,966 tests administered were confirmed positive) (Westra et al., 1994).

The following sections of this chapter examine information regarding the metabolism rates of alcohol and the other five drugs for which testing is conducted. There was some difficulty in interpreting the literature on

## Retention

drug retention because researchers use different methods in performing research and different terminology when reporting data. Unfortunately, the literature discussed in the following sections is not always clear on what analytic methods and quantitative levels to achieve "detection" are being used. For example, one author may define a "positive" drug test result as one that is above a particular drug's "standard" cutoff level while another researcher may use that same term to mean the amount of drug that is above the limit of detection (lowest detectable concentration) of the particular analytical method being used in the research. The same term is used even though there may be a considerable difference between the standard cutoff level and the analytical limit of detection. This lack of consistency in reporting study data means that detection data reported in this chapter are somewhat imprecise and not necessarily comparable in all instances. For purposes of consistency, study results are reported in nanograms per milliliter (ng/mL). The discussion of each drug contains a brief overview of that drug's pharmacodynamics which is intended to provide the reader with a clearer understanding of the mechanisms of drug retention. It is not the purpose of this chapter to delve into the detail of the kinetics, history, or use of the drugs discussed here. Rather, the focus is on how long a typical dose of a particular drug is normally retained in the body at levels detectable by urinalysis using immunoassay and chromatographic and spectrographic testing techniques.

All tests for drug use and retention are based upon urinalysis with the exception of alcohol, for which breath analysis is used to determine blood-alcohol concentration (BAC). Reported study results and conclusions based upon hair, saliva, or other methods are not included as only urine, blood, and breath analysis are, at this time, authorized by the NRC's FFD rule.

### 8.3: Marijuana (Cannabinoids)

The use of cannabinoids predates the fifth century AD (Blum, 1984). Marijuana is the term used when *Cannabis sativa-L* is prepared and smoked or eaten for the purpose of intoxication (Hawks, 1982). Marijuana, used recreationally or addictively, is the most common illicit drug in the U.S. In the 1993 National Household Survey on Drug Abuse, 9% of a sample representing the U.S. population aged 12 and older reported marijuana use in the past year (see Chapter 3). The popularity of marijuana is due in large part to its being relatively inexpensive and easy to obtain, its low incidence of addiction, and its being considered relatively harmless by most users.

The psychoactive ingredient in marijuana is D-9-tetrahydrocannabinol (THC). It is derived from the pistillate Indian hemp plant that grows in warm climates. The amount of active THC in a given marijuana cigarette or oral dose varies widely. The purity and strength of THC has steadily increased since 1975 with a concentration range of from 1 to 28 percent THC by weight of marijuana (Blum, 1984). The advent of sinsemilla (THC obtained from cannabis plants that have not been allowed to pollinate) results in THC concentrations 8 to 10 times higher than that from a normal plant. The Substance Abuse and Mental Health Services Administration (SAMHSA) cutoff levels for this drug are 50 ng/mL for the screen test and 15 ng/mL for the confirmation test. While sensitive testing equipment is able to detect metabolites of marijuana at much lower concentrations, the confirmation cutoff level used by most testing programs is relatively high, eliminating the possibility of positives caused by passive inhalation.

Marijuana is metabolized by the liver and eliminated in the feces and urine. The drug is changed into many different metabolites before elimination and, because it is lipid soluble, can be stored in the fat cells of the body. The amount of drug trapped in fat cells is directly proportional to the amount consumed, the time period during which the drug is consumed, and various body dynamics such as age, percent of body fat, and activity.

The symptoms of marijuana use include dry irritating cough, fatigue and lethargy, dilated pupils, and red eyes (Blum, 1984). The user experiences a mild state of euphoria similar to the effect of alcohol intake. Marijuana changes the way information is processed by the brain and can affect learning, memory, and emotional stability (NIDA, 1980). There is evidence that chronic marijuana use adversely affects fertility for both men and women and may cause birth defects in offspring (NIDA, 1989; NIDA, 1980). Barnett et al. (1985) found a consistent relationship between performance decrement and plasma THC levels. At low doses, marijuana seems to sharpen the senses and decrease inhibitions while at high levels it can generate a state of stupor. Normally, one average dose of marijuana will cause symptoms of use for two to four hours.

Three drug use conditions based upon frequency of use—occasional, moderate, and heavy drug consumption—are described in the following three subsections. The literature uses varying definitions for these three consumption patterns. For purposes of interpretation of the literature in this chapter, *occasional* use is defined as the inhalation of one marijuana cigarette

quarterly, *moderate* use is the inhalation of one marijuana cigarette weekly, and *heavy* use is the inhalation of one or more marijuana cigarettes daily. Other forms of marijuana consumption such as injection and ingestion will not be addressed here. These consumption forms are not common and result in pharmacokinetic patterns different for those produced by inhalation.

### 8.3.1: The Occasional User

The occasional user will excrete most, if not all, of the consumed THC and its metabolites within several hours of consumption. Morgan (1988) found that most occasional users will eliminate virtually all of the drug in four hours. Variability in metabolism and urine concentration, however, may allow detection of marijuana use in drug tests for up to five days. Schwartz and Hawks (1985), for example, state that "Generally, a single smoking session by a casual user of marijuana will result in subsequently collected urine samples being positive for one to five days . . . ." This statement, however, uses "positive" in the sense of the drug being detectable by reliable drug testing techniques rather than in the sense of producing a test result value above a 50 ng/mL cutoff level. A recent study indicates that the time period for detection after an occasional user smokes one marijuana cigarette can be less than two days using a 50 or 100 ng/mL cutoff level (Huestis et al., 1995).

### 8.3.2: The Moderate User

The moderate marijuana user's consumption profile is the most common. This person usually has a pattern of smoking a marijuana cigarette once a week, most often on the weekend. In such a consumption pattern, an immunoassay screen test on Monday might not produce a result above a 50 or 100 ng/mL cutoff level. Although a confirmation test using gas chromatography/mass spectrometry (GC/MS) is normally not performed when an immunoassay screen test proves negative, a GC/MS test of the moderate user's urine sample at the 15 ng/mL cutoff would probably prove positive. A second GC/MS test two to four days later would probably be negative for the average moderate user (Schwartz and Hawks, 1985; Hawks and Chiang, 1986).

### 8.3.3: The Heavy (Chronic) User

It is the chronic user, who smokes one or more marijuana cigarettes per day for several weeks or months, who will have detectable levels of the drug for long periods. Hawks and Chiang (1986), however, report that even drug

use as infrequent as two to three times per week will lead to positives most of the time. The persistence of THC in the fatty tissue of the body is the predominant reason for long detection times. According to Hawks and Chiang (1986), "THC tends to be stored in fatty tissue and can therefore accumulate faster than it can be eliminated in chronic (repetitive) smokers. This accumulation leads to much longer detection times in urinalysis for the chronic marijuana smoker than for the occasional smoker."

Various studies that have found detectable levels of the drug in chronic marijuana users from three to eight weeks after their having ceased drug use support the conclusion that longer detection times are found with heavy users. In one study of six heavy users, it was reported that detectable marijuana could be found for about three weeks after last use (Dakis, 1982). Detectability, in this study, was at the limit of detection rather than the SAMHSA 15 ng/mL of the THC metabolite confirmation cutoff level for marijuana. For 86 heavy marijuana smokers, Ellis et al. (1985) report a mean excretion time of 27 days after discontinuation of use. This was the length of time before a test result was reported as negative. Another study recorded positive drug test results for three weeks in heavy users who had abstained from marijuana use (Sunshine, 1985). A fourth study reports that concentrations of the THC metabolite greater than 20 ng/mL were found by GC/MS in very heavy users up to four weeks after their last drug use (Schwartz and Hawks, 1985). (This study did not report whether marijuana was detected by immunoassay screening at a level above 100 ng/mL.) In still another study involving chronic marijuana users, Morgan (1988) reports evidence of positive test results for incarcerated people for two months after sustained, heavy use.

Other researchers have examined marijuana retention periods by using an estimated half-life for elimination of this drug. Johansson and Halldin (1989), for example, report that the urinary excretion half-life of THC is as long as 9.8 days. In their study, detectable levels of metabolite were recorded four weeks after the last drug consumption. This half-life of 9.8 days reported by Johansson and Halldin is substantially longer than the 3.7 to 7.8 days, with an average of 4.5 days, elimination half-life reported by Baselt (1984). Baselt recorded drug level concentrations as high as 2,700 ng/mL in one chronic user. Using 4.5 days as the average elimination half-life suggests that it could take over 30 days of abstinence to reduce such a high level to below the confirmation test cutoff level of 15 ng/mL. If one assumes a 9.8-day estimation of elimination half-life, however, it would take about 70 days for the 2,700 ng/mL drug concentration to

decrease to below 15 ng/mL. The large difference between these reported study results, about 40 days, demonstrates the wide disagreement among researchers regarding how long a drug or its metabolites are retained by chronic users at concentrations that can be detected above a 15 ng/mL cutoff level.

The results of these studies indicate that a consistent, predictable pattern in the decrease in drug concentration following cessation of heavy marijuana use has not been demonstrated. A plotted curve showing decreasing amounts of detected drug will be different for various combinations of use patterns and body dynamics. Also, such a curve will not necessarily show a symmetrical decline over time but rather may have marked upturns as THC is slowly eliminated through fat cell metabolism and regeneration. Perhaps the most that can be said about the rate of decrease in drug concentrations in the systems of chronic marijuana users was stated by Schwartz and Hawks (1985) who found that "If the individual is a previously heavy chronic smoker who has in fact stopped, test samples may be positive for 3 or more weeks, but the concentrations should show a *generally* decreasing trend . . ."

### 8.3.4: Potential for Successive Positive Drug Test Results for Marijuana

The results of the studies discussed in the previous three subsections appear to have implications for determining whether a person can have two successive positive tests for marijuana without having taken the drug between the two tests. As discussed above, a person's age, body fat ratio, and other physiological factors, along with the intensity and duration of the person's marijuana use, will determine how quickly the drug is eliminated from the person's system. A lean, active, 25-year old will metabolize THC more rapidly than an overweight, sedate, 45-year old. Nevertheless, these studies indicate that the occasional marijuana user will excrete most, if not all, of the consumed THC and its metabolites within a period as short as several hours but no longer than five days after use. The amount of THC in an occasional user's system during most if not all of that period will, however, be at levels below the NRC's current 15 ng/mL confirmation testing cutoff level and probably below cutoff levels used for screening tests. Likewise, for the average moderate marijuana user, a GC/MS test two to four days after the last previous use of marijuana would probably prove negative for the THC metabolite at the 15 ng/mL cutoff level. Based on this information, it can be concluded that it would be very unlikely for an occasional or moderate user to incur a positive result from a GC/MS test using a

cutoff level of 15 ng/mL and spaced more than four days after a previous marijuana positive test unless that person had used the drug in the period between the two tests.

A chronic marijuana user, on the other hand, could plausibly incur a positive result from a second GC/MS test conducted as long as two to three weeks after a previous positive test without having taken the drug between the two tests. In fact, as long as the person's marijuana use up to the time of the first test had been sustained and heavy, the studies cited indicate that the period between two such positive tests during which the person has abstained can be as long as eight weeks. The length of time that a person who is abstaining from use after previous heavy use can still be susceptible to a positive result on a second drug test depends, among other things, upon what nanogram-per-milliliter level of drug detection is considered to be a "positive" test result. The lack of uniformity among the studies cited as to what drug concentrations the researchers considered to be positive indications of drug use, and the difficulty in comparing the findings of those studies, precludes a definitive statement as to the outer limit of a period during which a person's prior heavy marijuana use can be detected by urinalysis. Nonetheless, these studies indicate that the concentration of drugs in the system of a chronic user having a high drug concentration at the time of a first drug test will decline, with few exceptions, at a varying rate for a number of weeks after the initial test. Even if the person abstains from use during that period, he or she may be likely to have a detectable concentration around 15 ng/mL two to three weeks later and may have such a drug concentration at up to eight weeks after the first drug test.

## 8.4: Cocaine

Cocaine is a central nervous system stimulant that causes alertness, decreased appetite, and a feeling of increased energy. It is obtained from the coca bush and produced as a colorless, odorless powder that can be inhaled, swallowed, or injected. It is very soluble in both water and fat, allowing the drug to pass easily into the bloodstream as well as body tissue. Injected and inhaled cocaine is rapidly absorbed into the bloodstream where it has a stimulant effect on certain areas of the brain. The user typically experiences a "rush" described as a euphoric feeling.

The smoking of "free-base" or "crack" cocaine emerged in the late 1970s. The cocaine alkaloid is "freed" from the hydrochloride salt and burned in a pipe-like device. In tolerant crack abusers, as much as 252 mg of cocaine

are absorbed by the lungs (Fischman, 1984). Inhaled cocaine can reach the brain in as little as eight seconds and is relatively unchanged. Once ingested, crack cocaine is effective for only about five minutes. In contrast, taken orally, cocaine takes longer to affect brain cells and the drug is subjected to metabolic action in route (Fischman, 1984).

Although studies show some variability, they clearly indicate the rapidity with which cocaine is changed into its metabolites. Javaid et al. (1978) report that most cocaine users were past their "high" feeling and had returned to normal within 90 minutes of their last use. Fischman (1984) reports the half-life for elimination of cocaine to be about 40 minutes and Jones (1984) reports cocaine's half-life to be from 60 to 80 minutes regardless of how it is absorbed. These figures are estimates and their variability is due to the differences in individual patterns of metabolism, levels of dosage, and other factors.

Cocaine is rapidly metabolized and the pure drug is converted into several metabolites by the liver. It is unusual for a urine test to show cocaine in its pure form. The two most common metabolites are benzoylecgonine (BE) and ecgonine methyl ester (EME). These metabolites are filtered from the blood by the renal system and eliminated from the body via the urine. BE and EME are produced in almost equal amounts but BE can be detected for a longer period than EME because of the longer time required for BE to be excreted via the urine (Aubre, 1985). Detection times vary by person, but EME usually cannot be detected for more than 48 hours following a cocaine dose (Aubre, 1985). The National Center for Forensic Science (1992) reports that the approximate detection time of cocaine is 2 to 4 days. Another study (Baselt and Chang, 1987) also found that urine concentrations of BE could be detected at or above the SAMHSA 300 ng/mL screening cutoff level for 48 hours. (The SAMHSA GC/MS confirmatory test cutoff level for cocaine is 150 ng/mL.) Cone et al. (1988) presented a complex study that indicates the maximum detection time for a GC/MS cocaine metabolite test (at a level of 300 ng/mL) to be 52.9 hours after last exposure to the drug.

In addition to the problems in detecting cocaine use created by the drug's rapid elimination, there have been reports of people claiming that their positive cocaine drug test was due to consumption of a legally obtained beverage. ElSohly et al. (1986) tested a volunteer who drank one cup of "Health Inca Tea," a product imported into the U.S. and claimed to be decocainized. Analysis

showed that the tea bags contained an average of 4.8 mg of cocaine per bag. (The usual dose taken by a cocaine user ranges from 30 to 130 mg.) The subject's urine was collected on several occasions over a period of 29 hours. Using GC/MS, BE was detected at 360 ng/mL in a specimen collected 17 hours after ingestion of the tea. Twenty-two hours after ingestion the level of metabolite had dropped to 274 ng/mL. Thus it was demonstrated that GC/MS analysis can identify BE above the cutoff levels in a Health Inca Tea drinker but, again, the detectable window of time is narrow—about 18 to 20 hours after consumption.

Cone et al. (1995) reported that in studies conducted on passive inhalation of free-base cocaine, the cocaine absorption amounts were insufficient to produce positive urine tests under the SAMHSA screening (300 ng/mL) and confirmatory (150 ng/mL) cutoff levels. However, the researchers conclude that prolonged passive inhalation of cocaine in concentrations exceeding those in the current studies could result in cocaine-positive urine specimens.

Cocaine's popularity and ease of creating dependency, combined with its judgment altering effects, make it a particularly serious threat to the safe operation of complex equipment. Furthermore, the drug's rapid elimination from the body hinders the timely identification of the cocaine user through urinalysis drug testing. Despite the difficulty of detecting cocaine, however, only marijuana generates more positive drug test results than cocaine in the commercial nuclear power industry (Westra et al., 1994).

## 8.5: Alcohol

Alcohol (ethanol) is very soluble in water. It is rapidly absorbed from the digestive tract and enters the bloodstream unchanged. Ethanol affects the body in many ways, both physically and psychologically. Its addictive properties and popularity, together with its psychoactive effects on judgment and its social acceptance, make alcohol abuse the most pervasive and serious drug problem in the U.S. Of the drugs tested by NRC licensee FFD programs, alcohol has ranked third in the number of positive test results after marijuana and cocaine in each of the first four years of program operation (Westra et al., 1994).

The amount of alcohol in a person's blood is expressed as blood alcohol concentration (BAC). The NRC rule requires licensees to conduct alcohol testing using an evidential grade, analytical breathalyzer. The

confirmation test is performed using a second evidential grade breathalyzer. The NRC rule's current BAC cutoff level is 0.04 percent for both the screening and confirmation tests.

The breathalyzer measures BAC in arterial circulation, which is a more accurate indicator of the amount of alcohol affecting the brain than is either venous blood or urine. Using venous blood or urine to determine BAC is not as accurate due to the metabolic activity the alcohol undergoes in the liver and the time between peak level and detection in urine. It is conceivable that a slightly higher BAC could be noted in venous blood or urine for a longer period than through measuring arterial blood. The breathalyzer's ease of use and noninvasive nature combined with the positive correlation between impairment and arterial BAC make the use of a breathalyzer to measure BAC in arterial circulation the testing technique of choice when the subject is able to comply.

Like cocaine, alcohol's metabolism rate is relatively rapid. Alcohol is normally metabolized by the liver at the rate of 0.015 percent by volume or 150 mg/L per hour (Gerchow et al., 1985). In terms of number of drinks, alcohol is metabolized at a rate of about one drink per hour (DuPont, 1991). Normally, one can of beer, one glass of wine, or one drink of distilled spirits contain approximately one ounce of alcohol (Alcohol and Health, 1987). The approximate detection time in the urine for alcohol is three to four hours (National Center for Forensic Science, 1992). Body build, gender, physical activity, and age can dramatically change the rate of both absorption and elimination. Whether the stomach is empty or full before drinking begins and how rapidly alcohol is consumed can further affect how quickly the drug leaves the body. Ekman et al. (1963) administered enough ethanol to eight volunteer subjects to cause a 0.1 percent blood alcohol concentration. This means that each volunteer had enough alcohol in the blood to equal 0.1 percent of the person's total blood volume. A person having this BAC is considered to be "legally drunk" in every U.S. state. Within five hours, virtually all the alcohol the volunteers consumed had been metabolized and eliminated.

A technique known as back-extrapolation provides a means for drug testing programs to compensate somewhat for alcohol's relatively rapid metabolism rate. Through use of back-extrapolation, a person required to submit to a random or for-cause breathalyzer test 6 to 10 hours into a work shift, with a test result between 0.03 or 0.02 percent BAC, could be cited for a higher BAC at the

beginning of the work shift by applying the known average rate of alcohol metabolism to determine what the person's BAC probably was during the earlier part of the work shift. Using this technique, program personnel can determine whether the employee either arrived for work with a BAC over 0.04 percent or had been drinking while on duty.

### 8.6: Opiates

All opiate preparations are derived from the resin of the opium poppy plant. The most common drugs in this class are heroin (a semi-synthetic drug derived from opium), morphine, and codeine. Drugs in the opiate family are potent central nervous system depressants. There are also synthetic opiate-like drugs such as Demerol®, methadone, and fentanyl that have similar kinetic effects on body physiology (Ling and Wesson, 1990). Some over-the-counter (OTC) drug preparations contain dextromethorphan, a drug having codeine-like effects. Opiates can cause moderate to severe cognitive and physical impairment at relatively low doses in the nontolerant user. Heavy users of opiates develop a tolerance for the drugs and must increase intake in order to attain the same level of "high."

Many of the opiate drugs are used in medicine. Demerol®, morphine, and codeine are prescription drugs used for the relief of pain, cough, and diarrhea among other medically accepted applications (Struempler, 1987). Codeine is frequently prescribed by physicians and dentists as an oral medication while morphine and Demerol® are most often used in injection form in hospitals or other health care facilities. Pain and cough suppressants containing codeine can effect cognitive and psychomotor performance and cause impairment. Thus, a person who tests positive for opiates may often have a sound, medically legitimate reason for use of a drug in this class, but may still present a FFD issue that should be addressed.

During the four years of NRC FFD program operation, opiates have accounted for an average of 1.1 percent of each year's total number of confirmed positive drug test results among nuclear industry workers (Westra et al., 1994). Although this percentage of drug test positives for opiate consumption is quite low, FFD program personnel should not underestimate the potential for abuse of this large family of drugs.

All opiates are changed in the liver and excreted in urine as codeine or the metabolite morphine glucuronide. The source of the metabolite cannot be determined through

urinalysis (Paul et al., 1988). Generally, opiate use is effective in the body for about three to six hours (Bureau of Justice Statistics, 1992). Unlike marijuana, opiates are not stored in body cells for long periods. The drugs are metabolized and eliminated relatively rapidly. The National Center for Forensic Science (1992) and Cone and Welch (1989) report that opiate use can be detected through urinalysis for about two days. Fretthold (1990), on the other hand, reports that opiates can be detected above cutoff levels for four days after cessation of use.

The SAMHSA cutoff level for opiates is 300 ng/mL for both the screening and the confirmation tests. On November 16, 1995, DHHS published in the *Federal Register* (60 FR 57587) proposed changes in testing for opiates that would raise the screening test cutoff level for opiates to 2,000 ng/mL and the confirmatory test cutoff levels for morphine to 2,000 ng/mL, codeine to 2,000 ng/mL, and 6-AM to 10 ng/mL. It is intended that these changes in the testing cutoff levels will reduce the number of positive tests that are a result of legitimate uses of opiate-containing pharmaceuticals or the ingestion of poppy seeds.

Positive drug test results for opiates can occur after eating poppy seeds in quantities used in baked goods. The seeds contain forms of codeine and morphine at concentrations high enough to be reported as positive on a urinalysis drug screen. Struempler (1987) reports a codeine concentration of 214 ng/mL three hours after the ingestion of three poppy seed bagels. Morphine was determined to be 2,787 ng/mL. Twenty-two hours after consumption of the bagels, the morphine level was still over twice the 300 ng/mL cutoff level. In another report, Selavka (1991) reported urinalysis results of more than 4,000 ng/mL morphine and more than 2,000 ng/mL codeine after his subjects ate baked goods containing poppy seeds. With one type of Danish pastry the results were quite surprising: 11,571 ng/mL morphine and 4,861 ng/mL codeine. In this study, one subject's urine was still positive after 72 hours. Given this potential for poppy seed consumption to result in positive drug test results, it is extremely important to eliminate poppy seed consumption as an explanation when a person is reported as positive for opiate use (Struempler, 1987).

Because it is understood that there are a number of common opiates that can be prescribed by physicians and that certain foods can cause positive drug test results for opiates, the NRC FFD rule requires that the Medical Review Officer (MRO) perform an analysis in cases where a person has tested positive for opiates.<sup>3</sup>

Although there are a number of common opiates or opiate-like synthetic drugs that can be prescribed by physicians, there are no legal circumstances under which heroin can be prescribed or consumed. Therefore, a specific test that can differentiate heroin from the other opiates is desirable. The analysis for 6-acetylmorphine (6-AM), formerly referred to as 6-monoacetylmorphine (6-MAM), is such a test. Numerous studies, including those of Mulè and Casella (1988) and Cone and Welch (1989), have demonstrated that only a user of heroin will have the 6-AM by-product in the urine. This compound is not found in the urine after morphine, codeine, or poppy seed consumption.

Although the 6-AM test can be used to identify heroin users, in many cases the test may be neither practical nor cost effective due to the very short duration of the presence of 6-AM in the urine after heroin consumption. Fehn and Megges (1985) found that the half-life of 6-AM in whole blood is only 38 minutes. Projecting this very short half-life to what may be detected in the urine, it can be seen how quickly 6-AM will be eliminated from all bodily fluids and tissues. Cone and Welch (1989) found that levels of 6-AM in the urine could not be detected by urinalysis as little as eight hours after test subjects injected heroin. They state that "Overall, it appears that there is a very limited time window for detection of 6-MAM after heroin use." However, 6-AM tests may be effective when screening tests show high levels of morphine. Fortner et al. (1994) report that "... samples containing greater than 5,000 ng/mL (morphine metabolite) should be routinely recommended for 6-MAM analysis as ... there is a 77 percent chance that 6-MAM will be detected." If a person has signs and symptoms of heroin use, it is likely that a urine sample obtained within four to six hours will confirm positive for 6-AM if that substance is present.

In summary, the use of opiates continues to be a substantial FFD concern due to the ability of opiates to cause considerable cognitive and physical impairment even at relatively low doses combined with the widespread use of over-the-counter drug preparations containing opiates. The studies cited above indicate that opiates can be detected by urinalysis for only a relatively short period of two to four days. This rather rapid rate of metabolism makes detection of opiate use through urinalysis difficult. Because the consumption of poppy seeds can lead to positive drug test results for opiates, it is very important that MROs continue to conduct a special inquiry as to the cause of opiate positives. In cases where heroin use may have caused positive drug test results, the use of 6-AM analysis in confirmation

testing may not be practical unless the screening test results show high levels of morphine (i.e., at least 5,000 ng/mL).

### 8.7: Amphetamines

Amphetamines are a class of drugs in the stimulant group. As such, they differ from alcohol, opiates, and marijuana which are depressants. Amphetamines are structurally related to ephedrine and adrenaline. In low doses, amphetamine stimulants increase cortical alertness and attention to detail, i.e., they improve rather than impair performance. They also decrease fatigue and appetite. The effect of an amphetamine on the body usually lasts for about one to four hours (Bureau of Justice Statistics, 1992).

Amphetamines were first synthesized in 1887 (Jacobs and Fehr, 1987). Because of their ready availability, amphetamines have a great appeal for abuse (Holbrook, 1983). In the 1960s and 1970s, preparations containing amphetamines were very popular and physicians often prescribed them for dieters. Long-distance truck drivers used amphetamines to remain alert for 36 hours or more. Methamphetamine, a chemically related drug also known as "speed," has a long history of abuse in the U.S. During the first four years of NRC FFD program operation, amphetamines have accounted for an average of 2.4 percent of each year's total number of confirmed positive drug test results among nuclear industry workers (Westra et al., 1994).

Like the opiates, amphetamines cause the heavy user to become tolerant of their effects, requiring greater doses to achieve the same "high." The chronic user often resorts to injecting the drug as this means of consumption assures more rapid and efficient action (Caldwell, 1981). Even with heavy use, amphetamines will be detected in the urine for only 24 to 48 hours (National Center for Forensic Science, 1992). This rapid elimination rate makes it important for supervisors trained to recognize the signs of drug use to quickly decide whether people suspected of amphetamine use should undergo for-cause drug tests.

The legal use of substances containing amphetamines may produce an initial positive urine screen. These users can, however, be distinguished from illegal amphetamine users through confirmation testing using GC/MS testing. SAMHSA's cutoff levels for amphetamines are 1,000 ng/mL in urine for the screening test and 500 ng/mL for the confirmation test.

Many chemical preparations, including amphetamine and methamphetamine, exist in various chemical forms known as isomers. Isomers are compounds with exactly the same ingredients in exactly the same amounts but with their molecules arranged differently. Isomers are given names representative of some aspect of their diagrammed character such as dextro (*d*) and levo (*l*). Although similar in composition, isomers can have distinctly different pharmacological properties that, in turn, produce radically different effects on the user. The *d* isomer of methamphetamine, for example, is far more psychoactive than methamphetamine's *l* isomer. As such, it is a powerful central nervous system stimulant having high abuse potential. The *l* isomer, on the other hand, has primarily peripheral action and is the active ingredient in the Vicks™ Inhaler (NIDA, 1991).

Illegal drugs containing methamphetamine will often show both the *d* isomer and the *l* isomer in urinalysis. Therefore, the presence of the *l* form of methamphetamine in a urine sample does not by itself rule out illegal drug use (NIDA, 1991). For example, if greater than 80 percent of the total methamphetamine present in a urine is of the *l* type, it is probable that the Vicks® Inhaler, a legal drug, has been used (NIDA, 1991). Only *d*-methamphetamine is detected in the urine of crystal methamphetamine ("ice") smokers, however.

Although normal GC/MS testing methods cannot routinely differentiate between the *d* isomer and the *l* isomer, the *d* and *l* molecules can be technically separated and identified by additional confirmation analysis. Therefore, checking a confirmed positive test for amphetamine or methamphetamine requires a second, specialized GC/MS test to quantify the amounts and kind of isomers present. The National Institute on Drug Abuse recommended in 1991 that, when there is reason to suspect that a positive test result for methamphetamine may be associated with the use of Vicks® Inhaler, MROs should request the certified laboratory to conduct a reconfirmation analysis which differentiates between the *d* and *l* isomers (NIDA, 1991). If the reconfirmation analysis indicates that greater than 80 percent of the total methamphetamine (*d* plus *l*) present in the urine is of the *l* type, it is probable that the Vicks® Inhaler has been used.

In summary, the ready availability of amphetamines provides significant potential for abuse. Like cocaine and opiates, the rapid metabolism rate of amphetamines provides only a short period after use during which urinalysis can detect such use.

## 8.8: Phencyclidine (PCP)

Phencyclidine, known as "PCP" or "angel dust," was introduced in 1957 as an anesthetic. Unwanted side effects including schizophrenia-like psychosis and excitation made PCP unsuitable for anesthetic use except in veterinary medicine. PCP is usually consumed orally or added to leaves and smoked as a cigarette. It is easily absorbed into the bloodstream. The usual PCP dose is 5 to 10 mg and when smoked the effects of PCP reach a maximum in about 30 minutes and gradually decline over 24 hours (Jacobs and Fehr, 1987). At lower dosages (5 to 20 mg), the drug often produces a state of confusion that lasts several hours (Marwah and Pitts, 1986). PCP has become one of the major illicit drugs of use on the street but is mostly confined to the inner-city poor (Crider, 1986).

Any use of PCP can be viewed as potentially dangerous. The unpredictable nature of this drug may cause drastic effects on performance. Short-term job impairment can result from even occasional PCP use and there is a possibility that long-term impairment may result from flashbacks, memory loss, and personality changes (Barnes et al., 1988). PCP can cause agitation, hallucinations, and muscle rigidity (Marwah and Pitts, 1986). At doses greater than 20 mg, PCP can cause serious cardiovascular and brain wave changes. Behavioral dependence on PCP has been reported (Tennant et al., 1981) but physical dependence appears to be rare (Balster, 1986). There is evidence that users generate a tolerance for the drug and must increase their intake in order to obtain the same effect (Marwah and Pitts, 1986).

During the first four years of drug and alcohol testing under the NRC's FFD rule, PCP positive test results have accounted for an average of only 0.4 percent of each year's total confirmed positive drug test results (Westra et al., 1994). The SAMHSA cutoff level for both the screening and confirmation urine tests is 25 ng/mL.

Because PCP is metabolized and eliminated more slowly than alcohol, cocaine, amphetamines, or opiates, a very large dose may be detected via a urine drug test for as long as eight days (National Center for Forensic Science, 1992). Gorelick et al. (1986) reported that measurable levels of PCP remain in the urine for 4 to 5 weeks after last consumption. This duration estimate may be inaccurate, however, because it was based on patient self-reporting of the time and amount of last drug use.

## 8.9: Summary and Conclusions

As illicit drug use and alcohol abuse have increasingly been recognized as threats to workplace safety and efficiency across the country, many employers have initiated FFD programs that incorporate drug testing based upon urinalysis and alcohol testing based upon breath analysis as the primary means of detecting drug and alcohol abuse. Many of these programs test for use of the five illicit drugs discussed in this chapter as well as for abuse of alcohol. The literature reviewed in this chapter indicates, however, that these drugs or their metabolites remain in users' bodies only a relatively short time after consumption. Table 8.1 shows approximately how many days each of these drugs may be retained at a sufficiently high concentration to measure above the SAMHSA cutoff levels. It should be noted that these figures may be conservative in cases where testing laboratories are able to test for drugs at their testing method's level of detection rather than at the SAMHSA cutoff level.

The literature reviewed for this chapter indicates that research on the retention times of these drugs has been unable to produce uniform and precise time periods during which most of these drugs can reliably be detected by urinalysis. Numerous physical conditions and behavioral factors including the amount and purity of drug taken, frequency and nature of drug use, and drug users' physical characteristics create a complex of variables that affect how long a drug or metabolite may be retained in the body. Nonetheless, these estimates clearly demonstrate that the drugs of abuse that are targeted by testing programs are relatively quickly metabolized and eliminated by normal body physiology. Marijuana and PCP provide exceptions to this to some extent, although even marijuana use by occasional users becomes undetectable in a relatively short period after consumption.

These short retention times indicate that use of urine or breath testing as the primary or only tools for identifying drug and alcohol abusers may have limited value. This may be particularly true for reliance upon random testing for detecting drug and alcohol use. Except in the case of people who use drugs frequently enough that their use is almost continuously detectable, random testing may be too infrequent to effectively identify most users since the retention time of the drugs of interest is very short. Insofar as for-cause testing is concerned, these retention

## Retention

**Table 8.1: Summary of Approximate Drug Detection Time in Days<sup>a</sup>**

DRUG	DAYS
Alcohol	<1
Amphetamines	1
Opiates	2 <sup>b</sup>
Marijuana (Occasional User)	2
Cocaine	2
Phencyclidine (PCP)	8
Marijuana (Chronic User)	21

- a. Detection times in urine, with the exception of alcohol which is breath.
- b. Poppy seed ingestion may extend this to over 3 days (see text).

figures indicate that the interval between identification of possible impairment or an event that triggers testing and the time of test should be as short as is practicable. Furthermore, given how rapidly drug metabolites leave

the body, it is probable that confirmed positive test rates for all types of urine-based drug testing underestimate drug use prevalence in tested populations.

## Endnotes

1. Ronald Wilson, Battelle Seattle Research Center
2. Recent changes in immunoassay and gas chromatography/mass spectrometry (GC/MS) technologies, reported by Marilyn A. Huestis, John Mitchell, and Edward J. Cone (Huestis, Mitchell, and Cone, 1995) increased the specificity of testing and, as a result, decreased the ability to detect marijuana at higher cutoff levels (see *MRO Alert* 5(8), October/November 1995 for a review and discussion of these changes).
3. Although the accuracy of modern testing equipment and techniques is high, the need for human interpretation of drug and alcohol test results is well accepted. The NRC rule, therefore, requires each licensee FFD program to include a Medical Review Officer (MRO). Before a drug test result designated as positive by a drug testing laboratory is reported to licensee management, the Medical Review Officer uses knowledge of drug pharmacology and human physiology to determine whether there may in fact have been a technical error in the testing process or a legitimate reason for the positive test result.

## References

60 FR 57587. (November 16, 1995.) Substance Abuse and Mental Health Services Administration, "Changes to the Testing Cutoff Levels for Opiates for Federal Workplace Drug Testing Programs." *Federal Register*.

*Alcohol and Health*. (1987). National Institute on Alcohol Abuse and Alcoholism. Rockville, MD: Department of Health and Human Services.

Aubre, J. (1985). The urinary excretion of cocaine and metabolites in humans: A kinetic analysis of published data. *Journal of Analytical Toxicology*, 9: 241-245.

Balster, R.L. (1986). Clinical implications of behavioral pharmacology research on phencyclidine. In *Phencyclidine: An Update* (NIDA Research Monograph #64 pp. 148-162). Rockville, MD: National Institute on Drug Abuse.

Barnes et al. (1988). *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues*. Washington, DC: Nuclear Regulatory Commission.

Barnett, G., Licko, V., and Thompson, T. (1985). Behavioral pharmacokinetics of marijuana. *Psychopharmacology*, 85(51): 51-56.

Baselt, R.C. (1984). Letter to the editor. *Journal of Analytical Toxicology*, 8: 16A.

Baselt, R.C., and Chang, R. (1987). Urinary excretion of cocaine and benzoylecgonine following oral ingestion in a single subject. *Journal of Analytical Toxicology*, 11: 81.

Blum, K. (1984). *Handbook of Abusable Drugs*. New York: Gardner Press.

Bureau of Justice Statistics. (1992). *Drugs, Crime, and the Justice System*. Washington, DC: U.S. Department of Justice.

Caldwell, J. (Ed.). (1981). *Amphetamines and Related Stimulants: Chemical, Biological, Clinical, and Sociological Aspects*. Boca Raton, FL: CRC Press.

Cody, J.T., and Schwarzhoff, R.H. (1989). Impact of adulterants on RIA analysis of urine for drugs of abuse. *Journal of Analytical Toxicology*, 13: 277-284.

Cone, E.J., Menchen, S.L., Paul, B.D., Mell, L.D., and Mitchell, J. (1988). Validity testing of commercial urine cocaine metabolite assays: I. Assay detection times, individual excretion patterns, and kinetics after cocaine administration to humans. *Journal of Forensic Sciences*, 15-31.

Cone, E.J., and Welch, P. (1989). Time course of detection of 6-acetylmorphine in urine after heroin administration. In L.S. Harris (Ed.) *Problems of Drug Dependence 1989* (NIDA Research Monograph #95, p. 449).

Cone, E.J., Yousefnejad, D., Hillsgrave, M.J., Holicky, B., and Darwin, W.D. (1995). Passive inhalation of cocaine. *Journal of Analytical Toxicology*, 19, 399-411.

Crider, R. (1986). Phencyclidine: Changing abuse patterns. In *Phencyclidine: An Update* (NIDA Research Monograph #64 pp. 163-173). Rockville, MD: National Institute on Drug Abuse.

Dakis, C.A. (1982). Persistence of urinary marijuana levels after supervised abstinence. *American Journal of Psychiatry*, 139(9): 1196-1198.

DuPont, R.L. (1991). *Drug-Caused Impairment in the Workplace*. Rockville, MD: Bensinger, DuPont Associates, Inc.

Ekman, G., Frankenhaeuser, M., Goldberg, L., Bjerver, K., Järpe, G., and Mystern, A. (1963). Effects of alcohol intake on subjective and objective variables over a five-hour period. *Psychopharmacologia*, 4: 28-38.

Ellis, G.M., Mann, M.A., Judson, B.A., Schramm, N.T., and Tashehian, A. (1985). Excretion patterns of cannabinoid metabolites after last use in a group of chronic users. *Clinical Pharmacology and Therapeutics*, 38: 572-578.

ElSohly, M.A., Stanford, D.F., and ElSohly, H.N. (1986). Coca tea and urinalysis for cocaine metabolites (Letter to the Editor). *Journal of Analytical Toxicology*, 10: 256.

Fehn, J. and Megges, G. (1985). Detection of D<sup>6</sup>-Monoacetylmorphine in urine samples by GC/MS as evidence for heroin use. *Journal of Analytical Toxicology*, 9, 134-138.

Fischman, M.W. (1984). The behavioral pharmacology of cocaine in humans. *NIDA Research Monograph Series #50* (pp. 72-92). Rockville, MD: National Institute on Drug Abuse.

Fortner, N.A., Moezpoor, E.L., Warren, M., Fogerson, R.S., and Wade, N.A. (Presentation at the January 1994 American Academy of Forensic Sciences Conference). Urine morphine and 6-MAM levels from a parolee population.

Fretthold, D.W. (1990). Drug testing methods and reliability. *Journal of Psychoactive Drugs*, 22(4): 419-428.

Gerchow, J., Heifer, U., Schewe, G., Schwerd, W., and Zink, P. (1985). Calculating the maximum blood alcohol concentration and its application in evaluating the degree of impairment of responsibility. *Blutalkohol*, 22(1): 77-107.

Gorelick, D.A., Wilkins, J.N., and Wong, C. (1986). Diagnosis and treatment of chronic phencyclidine (PCP) abuse. *NIDA Research Monograph Series #64* (pp. 218-228). Rockville, MD: National Institute on Drug Abuse.

Hawks, R.L. (1982). The constituents of cannabis and the disposition and metabolism of cannabinoids. In R.L. Hawks (Ed.), *The Analysis of Cannabinoids in Biological Fluids* (NIDA Research Monograph #42, pp. 125-137). Rockville, MD: National Institute on Drug Abuse.

Hawks, R.L., and Chiang, C.N. (1986). Examples of specific drug assays. In R.L. Hawks and C.N. Chiang (Eds.), *Urine Testing for Drugs of Abuse* (NIDA Research Series #73, pp. 84-112). Rockville, MD: National Institute on Drug Abuse.

Holbrook, J.M. (1983). CNS stimulants. In G. Bennett, C. Vourakis, and D.S. Woolf (Eds.) *Substance Abuse: Pharmacologic, Developmental, and Clinical Perspectives*. New York: John Wiley.

Huestis, M.A., Mitchell, J.M. and Cone, E.J. (1995). Detection times of marijuana metabolites in urine by immunoassay and GC-MS. *Journal of Analytical Toxicology* (19): 443-449.

Jacobs, M.R., and Fehr, K.O. (1987). *Drugs and Drug Abuse*. 2nd Edition. Toronto: Alcoholism and Drug Addiction Research Foundation.

Javaid, J.E., Fischman, M.W., Schuster, C.R., Dekirmenjian, H., and Davis, J.M. (1978). Cocaine plasma concentration: Relation to physiological and subjective effects in humans. *Science*, 202: 227-228.

Johansson, E., and Halldin, M.M. (1989). Urinary excretion half-life of D<sup>1</sup>-tetrahydrocannabinol-7-oic acid in heavy marijuana users after smoking. *Journal of Analytical Toxicology*, 13: 218-223.

Jones, R.T. (1984). The pharmacology of cocaine. *NIDA Research Monograph Series #50* (pp. 34-53). Rockville, MD: National Institute on Drug Abuse.

Ling, W., and Wesson, D.R. (1990). Drugs of abuse: Opiates. *Western Journal of Medicine*, 152(5): 565-572.

Marwah, J., and Pitts, D.K. (1986). Psychopharmacology of phencyclidine. In *Phencyclidine: An Update* (NIDA Research Monograph #64 pp. 127-133). Rockville, MD: National Institute on Drug Abuse.

Morgan, J.P. (1988). Marijuana metabolism in the context of urine testing for cannabinoid metabolite. *Journal of Psychoactive Drugs*, 20(1): 107-115.

Mulè, S.J., and Casella, G.A. (1988). Rendering the "Poppy-Seed Defense" defenseless: Identification of 6-monoacetylmorphine in urine by gas chromatography/mass spectroscopy. *Clinical Chemistry*, 34(7): 1427-1430.

National Center for Forensic Science. (1992). *Drugs of Abuse: Reference Chart*. Baltimore, MD: Maryland Medical Laboratory, Inc.

National Institute on Drug Abuse (NIDA). (1989, May). Marijuana Update. *NIDA Capsules*. Rockville, MD: Author.

National Institute on Drug Abuse (NIDA). (1991). *Technical Advisory* (March 11, 1991). Rockville, MD: Author.

Paul, B.D., Mitchell, J.M., and Mell, L.D., Jr. (1988). Gas chromatography/electron impact mass fragmentometric determination of urinary 6-acetylmorphine, a metabolite of heroin. *Journal of Analytical Toxicology*, 13: 2-7.

Schwartz, R.H., and Hawks, R.L. (1985). Laboratory detection of marijuana use. *JAMA*, 254(6): 788-792.

## Retention

Selavka, C.M. (1991). Poppy seed ingestion as a contributing factor to opiate positive urinalysis results: The Pacific Perspective. *Journal of Forensic Sciences, JFACA*, 36(3): 685-696.

Struempler, R.E. (1987). Excretion of codeine and morphine following ingestion of poppy seeds. *Journal of Analytical Toxicology*, 11: 97-99.

Sunshine, I. (1985). *Monograph: Marijuana (THC)*. Kansas City, MO: Analytical Systems.

Tennant, F.S., Jr., Rawson, R.A., and McCann, M.A. (1981). Withdrawal from chronic phencyclidine (PCP) dependence with desipramine. *American Journal of Psychiatry*, 138: 845-847.

Westra, C., Forslund, C., Field, I., Gutierrez, J., Durbin, N., Grant, T. and Moffitt, R. (1994). *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports CY 1993* (NUREG/CR-5758, Vol. 4). Washington, D.C.: U.S. Nuclear Regulatory Commission, Office of Nuclear Reactor Regulation.

## Bibliography on Polydrug Use

Bailey, D.N. (1984). Comprehensive toxicology screening: The frequency of finding other drugs in addition to ethanol. *Clinical Toxicology*, 22(5): 463-471.

Bailey, D.N. (1990) Drug use in patients admitted to a university trauma center: Results of limited (rather than comprehensive) toxicology screening. *Journal of Analytical Toxicology*, 14: 22-24.

Chesher, G.B. (1986). The effects of alcohol and marijuana in combination: A review. *Alcohol, Drugs, and Driving*, 2(3-4): 105-119.

Evans, M.A., Martz, R., Rodda, B.E., Lemberger, L., and Forney, R.B. (1976). Effects of marijuana-dextroamphetamine combination. *Clinical Pharmacology and Therapeutics*, 20(3): 350-358.

Feigelman, W. (1987). Day-care treatment for multiple drug abusing adolescents: Social factors linked with completing treatment. *Journal of Psychoactive Drugs*, 19(4): 335-343.

Hawkins, M.R., Kruzich, D.J., Smith, J.D., and Silsby, H.D. (1985). Prevalence of polydrug use among alcoholic soldiers. *American Journal of Drug and Alcohol Abuse*, 11(1-2): 27-35.

Kaufman, E. (1982) The relationship of alcoholism and alcohol abuse to the abuse of other drugs. *American Journal of Drug and Alcohol Abuse*, 9(1): 1-17.

Lemberger, L., Dalton, B., Martz, R., Rodda, B., and Forney, R. (1976). Clinical studies on the interaction of psychopharmacologic agents with marihuana. *Annals of the New York Academy of Sciences*, 281: 219-228.

National Institute on Drug Abuse (NIDA). (1980, June). *Marijuana Research Findings, 1980*. (NIDA Research Monograph Series #31). Robert C. Petersen (Ed.). Rockville, MD: Author.

Peck, R.C., Biasotti, A., Boland, P.N., Mallory, C., and Reeve, V. (1986). The effects of marijuana and alcohol on actual driving performance. *Alcohol, Drugs, and Driving*, 2(3-4): 135-154.

Seivewright, N., and Dougal, W. (1993). Withdrawal symptoms from high dose benzodiazepines in poly drug users. *Drug and Alcohol Dependence*, 32: 15-23.

Skinner, M.H., and Thompson, D.A. (1992). Pharmacologic considerations in the treatment of substance abuse. *Southern Medical Journal*, 85(12): 1207-19.

Sweeney, J.A., Meisel, L., Walsh, V.L., and Castrovinci, D. (1989). Assessment of cognitive functioning in poly-substance abusers. *Clinical Psychology*, 45(2): 346-357.

# CHAPTER 9: EFFECT OF STORAGE TEMPERATURE AND EXTREME pH LEVELS ON MARIJUANA AND COCAINE METABOLITE CONCENTRATIONS

Arthur M. Zebelman and Ronald A. Wilson<sup>1</sup>

## 9.1: Introduction

Many drug testing programs, including the fitness-for-duty (FFD) programs of nuclear power plant licensees required by the Nuclear Regulatory Commission (NRC), rely on immunoassay and GC/MS testing to identify abusers of marijuana and cocaine. The ability of these tests to identify drug use by detecting drug metabolites in urine is critical to the success of any testing program. Concern has been expressed regarding whether specimen deterioration prior to testing may reduce the ability of the testing procedures to detect drug metabolites. In particular, the effect of temperature on the concentration of cocaine and marijuana metabolites in urine samples is of interest to both testing laboratories and to organizations with drug testing programs. Heat may cause specimen deterioration that could result in a reduction of the level of drug metabolite in a specimen or could affect the ability of immunoassay tests to detect a drug metabolite. Hence, storage or transportation in a heated condition may potentially result in a false negative test result. In addition, the pH level of a specimen may affect the outcome of testing (see Chapter 6 for a discussion of testing for pH). Although these risks exist in theory, the empirical effect of temperatures on urine specimens with various pH concentrations has not been well documented in the literature.

Although little research exists on the effects of heat on the quality of urine samples, concern over the effects of temperature on the quality of urine samples is reflected in the standard practices of laboratories, which reduce potential deterioration by storing specimens at cool temperatures and testing specimens quickly. However, many drug testing programs collect specimens a substantial distance from the testing laboratory and specimens are neither kept at cool temperatures nor tested quickly.

The question of the effect of temperature and time on specimen quality is of particular concern to those responsible for the U.S. Nuclear Regulatory Commission's (NRC) fitness-for-duty (FFD) rule (10 CFR 26) implementation. This rule provides guidelines for handling and storing urine specimens collected at commercially licensed nuclear power facilities. The guidelines recommend transporting specimens immediately for testing and require that if six hours or more will elapse before the specimen is shipped, then the specimen must be stored at or below 6° Centigrade (C).

Similarly, once specimens have arrived at the laboratory, those not screened within seven days must be kept at or below 6°C.<sup>2</sup> However, there are currently no guidelines requiring that specimens be transported within a specified period of time nor regarding the condition of the specimens during transport.

Under these guidelines, it is possible for urine specimens to be stored and transported at high temperatures prior to testing and for several days to elapse from the time the specimens are collected to the time when the specimens are tested for the presence of drug metabolites. Specimen transportation (e.g., via courier, Postal Service, or private forwarding companies) can require from one to four days to reach the testing laboratory. Specimens may be subjected to high temperatures for extended periods of time if, for instance, the specimen is transported in the cargo hold of an aircraft or in a courier's vehicle during the summer. Upon arriving at the laboratory facility, the current NRC FFD regulation allows specimens to be stored for up to seven days without refrigeration before being tested. Hence, in an extreme case, prior to testing, a urine specimen could be subjected to four days shipment plus another seven days at the laboratory without being refrigerated.

In order to address whether these conditions actually result in a loss of drug metabolites in urine samples, the NRC commissioned a small pilot study of the effects of temperature and pH on the level of drug metabolites in urine samples. This chapter reports the results of this NRC-sponsored pilot study. The pilot study was limited to the effects of temperature and pH on marijuana and cocaine metabolites in urine. Together, these two drugs accounted for about 73 percent of the reported positive drug tests in the nuclear power industry in 1993 (Westra et al, 1994). The other three drug classes for which urine testing is required in FFD programs—opiates, amphetamines, and phencyclidine—represent only 4 percent of reported positive tests, and were not examined in this study. (Alcohol testing, which is conducted with breath alcohol testing devices, not urine analysis, accounted for the remaining 23 percent of confirmed positive test results reported by nuclear power licensees in 1993 [Westra et al., 1994].)

The pilot study consisted of two experiments. The first experiment tested the effects of temperature on the stability in urine of benzoylecgonine (BE)—a metabolite of cocaine, and delta-9-tetrahydrocannabinol-9-carboxylic

## Specimen Deterioration

acid (THC)—a metabolite of the principle psychoactive chemical in marijuana. Stability was tested over a three-day period. Samples were kept at four different temperatures: -70°C (solidly frozen), 4°C (refrigerated), 24°C (room temperature), and 37°C (a temperature that would be reached on a hot day or in an enclosed car left in the sun). Since variation in pH can occur across samples, and because such variation may affect the stability of the sample, variations in pH as well as variations in temperature were examined. The second experiment examined the effects of storage at 37°C for 8-, 12-, 36-, and 72-hour periods on the stability of BE in urine samples having high pH levels.

### 9.2: Specimen Stability Pilot Study

The purpose of the first experiment was to evaluate the effect of various temperatures on cocaine and marijuana metabolite concentrations in urine specimens with various pH levels over a three-day (72 hour) period.

#### 9.2.1: Approach

Spiked solutions consisting of THC and BE were obtained from the Radian Corporation<sup>3</sup>. Two liters of drug-free urine were collected from several volunteers. The pool was mixed and divided into 3 portions. The pH of the apportioned urine was adjusted to 5.0, 7.0, and 9.0, respectively, using 0.1 Molar (M) of acetate and carbonate buffers. The three pH adjusted portions were each then divided into four batches—resulting in three sets of four batches. One batch of each set was maintained as a blank. The remaining three batches were each divided in half, with half spiked with methanolic stock solutions of THC and the other half spiked with BE to achieve drug concentrations approximately two, three, and five times the NRC screening cutoff levels for cocaine and marijuana metabolites.<sup>4</sup> This resulted in twelve conditions for each drug metabolite. The pH-adjusted specimens were further divided into and then stored in four separate 30 mL aliquot polyethylene bottles (this type of bottle is routinely used for urine drug collection and storage). This provided four test samples for each of the 12 conditions for each drug metabolite. These sets of blank and spiked specimens at pH 5.0, 7.0, and 9.0 were stored at four different temperatures, -70°C, 4°C, 24°C, and 37°C, for 72 hours. Hence, for each drug metabolite, 48 different conditions were tested. All of the resulting categories are listed in Table 9.1.

After storage at the four different temperatures for 72 hours, all specimens were frozen at -70°C until analysis

by immunoassay and gas chromatography/mass spectroscopy (GC/MS). The specimens were analyzed singly in five separate GC/MS and immunoassay runs either for BE or for THC. The BE and THC GC/MS procedures were those routinely used for forensic specimens performed under the Substance Abuse and Mental Health Services Administration (SAMHSA) guidelines (National Institute on Drug Abuse, 1980). Three levels of controls (i.e., specimens not spiked with drug metabolites, or blank specimens) were included with each batch of specimens run through GC/MS testing.

#### 9.2.2: Results

The results of the immunoassay and GC/MS analysis of the blank and spiked, pH-adjusted urine specimens are provided in Tables 9.1 and 9.2, respectively. The results are summarized below.

##### 9.2.2.1: Immunoassay Results

As expected, the immunoassay results were negative for all blank specimens for BE and THC (see Table 9.1).

The BE immunoassay results were positive for all specimens other than the blanks, regardless of pH and storage conditions. Similarly, all THC immunoassay results were positive when a cutoff level of 20 nanogram per milliliter (ng/mL) was used. Using the higher THC cutoff level of 100 ng/mL (the level specified at the time of the study in DHHS guidelines), two of the three urine specimens spiked at double the cutoff level (200 ng/mL) with a pH level of 5 were screened by immunoassay as negative in spite of the fact that their concentration was measured by GC/MS as greater than 100 ng/mL of THC, i.e., false negative screening results. This may have been due to the known problem that low pH causes with enzyme immunoassay. The specimens spiked with THC at three and five times the cutoff level were positive for marijuana except for the specimen spiked at three times the cutoff level at pH 5 stored at 24°C. This latter result may be due, once again, to the low pH of the specimen, or due to the fact that the immunoassay tests for these specimens were delayed for one day after the GC/MS testing and the volume of urine remaining after the GC/MS analysis was less than two mL. There may also have been some losses of THC as a result of THC binding to the walls of the specimen container.

Table 9.1: Experimental Results: EMIT® Immunoassay

		THC						BE			
pH	Target Concentration (ng/mL)	Temperature (°C)				pH	Target Concentration (ng/mL)	Temperature (°C)			
		-70	4	24	37			-70	4	24	37
		delta A/min at 100 ng/mL cutoff level and at 20 ng/mL cutoff level						delta A/min at 300 ng/mL cutoff level			
5	Blank (0)	-620 -502	-605 -510	-403 -329	-495 -542	5	Blank (0)	-235	-234	-256	-228
	200	-51 736	-357 643	-315 480	669 715		600	110	108	124	130
	300	668 733	287 693	-303 689	875 701		900	181	190	182	198
	500	694 730	802 726	321 890	852 658		1,500	231	249	231	253
7	Blank (0)	-417 -371	-478 -611	-525 -449	-529 -562	7	Blank (0)	-238	-244	-247	-238
	200	306 779	514 649	436 763	608 625		600	100	86	97	104
	300	876 738	818 678	723 711	818 608		900	175	177	138	169
	500	872 785	862 651	727 777	854 625		1,500	245	243	243	239
9	Blank (0)	-432 -429	-502 -579	-520 -438	-564 -474	9	Blank (0)	-239	-218	-234	-223
	200	615 884	751 735	415 785	291 714		600	107	105	102	74
	300	904 886	883 708	773 778	769 718		900	180	177	145	123
	500	891 880	871 772	746 794	787 669		1,500	243	267	240	200

## Note:

The immunoassay results are expressed as a reaction rate relative to the cutoff levels specified at the time of the study by the Department of Health and Human Services (DHHS), Substance Abuse and Mental Health Services Administration (SAMHSA), 100 ng/mL and 300 ng/mL for THC and BE, respectively. Experiments were also conducted using a 20 ng/mL THC cutoff for comparison purposes. Immunoassay results that indicate a metabolite concentration less than the cutoff level are shown as negative numbers.

## Specimen Deterioration

### 9.2.2.2: GC/MS Results

The GC/MS results are provided in Table 9.2. Once again, as expected, the blank specimens were negative for BE and THC metabolites by GC/MS analysis. All specimens spiked with THC were positive at a 100 ng/mL cutoff level and all specimens for BE were positive at a 300 ng/mL cutoff level. However, two specimens at 37°C and pH 9 showed significantly lower levels of BE than expected. In order to further explore this finding, a second pilot study was conducted.

## 9.3: Continuation of Specimen Stability Pilot Study

Based on the results of the first experiment, a second experiment was undertaken to determine if there was a significant loss of BE in spiked urine samples when the pH is 8 or 9 and the samples are stored at 37°C for times varying from 8 to 72 hours. The purpose of this effort was to determine how long it takes for BE metabolite

concentrations to degrade in high pH specimens subjected to high temperature conditions.

### 9.3.1: Approach

Similar to the approach used in the first experiment, a drug-free urine pool was collected. Two portions of the drug-free urine were prepared at pH 8 and 9, respectively. Two batches at each pH were prepared, one batch from each pH level was left blank and the other was spiked with BE at a concentration targeted at 300 ng/mL, the screening cutoff level for cocaine mandated by DHHS guidelines at the time of the study. Two specimens were prepared at each pH from the spiked batches—resulting in two identical specimens in each testing category (see Table 9.3). Two specimens at each pH were stored at 37°C for different time intervals—8, 12, 36, and 72 hours. Two specimens at pH 8, and two at pH 9 were frozen immediately for control purposes. At the end of each time period all specimens were frozen at -70°C until analysis by GC/MS.

**Table 9.2: Experimental Results: Gas Chromatography/Mass Spectroscopy (GC/MS)**

		THC						BE			
pH	Target Concentration (ng/mL)	Temperature(°C)				pH	Target Concentration (ng/mL)	Temperature(°C)			
		-70	4	24	37			-70	4	24	37
		{Actual Metabolite Levels} [ng/mL]						{Actual Metabolite Levels} [ng/mL]			
5	Blank (0)	0	0	0	0	5	Blank (0)	0	0	0	0
	200	186	122	179	222		600	555	586	566	549
	300	277	230	264	315		900	773	848	816	728
	500	452	343	408	481		1,500	1,313	1,378	1,323	1,308
7	Blank (0)	0	0	0	0	7	Blank (0)	0	0	0	0
	200	201	200	212	216		600	541	531	559	504
	300	285	286	289	223		900	807	782	786	800
	500	459	424	472	421		1,500	1,301	1,308	1,223	1,197
9	Blank (0)	0	0	0	0	9	Blank (0)	0	0	0	0
	200	235	208	197	197		600	536	495	443	387
	300	344	321	301	266		900	814	740	1,002	647
	500	588	423	470	464		1,500	1,398	1,230	1,635	932

### 9.3.2: Results

The results of the duplicate GC/MS analysis of each of the specimens are provided in Table 9.3. The BE concentrations for the nonzero storage times were compared to the concentrations of BE measured in the specimens that were frozen immediately. At pH 8 the two samples stored for 36 hours at 37°C were negative. At pH 9 six of the eight specimens tested negative after storage.

## 9.4: Discussion

The results from this pilot study are limited by the small number of specimens analyzed. Nevertheless, several findings emerged. First, there were a number of false negative test results for THC using immunoassay tests when the pH level of the sample was 5. A pH level of 5 is within the normal range of pH for urine. Although a number of specimens tested negative in the immunoassay test when the pH level was 5, THC did not show a significant deterioration during storage, as shown in Tables 9.1 and 9.2. This would suggest that although THC metabolite concentrations are not severely affected by heat, high cutoff levels combined with a low pH of specimens may lead to false-negative results. Second, although storing a specimen with pH level of 9 at 37°C (approximately 99°F) did not result in false-negative test results using GC/MS when the specimen was spiked at two or more times the cutoff level required by DHHS guidelines, the difference in the BE concentration after storage showed a significant reduction in the concentration of the cocaine metabolite BE in two of three specimens tested in the first experiment. There was no significant reduction in the BE concentration in

specimens of pH 5, 7, or 9 when the specimens were stored at refrigerator temperature (4°C) or at normal room temperature (24°C).

The second experiment involving BE metabolite concentration degradation at higher temperatures provided several results. Two of the eight specimens having a pH level of 8 were lower than the cutoff level, however, these specimens were still within the normal range of testing variation and there does not appear to be a significant loss of BE at pH 8 for storage times up to 72 hours at 37°C. At pH 9 six of the eight samples (at 8, 36 and 72 hours) resulted in BE concentrations lower than 300 ng/mL, however, there does not appear to be a significant loss of BE for at least 12 hours of storage at 37°C. The lower concentrations of BE at 36 and 72 hours indicated a significant loss of BE at pH 9 after 36 hours of storage at 37°C. Cases where the BE metabolite concentration is close to the cutoff level of 300 ng/mL in urine specimens with a high pH and high storage or transportation temperatures would probably result in false-negative test results.

Although uncommon, a urinary pH of 8 or 9 is possible. In a review of 250 consecutive urine samples by one of the authors, 2.8 percent had a pH of 8. There were no samples with a pH of 9 in the particular lot of 250 urine specimens. However, urine measured at pH 9 has been reported. The primary cause of high pH in urine is through the action of bacteria (the production of ammonia from urea). Specimens that are allowed to remain at room temperature or higher for 72 hours or longer will usually have a higher than normal pH. A worker with a moderate to severe urinary tract infection will usually submit a specimen already heavily inoculated with live bacteria. Should such a specimen remain in the holding,

Table 9.3: Benzoylecgonine (BE) Concentration Following Storage at 37°C for Various Times

Storage Time (hours)	Concentration (ng/mL) pH 8		Concentration (ng/mL) pH 9	
	0		0	
Blank				
0 (immediately frozen)	333	297	311	322
8	333	366	296	298
12	329	307	368	329
36	273	276	265	270
72	319	315	253	253

## Specimen Deterioration

transportation, and storage process for up to 7 days (which is possible under current NRC guidelines), the concentration of bacteria could dramatically alter the specimen pH level. If not detected prior to testing, this could result in a lower BE metabolite concentration, resulting in a false-negative test result.

As with all experiments, these experiments do not fully reflect actual conditions. In addition to the use of pH levels that are uncommon in the population, this study used pooled urine samples. Interindividual variations in the urine matrix might yield different results in stability studies carried out with a larger number of individual specimens from different donors. It is unlikely, however, that BE would be found to be more stable than indicated in this study.

### 9.5: Conclusion

This pilot study was designed to address whether specimens stored and/or shipped at different temperatures may deteriorate over time. It is the underlying assumption in testing programs that if a drug or drug metabolite is present in a freshly voided specimen, it should remain in the specimen in the same concentration i.e., unchanged, when tested several days later at the

laboratory. Although gross deterioration and loss of drug metabolite do not occur rapidly, this pilot study provides initial empirical support that urine specimen test results containing the cocaine metabolite BE may be adversely affected if subjected to high temperature conditions for several days prior to testing. Reductions in BE concentrations due to high heat could result in false-negative test results. Proper shipping of urine specimens when testing for the presence of BE should ensure that the specimen is protected from temperatures above 24°C. If temperatures are likely to approach 37°C, provision should be made to cool the specimens. None of the specimens containing the marijuana metabolite THC contained significantly less THC after being subjected to high temperatures for several days. However, false-negative results for THC were found in samples with pH levels of 5.

This was a limited pilot study. The results indicate that additional research may be justified using a much larger number of trials. In particular, the deterioration of drug or drug metabolites in urine specimens due to environmental changes such as continuous variations in temperature, container material, and handling are areas that would benefit from further examination.

## Endnotes

1. Arthur M. Zebelman, Ph.D, Director of Clinical Operations and Scientific Director, Dynacare Laboratory of Pathology of Seattle, Inc.  
Ronald A. Wilson, Battelle Seattle Research Center
2. 10 CFR Part 26, Appendix A, § 2.7 (c).
3. DL-11-nor-9-carboxy-delta-9-THC, catalogue no. T-006, 100 ug/mL in methanol, and benzoylecgonine, catalogue no. B-004, 1 mg/mL in methanol. Radian Corporation.
4. The maximum NRC cutoff level for marijuana metabolite was 100 nanograms per milliliter (ng/mL) at the time of this study. The NRC cutoff level for cocaine metabolite is 300 ng/mL. That is, urine specimens containing drug metabolites in concentrations above the cutoff level indicate that the specimen is positive for the drug in question. Initial positive results are then confirmed with a more rigorous gas chromatography/mass spectroscopy (GC/MS) analysis (National Institute on Drug Abuse, 1980).

## References

10 CFR Part 26 (January 1, 1990). U.S. Nuclear Regulatory Commission. "Fitness for duty programs." *U.S. Code of Federal Regulations*.

Fleming, T., Westra, C., Field, I., Forslund, C., Durbin, N., Grant, T., and Moffit, R. (1993). *Fitness for Duty in the Nuclear Power Industry: Annual Survey of Program Performance Reports - CY 1992 - NUREG/CR-5758, Volume 3*. Washington, DC: U.S. Nuclear Regulatory Commission.

National Institute on Drug Abuse. (1980). *GC/MS Assays for Abused Drugs in Body Fluids* (NIDA Research Monograph #32), Rockville, MD: Author.

Westra, C., Durbin, N., Field, I., Wilson, R., Battelle Seattle Research Center, Hattrup, M., Cunningham, M., Pacific Northwest Laboratory (August 1995). *Fitness for Duty in the Nuclear Power Industry: Annual Summary of Program Performance Reports CY 1994* (NUREG/CR 5758, Vol. 5). Washington, D.C.: U.S. Nuclear Regulatory Commission, Office of Nuclear Reactor Regulation.

# CHAPTER 10: PERFORMANCE-BASED ASSESSMENT TESTS

**Jody Toquam and Alvah Bittner<sup>1</sup>**

Performance-based assessment tests can provide direct measures of an individual's ability or inability to perform specific tasks required on the job. These tests have been proposed to detect worker performance impairment, a key concern in many industries. For example, for jobs in the commercial nuclear industry, the Nuclear Regulatory Commission's (NRC) fitness-for-duty (FFD) regulation specifies two primary program performance objectives. These include assuring that nuclear power plant personnel are free from any mental or physical impairment that would in any way adversely affect their ability to safely and competently perform their job duties, and that they will perform their tasks in a trustworthy and reliable manner. If empirical evidence for performance-based tests demonstrate that these tests can be used to detect impairments, such tests could be used to address the primary purpose of the FFD program—identification of impaired workers. In addition, performance-based testing may be used as part of a comprehensive program to assure that workers do not violate the provisions of FFD programs that prohibit the use of illegal drugs and the misuse of alcohol.

Previous reviews of the technical issues surrounding the NRC's FFD program examined performance-based tests and test batteries and concluded that the empirical data did *not* support using these tests in a FFD program (Moore et al., 1989; Durbin et al., 1991). Current research, however, supports the reexamination of these tests. For example, during the past few years, continued research using performance-based testing technology has advanced this area by examining performance-based testing methods and outcomes. In addition, the number of vendors marketing these tests and test batteries for purposes of FFD testing has increased. Government agencies and research programs are currently evaluating this body of literature and corresponding vendor materials to determine the feasibility of performance-based tests for safety-sensitive jobs (e.g., Federal Aviation Administration and National Research Council).<sup>2</sup> There is also evidence indicating that these tests are being used in some industries to screen out impaired workers on a daily basis. Results from the American Management Association survey of drug testing policies in 1993 indicated that about one percent of the companies surveyed used performance-based testing, but these tests were often used in conjunction with other FFD testing procedures (American Management Association, 1993. See also Chapter 3, Prevalence of Drug Use, in this volume for a more detailed discussion of this survey). Finally, a recent informal survey of performance-based test vendors indicates that many process-control industries and other organizations are currently using performance-

based testing to ensure worker fitness for duty (see Section 10.2.2 for details from vendors).

The purpose of this chapter is to provide a detailed examination of performance-based tests that may potentially have a role in the commercial nuclear power industry's FFD testing programs. Toward this end, this examination identifies and describes a variety of specific tests that have been developed and have potential for use in an FFD program. An evaluation of these tests is provided in terms of the ability of each test to identify impaired employees (i.e., those who cannot adequately perform their required job duties).

This chapter is organized as follows: Section 10.1 describes test standards used to evaluate performance-based tests of performance impairment problems; Section 10.2 summarizes and discusses the current performance-based test literature with a focus on the desirable characteristics and test standards; Section 10.3 provides a demonstration of procedures for calculating detection accuracy for a selected performance-based test; and Section 10.4 summarizes the status of performance-based tests and discusses issues in performance-based testing that must be addressed before such tests can be implemented in a FFD program.

## **10.1: Scope and Optimal Characteristics of Performance-Based Tests**

This section provides preliminary information concerning the procedures used to evaluate performance-based tests. Initially, these tests can be evaluated by examining qualitative characteristics of the tests (e.g., scoring procedures). More thorough evaluations of these tests include examinations of empirical (quantitative) data obtained from research conducted to investigate psychometric characteristics of each test. Those qualitative and quantitative characteristics of interest for evaluating the feasibility of performance-based tests are described in detail in this section.

### **10.1.1: Scope of Performance-Based Tests**

Performance-based tests are designed to detect behavioral and physiological indications of impairment and, thereby, to assess an employee's current ability to perform effectively on the job. Performance-based tests differ from traditional urinalysis tests to detect drug use or breath tests to detect alcohol use in two important ways:

## Performance-Based Tests

- Performance-based tests have been shown to effectively detect impairment caused by both non-drug and drug agents. These nondrug-related impairment factors include sleep loss, stress, exposure to chemicals, and altitude changes, as well as interactions of combinations of these and other stressors. Such tests would also identify impairment caused by drugs that are often not included among those tested by urinalysis, such as certain over-the-counter (OTC) or prescription medications.
- Performance-based tests are designed to detect impairment, but not its cause.<sup>3</sup> Therefore, once an employee is deemed to be impaired, the organization must use additional tests and measures to identify the cause of impairment.

These differences have important implications for meeting the performance objectives of FFD programs in safety-sensitive industries. First, performance-based tests may provide the unique capability of detecting impairment that is unrelated to drug or alcohol use and abuse. Similar to FFD programs that rely on behavioral observations by supervisors, performance-based tests provide broader coverage of agents and sources of impairment (e.g., drug use, exposure to chemicals, sleep loss, illness, stress) than traditional FFD tests, such as urinalysis, which test only for use of a panel of drugs. Because sources of impairment other than drug and alcohol use can immediately affect job performance, improvements in detecting these types of impairment could enhance public safety. Second, performance-based tests *cannot* provide direct evidence that an individual has violated a policy against drug or alcohol use. Because performance-based tests are designed to assess impairment without necessarily attributing the cause of impairment, organizational decisions or actions taken as a result of test scores are generally not based on an attribution of illegal or inappropriate drug or alcohol use. Additional steps are needed to determine if the impairment was caused by illegal or inappropriate drug or alcohol use.

### 10.1.2: Optimal Characteristics of Performance-Based Tests

To identify and select the most appropriate tests for a performance-based FFD testing program, it is important to identify the critical tasks for a particular job and then identify the abilities and skills required to perform those critical tasks in a safe and competent manner. At this point, one must select performance-based tests that measure the required abilities and skills and that possess sound psychometric characteristics (e.g., appropriate test

length, high reliability and reliability efficiency, documented test validity and so on). In most organizations, the critical job duties and tasks, as well as the required abilities and skills have been identified and are provided in detailed job analyses documents. It is possible, therefore, to use this information to identify and select the most appropriate performance-based test(s) for each job of interest.

Test administration time and the time required to score and interpret test results must also be considered. Because administration time in any FFD program is generally limited, tests with short administration times are considered optimal (e.g., 3-5 minutes), while tests that require 10 or 20 minutes to administer are impractical for FFD testing purposes. In addition, tests that provide nearly immediate feedback on the employee's performance, so that appropriate actions can be taken in real time (i.e., immediately after completing the test), are preferred for FFD testing purposes. For this reason, tests that require complex administration procedures and apparatus, or that are difficult or time-consuming to score (e.g., tests requiring hand scoring), are generally not used in FFD programs.

In previous reviews of performance-based tests, many of the tests were administered via paper-and-pencil or apparatus. These tests often require longer than desirable administration times and time-consuming scoring and interpretation procedures. With advances in computer technology and test development, many of the performance-based tests evaluated in this review are computer-administered. Administration, scoring, and test score interpretation requirements are much shorter for computer-administered performance-based tests than that for paper-and-pencil tests. As a result, computer-administered performance-based tests currently available are more amenable to FFD programs because of the ease of administration, scoring, and interpretation, relative to previous performance-based tests (e.g., paper-and-pencil).

In sum, performance-based tests that 1) have been selected to measure required abilities and processes for a given job or group of jobs; 2) involve short administration times and simple, yet rapid scoring; and 3) provide immediate interpretation of test scores are most suitable for use in work places having significant safety considerations.

In addition to test administration and scoring considerations, empirical data (quantitative data) supporting the use of performance-based tests should also be examined. This body of empirical data generally

provides information about standards that tests should meet to ensure that these tests effectively detect impairment related to job performance requirements. In general, there are four standards on which tests should be evaluated, including reliability, sensitivity, accuracy, and validity. The importance of each of these four standards for tests, procedures for measuring each, and findings from research regarding current performance-based testing techniques in meeting these requirements are described in greater detail in Sections 10.1.2.1 through 10.1.2.4. Three of these standards, reliability, sensitivity and validity, are used in assessing specific types of performance-based tests in Section 10.2. The fourth standard, accuracy, is discussed and applied to a single selected performance-based test in Section 10.3.

#### 10.1.2.1: Reliability<sup>4</sup>

Test reliability indicates the degree to which a test produces the same result each time and is free from measurement error. Test measurement error can arise from several sources, such as problems in test construction (e.g., poorly written test items that may be interpreted differently by different subjects) or problems with test administration procedures (e.g., improper or inconsistent test administration conditions that may vary for different subjects) (Jensen, 1980). Practice effects are another source of measurement error that can occur in many performance-based tests. For example, some subjects may be more familiar with the test content and testing procedures because they have had the opportunity to practice completing these tests more than other subjects. Further, subjects benefit from practice at different rates. Therefore, it is important to assess the amount of practice time needed for subject's scores to stabilize (i.e., after sufficient practice time is provided, an individual subject's scores should be fairly consistent from one administration to the next). The amount of practice time required for subject's scores to stabilize is another important characteristic of a test.

To reduce measurement error and increase the reliability of any performance-based test, organizations should select well-constructed tests, strictly adhere to administration procedures and allotted times, and allow sufficient practice time for all subjects to reach a plateau before baseline performance data are collected. Estimates of test reliability indicate how well these measurement error problems have been avoided.

Reliability of a particular test is estimated by measuring test stability. This involves administering a test to subjects on one occasion (time 1), and then administering

the *same* test again to the *same* group of subjects a second time (time 2). In most studies, the delay period between the two test administrations ranges from one day to several weeks or months. Test scores are computed for subjects for each occasion (time 1 and time 2) and then the correlation between these scores is computed. This correlation can range from 0.00 (no similarity in performance of the test across administrations) to +1.00 (test performance is exactly the same across administrations),<sup>5</sup> with higher values indicating higher test stability. Values of 0.70 and above are most desirable. The performance-based test literature indicates that reliability is typically assessed using a delay of one or two days.

The literature also provides information about the time needed for test scores to stabilize in terms of the amount of practice required, practice in performing or completing the test (stabilization time). Stabilization time is determined by administering a test to the same group of subjects on several occasions. The correlation between test scores from two different occasions is computed (e.g., time 1 and time 2, time 2 and time 3, and so on) to identify the point at which test reliability stabilizes or changes very little from one occasion to the next. For example, if a 10-minute attention test is administered on six occasions and the correlations stabilize after time 3, stabilization time is 30 minutes (three occasions with 10 minutes per administration).

Because test length (administration time) is a critical factor for performance-based tests, and because administration times for tests vary greatly, researchers in this field have developed an efficiency metric (reliability efficiency) to compare reliability estimates across many different tests. A common metric is needed because as tests increase in length, reliability generally increases. Therefore, it is inappropriate to compare the reliability of a short test with that of a long test. The common reliability efficiency metric is produced by estimating the stability of the full test and then adjusting this estimate using the Spearman-Brown formula<sup>6</sup> to produce an estimate of test stability for a three-minute test (Ghiselli, Campbell and Zedeck, 1981). For example, if a test has a 20-minute administration time and a test stability estimated at 0.95, the reliability efficiency metric can be estimated for the same test with an administration time of three minutes. In this example, the test stability estimate for the shorter test is 0.74 using the Spearman-Brown formula. Estimating reliability efficiency, of course, is based on the assumption that the reduced test is parallel or nearly equivalent to the original test in terms of item content and item difficulty. Thus, the reliability

## Performance-Based Tests

efficiency metric provides a means to compare reliabilities for tests of varying lengths. This technique enables one to determine if a test can still be effective (sufficiently reliable) if test length is reduced. Acceptable reliability efficiency values given as estimates for tests three minutes in length are the same as those for stability estimates (0.70 to 1.00, with higher values being more desirable).

### 10.1.2.2: Sensitivity<sup>7</sup>

Sensitivity is a test characteristic that indicates whether a test can detect performance impairment due to some specific agent (e.g., fatigue or drug intake). Although performance-based tests are designed to simply detect impairment without identifying the cause, it is important to demonstrate that these tests can detect impairments that result from a variety of sources (e.g., drug use, exposure to chemicals, fatigue). For many of the performance-based tests, studies have been conducted to assess the impairment effects of drug or chemical use, fatigue, stress, or environmental exposure to chemicals on test performance. General procedures for assessing sensitivity are as follows: 1) subjects complete a test or series of tests before exposure to a drug, chemical, or condition (pre-exposure); 2) subjects are exposed to the drug, chemical, or condition; and 3) subjects complete the test one or more times to assess the effects of the drug, chemical or condition at different stages (e.g., peak exposure or post exposure—a time period following exposure to the drug, chemical, or physical condition). Researchers may vary the level of exposure to a drug, chemical, or condition to determine the level at which the performance-based test can detect impairment (e.g., varying levels of blood alcohol concentration, or BAC). The purpose of these latter studies is to demonstrate that performance-based tests can detect impairment from these substances at even low doses or short exposure periods. As noted earlier, although these tests can detect impairment from a particular substance, they cannot be used to identify the specific cause of impairment.

Impairment assessment is determined by comparing pre-exposure test scores with post-exposure test scores.

Conclusions about impairment drawn from test scores are influenced by the way in which scores are compared and the magnitude of the differences between test scores. For example, scores from tests administered for educational or occupational purposes are interpreted by using the normative method. This method defines how well a subject has performed relative to all others who have taken the test. For performance-based tests applied in a FFD program, it is more appropriate to use the ipsative

method to interpret test scores. In this method, a subject's test score is compared with previous scores from the *same* subject on the *same* test; previous performance on the test is defined as a subject's baseline performance. If a difference appears between a subject's current test score and baseline performance on the same test, then it is possible to conclude that the subject is impaired. The decision about impairment is a function of the magnitude of the difference between current test score and baseline performance. A baseline performance score is derived by computing a subject's mean score (average) and standard deviation (variation) across scores on the same test completed on several occasions. These values (mean and standard deviation) are used to determine if there is a statistical difference between the baseline and current scores. In most instances, the difference must be large to demonstrate impairment.<sup>8</sup> If a subject's current test score is three standard deviations higher or lower<sup>9</sup> than the baseline level, one can conclude that the person is impaired.

### 10.1.2.3: Validity

Validity indicates that the test measures what it was designed to measure. Test validity can be described in many ways. Two common ways are construct validity and criterion-related validity. Construct validity involves demonstrating that a newly-developed test designed to measure a particular ability does indeed measure that ability. Construct validity can be demonstrated using a wide variety of methods. One common method involves computing the correlation between scores on a newly-developed test with test scores from the same subjects on a well-established test. Another form of validity is criterion-related validity. This form of validity indicates that scores on a test can be used to predict job or task performance.<sup>10</sup> For both construct validity and criterion-related validity, correlations are used to demonstrate a relationship between scores on the performance-based test and either scores on a well-established test (construct validity) or scores from job performance measures (criterion-related validity). These correlations can range from -1.00 to +1.00.

Criterion-related validity (predictive validity) is highlighted in this chapter for two reasons. First, to demonstrate that a test is indeed performance based, it is necessary to demonstrate that performance on the test is linked with performance on the job. This is accomplished by presenting quantitative data that show a statistically significant correlation between test scores and job performance scores. The performance-based test literature does not address this directly. Instead,

researchers provide anecdotal data indicating that a particular test has been or can be used to predict performance in particular occupations. Thus, a second reason for highlighting criterion-related validity is that it represents a vital missing component in the performance-based test literature. To ensure that performance-based tests assess the ability to perform critical requirements of the job, it is important to follow standard procedures for demonstrating the criterion-related validity of a test.<sup>11</sup> These types of studies must be conducted to demonstrate in quantitative terms the relationship between scores from performance-based tests and performance in a particular job (e.g., nuclear power plant operations). Currently, the performance-based test literature provides only qualitative information demonstrating a link between performance-based test scores and job performance (e.g., jobs for which the test have been or are currently used).

#### 10.1.2.4: Accuracy

Accuracy represents the degree of association between decisions made using performance-based test scores (impaired or not impaired) and the actual condition of the subject. Detection accuracy for any performance-based test can be determined by conducting well-defined experiments. A sample experimental procedure for demonstrating accuracy is provided below:

- Randomly assign subjects to an experimental or control group. (Groups should be of approximate equal size.) Throughout the experiment, control subjects simply complete the performance-based tests without exposure to the drug or condition.
- All subjects (experimental and control groups) complete the performance-based test (Time 1).
- Administer drug or experimental condition to experimental group (dosage or length of exposure is selected to produce impairment in subjects).
- All subjects (experimental and control groups) complete the performance-based test (Time 2).
- For all subjects (experimental and control), compute a difference score between Time 1 and Time 2 performance-based test scores. Apply decision rules (e.g., three standard deviation rule) to categorize each subject as impaired or unimpaired.

Under optimal conditions, difference scores (Time 1 - Time 2) and decisions about impairment will correspond perfectly with experimental conditions. This means that

all subjects exposed to a drug or condition would be classified as impaired using these difference scores, while control subjects would be classified as unimpaired. Under such optimal conditions, the performance-based test would be viewed as highly accurate.

A quantitative metric for describing test score accuracy is computed by comparing decisions derived from the performance-based test scores with "true" impairment or non-impairment defined by the experimental conditions. If there is a one-to-one correspondence between decisions based on test scores and experimental conditions, then the test is 100% accurate. If, however, the two differ, the test is less than 100% accurate. Given that there is measurement error in all tests (see Section 10.1.1 on test reliability), we can expect that most tests are less than 100% accurate. Therefore, it is important to determine the level of test accuracy before implementing such tests in a performance-based test system.<sup>12</sup>

## 10.2: Performance-Based Testing Literature

This section summarizes the literature designed to investigate the feasibility of using performance-based tests to determine impairment. The literature reviewed here employs the first three test standards discussed in Section 10.1 (reliability, sensitivity, and validity) to evaluate performance-based tests. Accuracy, as defined in Section 10.1.2.4, is seldom reported in the performance-based test literature. For this reason it is treated separately in Section 10.3.1.

The information reported in this section was obtained from a variety of sources, including the academic literature, government reports, and documents and reports from vendors who market performance-based tests.<sup>13</sup> Results from this literature are provided in two sections: Section 10.2.1 contains a summary and discussion of findings from the academic literature, and Section 10.2.2 contains a summary of current trends and the types of tests available from vendors.

### 10.2.1: Review of Performance-Based Testing Literature

Three broad categories of performance-based tests are considered in this section as candidates for inclusion in FFD batteries. These categories include cognitive ability and processing tests, psychomotor tests, and physiological tests. Cognitive tests assess an individual's processes and abilities to perform mental tasks that range from very

## Performance-Based Tests

simple (e.g., reaction time) to very complex (e.g., inductive reasoning). Psychomotor tests assess one's ability to make a variety of coordinative, manipulative, repetitive, and/or precise body or limb movements. Physiological tests address primarily vision, visual acuity, and eye movement.

Tests falling into each of these categories have been examined as candidates for inclusion in performance-based test batteries that would have the potential to detect job performance impairment. Examples of each type of test are currently being marketed and are in use in perhaps three dozen or more locations nationally. In general, there is reason to expect that one or more of these tests could be used for commercial nuclear applications to assess employee performance impairment.<sup>14</sup>

In the following sections, specific tests that fall into the three broad categories are evaluated. Within the cognitive domain, 10 classes of mental abilities and processes are examined. All psychomotor tests are described within this single category, and the physiological domain includes two classes of test. The following subsections summarize the academic literature for 13 classes of processes and abilities that are important for functioning in the workplace. All measures are evaluated in terms of three of the standards discussed in Section 10.1, reliability, sensitivity, and criterion-related validity. A summary of this information is presented in Table 10.1 for each of the 13 processes and abilities identified in the literature. Discussions of this summary information are provided in Section 10.2.1.1 through Section 10.2.1.4. A more complete summary of each of the tests, as well as the sources of the summary information is provided in Appendix G.

### 10.2.1.1: Cognitive Tests

Ten different cognitive ability or processing tests are discussed in this section. These tests include: attention/divided attention, time estimation, perceptual speed, memory, perception, spatial ability, reasoning, fluency, number facility, and reaction time.

#### Attention/Divided Attention

Attention/divided attention is a cognitive process that involves concentrating to identify visual and auditory stimuli (attention), filtering out irrelevant stimulus input (selective attention), and performing two or more tasks simultaneously (divided attention/dual-task). Ten different attention tests were evaluated as potential

performance-based tests. Administration times range from one to 30 minutes for these tests, with a median administration time of five minutes. Stabilization time was reported for only one test; this test (dual-task) required 100 minutes of test time to stabilize. Reliability efficiency was also reported for only one test. This was estimated at 0.50 for that test (based on the standards for acceptable levels of reliability, reliability for this test is unacceptable). Attention/divided attention tests have been shown to be sensitive to impairments caused by alcohol, amphetamines, diazepam, marijuana, and sleep loss. Although no criterion-related validity information was provided, attention test scores have been shown to be related to job performance for pilots, truck drivers, and other professional drivers.

#### Time Estimation

Time estimation represents the process by which one can estimate elapsed time over a given period. Two tests measuring this cognitive process were identified in the literature review. Administration times for these tests, three and five minutes, are relatively short. Reliability estimates for these tests were not reported. Time estimation test scores have been shown to be sensitive to impairments caused by atropine and carbon monoxide. No criterion-related validity information was provided. Researchers did not identify occupations for which these tests would be useful as performance-based indicators. In general, although this type of test appears to be useful in terms of test length, there is very little empirical data to support using this type of measure as a performance-based test in a FFD program.

#### Perceptual Speed

Perceptual speed tests require subjects to scan material quickly and accurately to compare pairs of objects to determine if they are the same or different, or to identify figures or objects in a complex set of objects. The literature contained descriptions of 30 different tests measuring perceptual speed ability. These tests are generally of short duration, with administration times ranging from one to 50 minutes, with a median length of three minutes. Stabilization times range from 16 minutes to 121 minutes, with a median of 27 minutes. Stability estimates range from 0.55 to 0.90, with a median of 0.85. Reliability efficiency estimated for a three-minute test ranges from 0.52 to 0.93, with a median of 0.85. Perceptual speed test scores have been shown to be sensitive to impairments caused by alcohol, antihistamines, halon, motion sickness drugs, chemotherapy and changes in altitude. These tests are

**Table 10.1: Summary of Performance-Based Test Information**

Description of Process/Ability	# of Tests	Administration Time	Reliability	Sensitivity	Validity
<u>Attention/Divided Attention</u>  Ability to concentrate and identify visual and auditory stimuli; to filter out irrelevant stimulus input; or to perform two tasks simultaneously.	10	1-30 minutes	Stabilization Time: 100 minutes (dual task); Reliability Efficiency: 0.50	Alcohol, Amphetamines, Diazepam, Marijuana, Sleep loss	Pilot, Truck Driver, or Professional Driver
<u>Time Estimation</u>  Ability to estimate time lapse or time for two objects to intersect.	2	3-5 minutes	Not reported	Atropine, carbon monoxide	Not reported
<u>Perceptual Speed</u>  Ability to quickly and accurately compare pairs of objects or to identify figures or objects in a complex set of objects.	30	1-50 minutes	Stabilization Time: 16-121 minutes, median = 27 minutes; Stability: 0.55-0.90, median = 0.85; Reliability Efficiency: 0.52-0.93, median = 0.85	Alcohol, Antihistamines, Halon, Motion sickness drugs, Chemotherapy, Altitude	Jobs that require employee to react rapidly to visual and audio cues
<u>Memory</u>  Ability to retain and recall material over very short or long periods.	15	1-75 minutes	Stabilization Time: 15-20 minutes, median = 18 minutes; Stability: 0.67-0.98, median = 0.75; Reliability Efficiency: 0.65-0.84, median = 0.75	Alcohol, Marijuana	Not reported
<u>Perception</u>  Ability to visualize objects to dissemble from other material or visualize complete information from partial information.	6	30 seconds-3 minutes	Stabilization Time: 1.5-28 minutes, median = 3.5 minutes; Reliability Efficiency: 0.80-0.98, median = 0.96	Atropine, Depressants, LSD, Nicotine, Scopolamine, Age, Psychiatric disorders	Not reported
<u>Spatial Ability</u>  Ability to manipulate or transform the image of spatial patterns into other arrangements, involving two or three-dimensional space.	9	1-30 minutes	Stabilization Time: 14 minutes (1 test); Stability: 0.70-0.97, median = 0.88; Reliability Efficiency (2 tests): 0.79-0.93	Alcohol, Antihistamines, Halon, Motion sickness drugs, Chemotherapy	Not reported

## Performance-Based Tests

**Table 10.1: Summary of Performance-Based Test Information (Continued)**

Description of Process/Ability	# of Tests	Administration Time	Reliability	Sensitivity	Validity
<u>Attention/Divided Attention</u>  Ability to concentrate and identify visual and auditory stimuli; to filter out irrelevant stimulus input; or to perform two tasks simultaneously.	10	1-30 minutes	Stabilization Time: 100 minutes (dual task); Reliability Efficiency: 0.50	Alcohol, Amphetamines, Diazepam, Marijuana, Sleep loss	Pilot, Truck Driver, or Professional Driver
<u>Reasoning</u>  Ability to select and organize relevant information for the solution of a problem.	10	1-22.5 minutes	Stabilization Time: 18 minutes (1 test); Stability: 0.59-0.96, median = 0.85; Reliability Efficiency: 0.89-0.93	Alcohol, Antihistamines, Halon, Motion sickness drugs, Chemotherapy, Environmental stress, Road pollutants, Sleep loss	Not reported
<u>Fluency</u>  Ability to produce words that fit one or more structured, phonetic, or orthographic restrictions.	4	3-8 minutes	Stabilization Time: 1-72 minutes, median = 28 minutes; Reliability Efficiency: 0.66-0.79, median = 0.70	Not reported	Not reported
<u>Number Facility</u>  Ability to perform basic arithmetic operations with speed and accuracy.	7	1-90 minutes	Stabilization Time: 27-90 minutes, median = 48 minutes; Reliability Efficiency: 0.83-0.90, median = 0.87	Alcohol, Atropine, Ditran, Jet fuel, Methychloride, Scopolamine, Sleep loss	Not reported
<u>Reaction Time</u>  Ability to make simple responses upon presentation of a stimulus or stimulus set.	10	1-60 minutes	Stabilization Time: 35-50 minutes; Stability: 0.71 (1 test); Reliability Efficiency: 0.51-0.80, median = 0.69	Alcohol, Antihistamines, General anesthesia, Halon, Toluene, Chemotherapy, Sleep loss, Sustained physical effort	Not reported
<u>Psychomotor</u>  This includes all abilities involved in coordination, manipulation, repetition and precise body or limb movements; all involve motor control and movement.	37	1-25 minutes	Stabilization Time: 1-135 minutes, median = 42 minutes; Stability: 0.73-0.99, median = 0.87; Reliability Efficiency: 0.60-0.97, median = 0.90	Alcohol, Carbon monoxide, Marijuana, Motion sickness drugs, Secobarbital, Altitude, Chemotherapy	Validated for intoxicated drivers

Table 10.1: Summary of Performance-Based Test Information (Continued)

Description of Process/Ability	# of Tests	Administration Time	Reliability	Sensitivity	Validity
<u>Attention/Divided Attention</u>  Ability to concentrate and identify visual and auditory stimuli; to filter out irrelevant stimulus input; or to perform two tasks simultaneously.	10	1-30 minutes	Stabilization Time: 100 minutes (dual task); Reliability Efficiency: 0.50	Alcohol, Amphetamines, Diazepam, Marijuana, Sleep loss	Pilot, Truck Driver, or Professional Driver
<u>Vision/Visual Acuity</u>  Normal visual acuity, including peripheral vision, depth perception, light adaptation.	13	3 minutes to 1.5 hours	Stabilization Time: 1 minute (5 tests); Reliability Efficiency: 0.51-0.75; median = 0.53	Alcohol, Marijuana for some tests; others not significant. Diazepam, Amphetamine	Not reported
<u>Ocular Movement</u>  Tests are designed to measure eye movements (gaze nystagmus).	19	8-14 minutes	For most tests—not reported. Stability (1 test): 0.60-0.99	Alcohol, Antihistamines, Barbiturates, Cocaine, Diazepam, Hydrochloride, Inhalants, Ketamine hydrochloride, Methadone, PCP; no significant effects for sleep loss	Validated for intoxicated drivers

useful for jobs that require one to respond rapidly to visual or audio cues (e.g., air traffic controllers).

### Memory

Memory tests assess one's ability to retain and recall information over very short or long periods. Fifteen memory tests were identified in the literature. These tests require from one minute to 75 minutes to administer, with a median test length of 2.5 minutes. Stabilization time is generally lower for these tests (15 to 20 minutes), relative to other cognitive ability tests. Stability test estimates range from 0.67 to 0.98 (median = 0.75). Reliability efficiency ranges from 0.65 to 0.84, with a median of 0.75. Memory tests have been shown to be sensitive to impairments caused by alcohol and marijuana. No criterion-related validity information was provided for these tests, nor did researchers identify occupations for

which these tests would be useful performance-based test candidates.

### Perception

Perception tests assess the ability to visualize objects, to disassemble objects from complex material, or to visualize complete information from partial information. Administration times for the six tests described in the literature are of short duration, ranging from 30 seconds to three minutes. Stabilization time is also short (1.5 to 28 minutes, with a median of 3.5 minutes). Reliability efficiency ranges from 0.80 to 0.98, with a median of 0.96. Tests of perception have been shown to detect impairments caused by drug and chemical use (nicotine, depressants, LSD, scopolamine, atropine), and impairment due to age and psychiatric disturbance. No job-related validity information was reported.

## Performance-Based Tests

### Spatial Ability

Spatial ability tests measure the ability to manipulate or transform the image of spatial patterns into other arrangements in two- or three-dimensional space. Nine spatial ability tests were evaluated. These require generally short administration times, ranging from one to 30 minutes, with a median test length of 3.5 minutes. Stabilization time for one test was reported at 14 minutes. Test stability was reported as ranging from 0.70 to 0.97 (median = 0.88). Reliability efficiency was reported for two tests; these values were 0.79 and 0.93. Scores on spatial ability tests are sensitive to impairments caused by alcohol, antihistamines, halon, motion sickness drugs, and chemotherapy. No criterion-related validity information was reported for these tests.

### Reasoning

Reasoning tests are designed to assess the ability to select and organize relevant information to solve problems. A total of 10 reasoning tests were identified in the literature. These generally involve short administration times (1 to 22.5 minutes, with a median test length of 5.5 minutes). Stabilization time for one reasoning test was estimated at 18 minutes. Test stability was estimated at 0.59 to 0.96, with a median value of 0.85. Reliability efficiency estimates range from 0.89 to 0.93. Research has shown that reasoning test scores are influenced by alcohol, antihistamines, halon, motion sickness drugs, chemotherapy, environmental stress, road pollutants, and sleep loss. No job-related validity information was reported for these tests.

### Fluency

Fluency tests measure the ability to produce words that fit one or more structured, phonetic, or orthographic restrictions. Four fluency tests were described in the literature; these tests require from three to eight minutes to administer, with three of the four tests requiring an eight-minute administration time. Stabilization time ranges from one to 72 minutes (median value is 28 minutes). Reliability efficiency ranges from 0.66 to 0.79, with a median value of 0.70. Although these tests have been examined as potential FFD tests, researchers have not identified the types of conditions or drugs that influence performance on these tests. Further, job-related validity information was not reported for these tests.

### Number Facility

Number facility tests assess the ability to perform basic arithmetic operations quickly and accurately. Seven number facility tests are summarized in the literature review. These tests generally require very short test administration times, ranging from one minute to 90 minutes, with a median test length of four minutes. Stabilization time ranges from 27 to 90 minutes (median value is 48 minutes). Reliability efficiency ranges from 0.83 to 0.90, with a median value of 0.87. Scores on number facility tests have been shown to be influenced by the use of or exposure to alcohol, atropine, ditran, jet fuel, methylchloride, scopolamine, and sleep loss. Job-related validity information was not reported for these tests.

### Reaction Time

Reaction time tests assess the ability to make simple responses upon presentation of a stimulus or stimulus set. A total of 10 tests is described in the literature review. Administration time varies from one minute to 60 minutes, with a median test length of five minutes. Stabilization time for reaction time tests ranges from 35 to 50 minutes. Stability reported for only one test is 0.71. Reliability efficiency values range from 0.51 to 0.80, with a median value of 0.69. Reaction time test scores have been shown to be influenced by alcohol, antihistamines, general anesthesia, halon, toluene, chemotherapy, sleep loss, and sustained physical effort. Job-related validity information was not reported for reaction time tests.

#### 10.2.1.2: Psychomotor Tests

Psychomotor tests include tests of all abilities involved in coordinative, manipulative, repetitive, and precise body or limb movements. Psychomotor tests assess a broad range of motor control and movement capabilities, including eye-hand coordination, coordination and balance, tracking, and manual dexterity. A total of 37 psychomotor tests was identified in the literature review. Administration times for these tests range from one to 25 minutes, with a median test length of two minutes. Stabilization time ranges from one minute to 135 minutes, with a median value of 42 minutes. Stability estimates for psychomotor tests range from 0.73 to 0.99 (median value is 0.87). Reliability efficiency estimates range from 0.60 to 0.97,

with a median value of 0.90. Scores on psychomotor tests have been shown to be influenced by alcohol, carbon monoxide, marijuana, motion sickness drugs, altitude, secobarbital, and chemotherapy. These tests have been validated as effective for detecting intoxicated drivers.

#### 10.2.1.3: Physiological Tests

Physiological tests include tests involving vision and visual acuity, as well as tests for ocular movement. These tests are discussed below.

##### Vision and Visual Acuity

Vision and visual acuity measures assess normal visual acuity processes, such as peripheral vision, depth perception, and light adaptation. A total of 13 vision and visual acuity tests were identified in the literature review. These require from three minutes to 1.5 hours to administer with a median time of three minutes. Stabilization time is one minute for five tests. Reliability efficiency ranges from 0.51 to 0.75, with a median value of 0.53. Scores on these tests have been shown to be influenced by alcohol, marijuana, diazepam, and amphetamines for some visual acuity tests (e.g., visual autokinesis and dark adaptation), but this is not true for all vision and visual acuity tests (e.g., vertical phoria). No job-related validity information was presented for these tests.

##### Ocular Movement

Ocular movement tests are designed to measure eye movements (gaze nystagmus). A total of 19 ocular movement tests was reviewed. Test administration times range from eight to 14 minutes. For most ocular movement tests, reliability information was seldom reported. Stability reported for one test ranges from 0.60 to 0.99 (EM2 test; see Appendix G). Scores on these tests are sensitive to impairments caused by drug and chemical intake, such as alcohol, antihistamines, barbiturates, cocaine, diazepam, hydrochloride, inhalants, ketamine hydrochloride, methadone, and PCP. Research has not, however, shown significant effects for sleep loss. These tests have been validated for detecting intoxicated drivers.

#### 10.2.1.4: Summary of Performance-Based Testing Literature

Overall, researchers have investigated a large number of process and ability tests that may serve as performance-based tests in the work place. As indicated in the above

discussion, much of the research has focused on demonstrating that scores on these tests are influenced by drug intake, environmentally-induced chemical intake, and physical conditions. Administration time for many of the tests is quite low (see Appendix G for specific times), although in some instances administration times are prohibitively long. The shortest administration times appear for time estimation, perception, spatial ability, fluency, and ocular movement tests. Stabilization time represents the amount of practice that is required for scores to stabilize and to eliminate practice effects. For most process and ability tests, practice influences test scores. The smallest practice effects appear for memory, perception, spatial ability, reasoning, and vision and visual acuity tests. Test stability and reliability efficiency estimates are generally high for most tests. The highest values appear for perceptual speed, perception, spatial ability, reasoning, number facility, and psychomotor ability tests. For most process and ability tests, job-related validity information was not reported. Thus, this is one area in which more research is needed to substantiate claims that such tests can identify impairment in abilities and processes required to perform critical tasks in specific jobs and occupations.

#### 10.2.2: Performance-Based Tests Available for Implementation

At the present time, vendors are marketing performance-based tests for use in the work place. To assess the status of performance-based tests for application to the nuclear industry, we contacted 10 vendors from around the country and requested materials describing their performance-based tests. From these requests, we received marketing materials from seven vendors. Materials from a sample of five vendors are summarized in this section. Tests offered by these five vendors were selected because along with the tests, vendors also supplied details about test development, procedures for test administration in a FFD program, or research that support the tests in FFD programs (see Appendix G).

Vendors' marketing materials generally describe the type of test, provide some research references documenting support for their test, and describe the types of occupations in which the test is currently being used, if available. It is important to note that the marketing materials vary in terms of the types and amount of information presented for the test. Some materials simply promote the performance-based test without providing specific references or sources to support their claims. In general, most vendors refrained from identifying specific companies or organizations that currently use or are

## Performance-Based Tests

planning to use their performance-based tests. This section includes a description of the test or tests each vendor is marketing and information supporting the use of these tests (i.e., reliability, sensitivity, and validity). These vendor summaries are followed by an overall summary.

### 10.2.2.1: SEDI Corporation

SEDI Corporation currently markets a computer-administered test that assesses divided attention. This *Simulated Evaluation of Drug Impairment (SEDI)* test requires a three-minute administration period. The vendor recommends using this test to determine if the employee is impaired at the start of the work shift and recommends using this test in conjunction with drug testing (National Institute of Drug Abuse, 1991). The SEDI Corporation states that this test can be used to detect impairment from alcohol, illegal drugs, stimulants, and antihistamines. An annotated bibliography provides research references documenting these claims. This company also provides normative data against which employees' scores can be compared (SEDI Corp., 1991/1992).

### 10.2.2.2: Essex Corporation

The Essex Corporation has constructed a battery of cognitive and psychomotor ability tests that can be used in various combinations to assess FFD. The *Delta: Human Performance Test and Measurement System Battery* includes tests designed to observe and quantify cognitive, perceptual, and psychomotor changes in human performance due to physical, chemical, and psychological stressors (Banisch, 1992). The battery includes 15 different tests that assess perceptual speed, short-term memory, reasoning, number facility, spatial ability, reaction time, and psychomotor ability. According to the developers, there has been over 20 years of research support using these tests. In fact, as reviewed in a previous report (Moore et al., 1989), these tests were primarily derived from those tests evaluated in the Performance Evaluation Tests for Environmental Research (PETER) Program (Bittner et al., 1986) and implemented into the Automated Portable Test System (APTS) Battery (Bittner et al., 1985). Some tests were also drawn from the Unified Tri-Services Performance Assessment Battery (UTS-PAB) (Englund et al., 1987). All tests can be administered using a desk top or portable computer; test performance data are available immediately upon completion of the test or test battery. Developers describe these tests as sensitive, stable, and reliable and requiring only a few minutes to administer (Essex, 1992).

### 10.2.2.3: Nova Technology Incorporated

Nova Technology Incorporated (NTI) is currently marketing another test battery loosely based on the UTS-PAB (Englund et al., 1987). This battery, *NOVASCAN: Employee Performance Evaluation System*, contains three computer-administered tests that assess spatial ability, attention, and psychomotor ability (tracking). Test developers reported that these tests are based on tests developed for UTS-PAB; however, in many cases the tests differ markedly in appearance. Each test requires from three to five minutes to administer, not including equipment setup times. For industrial applications, NTI recommends selecting one or more tests based on job performance requirements. Test developers stated that battery test scores have been shown to be influenced by alcohol use; illicit, prescription, and over-the-counter drug use; illnesses; stress; sleep loss; and personal or family problems (Nova, 1993). An employee's test scores are available immediately upon completing the test, and scores are compared with his or her own previous scores over the past 30 trials (i.e., baseline data are based on 30 administrations of each test). The computer program, after detecting a potentially impaired employee, provides a broad assessment of the potential problem (e.g., a certain pattern of test scores has been shown to correspond with fatigue, while another pattern may correspond with alcohol or drug use). Developers also report that these tests can be used to identify impairment for employees required to drive trucks or pilot airplanes and for industrial applications, such as shipbuilding, heavy construction, and process industries.

### 10.2.2.4: Performance Factors Incorporated

Performance Factors Incorporated (PFI) is marketing a computer-administered psychomotor ability test that assesses eye-hand coordination in a tracking test (Hamilton, 1991). The test is also referred to as the critical tracking task (see Appendix G for details of this test). This test is currently being used in industries with employees performing in a wide range of jobs, such as machine tooling, driving tour buses, handling poisonous gases, working with heavy equipment, as well as in highly hazardous process control industries. Employee data collected from these industries indicate that most failures are not due to drug use, but instead are attributed to fatigue and illness.

### 10.2.2.5: Eye Dynamics, Inc.

Eye Dynamics, Inc. (EDI)<sup>15</sup> currently markets a test to assess eye movements, the *EPS-100* (ISCAN Eye

Dynamics Research Laboratory, 1992). This computer-administered test assesses ocular movement using a video monitor; this test is similar to those that fall within the ocular movement category described in Table 10.1. Research supporting the use of this test comes from a variety of sources, including the Drug Recognition Expert Program sponsored by the National Highway Traffic Safety Administration (NHTSA) (see Durbin et al., 1991) and studies sponsored by the National Institute of Drug Abuse (NIDA). Test scores have been shown to be influenced by buprenorphine, hydrochloride, cocaine, methamphetamine, marijuana, alcohol, and fatigue. An employee's test scores are compared with scores obtained under "clean" conditions. Test developers state that this test is designed to be used in conjunction with other drug testing methods, such as urinalysis. For example, this vendor states that the EPS-100 test can be administered on a daily basis before work begins, "for cause," after an accident, or on a random basis. If an employee fails the EPS-100 test, the organization could then "proceed with a urine-based screening and confirming test, as well as appropriate medical or EAP evaluations to determine the reason the subject failed the EPS-100 test" (DDSI, 1992, p. 9).

#### 10.2.2.6: Summary of Currently Available Performance-Based Tests

Based on input from vendors, it is clear that performance-based testing devices are available for evaluating employee fitness for duty. The tests currently marketed have been, for the most part, studied under different circumstances to assess the influence of a variety of drugs and conditions on test performance. Data presented by vendors indicate that performance-based tests can and are currently used to detect performance impairments in the work place. Indeed, many of the vendors describe their tests as serving as the first step in a drug testing program, wherein a failure on the performance-based test is used to indicate if urinalysis test should be conducted to complete the evaluation of fitness.

### 10.3: Performance-Based Test Accuracy

As described in Section 10.1.2.4, a critical feature of tests designed to assess performance impairment is test accuracy. Therefore, when determining if a performance-based test can be of practical use in a FFD program, it is important to ask if the test is sufficiently accurate for this purpose. Within the body of literature reviewed for this

chapter, test accuracy is almost always described in qualitative terms (e.g., tests detect drug or alcohol use). Seldom do researchers report test accuracy in quantitative terms (e.g., tests are 95 percent accurate in detecting BAC levels of 0.04 or above). Thus, potential users may have a difficult time comparing the costs and benefits of performance-based tests over the more traditional FFD tests.

This section addresses test accuracy by demonstrating procedures for providing a quantitative estimate of test accuracy for a single performance-based test. The purpose is to demonstrate that a quantitative accuracy estimate can be computed for performance-based tests. Thus, this section addresses test accuracy by:

- defining test accuracy in terms that are suitable for determining test accuracy in quantitative terms (Section 10.3.1),
- using this definition to select a performance-based test for which detection accuracy can be calculated (Section 10.3.2), and
- quantitatively evaluating the practical utility of the selected test (Section 10.3.3).

#### 10.3.1: Test Accuracy in Quantitative Terms

Test accuracy is a characteristic that reflects the probabilities of both correct and incorrect detections (Elgin, 1971; Bittner, Kennedy and McCauley, 1980).<sup>16</sup> Accuracy can be defined in an equation format, with the likelihood of correctly detecting an impaired person defined as the *probability of correct-detection* ( $p_d$ ). This probability, ( $p_d$ ), corresponds to the probability of correctly identifying an impaired worker (e.g., identifying impairment in a worker with a known source of impairment such as a BAC of 0.04). The likelihood of incorrectly identifying the unimpaired worker as impaired is defined as a *probability of a false alarm* ( $p_f$ ). Ideally, a test increases the probability of correct-detection ( $p_d$ ) while decreasing the probability of false alarm ( $p_f$ ). However, these two probabilities,  $p_d$  and  $p_f$ , are potentially interdependent. If the score that determines pass or failure on a test (a cutoff score) is set very high (e.g., almost everyone fails), then the likelihood of correct detection is increased. The likelihood of false alarm, however, is also increased. Similarly, a low cutoff score (i.e., everyone passes) reduces the probability of a false alarm but also reduces the probability of correct detection. To capture this interdependence, and also to reflect

differences in test accuracy, a mathematical measure of test accuracy ( $d'$ ) has been developed:

$$d' = Z(p_d) - Z(p_f) \quad (\text{Equation 10.1})$$

where  $Z(p)$  is the function relating cumulative unit-normal probabilities (e.g.,  $p_d$  and  $p_f$ ) to their associated standard score values. This function is particularly useful because it provides a means of determining the magnitude of the probability of correct detection ( $p_d$ ) when  $d'$  is known and false alarms are restricted to a reasonable level (e.g.,  $p_f = 1/200$ ). The values of  $d'$  across a number of performance tests also permit the direct comparison of their sensitivities at a level of impairment equivalent to a specified level of an impairing factor (e.g., 0.04 BAC). Bittner and his colleagues (1980) previously employed the  $d'$  metric in a function relating drug performance effects to dose level.

### 10.3.2: Selecting a Test for Accuracy Calculations

This section describes the procedures used to select a performance-based test that is highly accurate ( $d'$ ) in detecting alcohol-induced performance impairments at differing BACs.<sup>17</sup> To assist in selecting such a test, it was first necessary to evaluate the published test data in terms of  $d'$  (where possible), and then to select from available candidates the test(s) with the largest functional values over a range of BACs (ranging from 0.00 to 0.15 BAC). A psychomotor test, two-hand tapping, was selected for accuracy assessment. This test was selected for several reasons: 1) it is a relatively short test with high reliability (see Appendix G); 2) this type of test has been shown to be predictive of job performance in a variety of occupations (e.g., technical, protective service, maintenance; McHenry and Rose, 1987); and 3) scores on this test have been shown to be influenced by moderate BAC levels within the range of concern of safety-sensitive industries. Further, although this test is classified as a psychomotor test, it also has been shown to assess cognitive elements.

#### 10.3.2.1: Computing Test Accuracy ( $d'$ )

The process of computing test accuracy values from published data proved particularly challenging because test accuracy was expressed in the literature in a wide variety of forms. Sometimes, the published reports provided false-alarm (false-positive) and correct-detection data for one or more BACs and, in those cases, the formulation discussed earlier could be used [ $d' = Z(p_d) -$

$Z(p_f)$ ]. For accuracy data for other types of tests, either a more complex analysis would be required to derive relevant values or no suitable means was readily available. One reported analysis resulted in the derivation of the functional relationship between  $d'$  and BAC level from the published  $F$ -statistic:

$$d'(j) = x_j [Q(K-1)(F-1)/[N \sum_j (x_j - M_x)^2]]^{1/2} \quad (\text{Equation 10.2})$$

where  $d'(j)$  is the accuracy for BAC level  $x_j$ ,  $K$  and  $N$  are the number of treatment (BAC) levels and subjects,  $M_x$  is the arithmetic mean of the  $x_j$ s, and  $Q$  is the ratio of a desired test time to that upon which the  $F$  is based.<sup>18</sup>

Applying this functional relationship to the selected test (the 20-second two-handtapping test; Kennedy, Turnage and Dunlap, 1992) with  $F = 15.65$ ,  $K = 7$  (see Endnote 19) and  $N = 18$ ,  $\sum_j (x_j - M_x)^2 = 0.05625$  and sequentially  $Q = 9$  (for a test length of 180 seconds) and  $Q = 13.5$  (for a test length of 270 seconds), a pair of relationships result:

$$d'(j) = 27.93x_j \quad (\text{Equation 10.3})$$

$$d'(j) = 34.21x_j \quad (\text{Equation 10.4})$$

where  $x_j$  = BAC level. These relationships are applicable to two-hand tapping tests of different lengths, where Equation 10.3 applies to a three-minute test (180 seconds) and Equation 10.4 applies to a four-and-one-half minute test (270 seconds). These formulae are employed and their implications are explored in the next section.

### 10.3.3: Accuracy Evaluation of the Selected Test

Evaluations of the detection accuracies for the two-hand tapping test ( $p_d$ ) were computed for two false-positive rates ( $p_f = 0.005$  and  $0.010$ ) and two test lengths (180 seconds and 270 seconds) after substituting the appropriate equation (Equation 10.3 or Equation 10.4) into Equation 10.1 and solving for  $p_d$ . Results from these calculations are presented in Table 10.2. From this table, it can be seen that test score accuracy for all test combinations would be good to superb when the BAC is 0.15 ( $p_d = 0.947$  to  $0.998$ ; where 1.00 indicates perfect accuracy). But, accuracy detection drops precipitously at BAC 0.100 ( $p_d = 0.586$  and  $0.680$  for a 180 second test, with higher values  $0.801$  and  $0.863$  for 270 second tests). With BACs at 0.05 or lower, detection accuracy is

**Table 10.2: Evaluation of Two-Hand Tapping Test Accuracy: Proportion Detected of BAC Equivalent Impairments for Two Test Lengths and Two False Alarm Rates**

BAC	Proportion Detected			
	180 Sec. (3 Min.) Test Length		270 Sec. (4.5 Min.) Test Length	
	False Alarm Rate 0.005	False Alarm Rate 0.010	False Alarm Rate 0.005	False Alarm Rate 0.010
0.150	0.947	0.967	0.995	0.998
0.100	0.586	0.680	0.801	0.863
0.075	0.315	0.409	0.496	0.595
0.050	0.119	0.176	0.194	0.269
0.025	0.030	0.052	0.043	0.071
0.000	0.005	0.010	0.005	0.010

unacceptable ( $p_d = 0.269$  and below). Results presented in Table 10.2 indicate that this performance-based FFD test would accurately detect higher levels of impairment represented by BACs of 0.15 relatively well, but would not be very accurate for BACs of 0.05 and below.

One final note concerning the calculation of detection accuracy for performance-based tests. The procedures and formulae described in this section represent another critical variable that should be available for any performance-based test used in a FFD program.

Accuracy calculations provide support for using these tests in a multiple-hurdle FFD program (e.g., performance-based tests administered, if results suggest impairment then urinalysis or other FFD tests are administered). Even though the literature provides qualitative information about test accuracy, quantitative accuracy information will provide greater assurances concerning the merits of performance-based tests.

## 10.4: Summary and Implications for the Nuclear Industry

The first part of this section contains a summary of the performance-based test information provided in this chapter (Section 10.4.1). The final part of this section identifies issues and questions that should be addressed before performance-based tests can be implemented in the commercial nuclear power industry (Section 10.4.2).

### 10.4.1: Overall Summary of the Status of Performance-Based Tests

This chapter has systematically considered the status of performance-based tests with regard to their potential FFD use in the nuclear industry. Setting the stage for later sections, the desired psychometric characteristics for performance-based tests were delineated in Section 10.1, which pointed out that performance-based tests differ from drug and alcohol tests in three important ways:

- Performance-based tests have been shown to identify impairment caused by both drug and non-drug agents (e.g., sleep loss, stress, exposure to chemicals, etc.).
- Performance-based tests are designed to detect impairment, but not its cause, and
- Performance-based tests consequently provide a broad coverage of agents and sources that influence job performance and public safety.

A summary and discussion of the current literature regarding the feasibility of performance-based tests was provided in the second section of this chapter (Section 10.2). The initial examination of the performance-based testing literature (Section 10.2.1) revealed the following:

- Research on performance-based tests includes studies of their suitability for and sensitivity to drug use,

## Performance-Based Tests

environmentally-induced chemical intake, and physical conditions.

- Administration times for many of the tests are in the range of one minute to five minutes, which would be practical in industrial settings.
- Amount of practice required for scores to stabilize is relatively short in many cases, and reliability efficiency estimates are generally high.
- More research is needed to substantiate claims that such tests can identify performance impairment in specific jobs and occupations, and that test accuracy is sufficiently high to warrant use as the first test in a multiple-tiered FFD program.

The subsequent review of vendor literature (10.2.2) further examined the current practical utility of performance-based testing. Vendors present the following evidence:

- Performance-based testing is becoming a more commercially recognized procedure for evaluating employee fitness for duty.
- Performance-based tests can and are currently used to detect performance impairments in the work place.

These findings indicate that even with some questions yet unanswered by research (e.g., criterion-related validity), some U.S. industries are investing in performance-based FFD programs to assist in identifying worker impairment that may contribute to reduced safety on the job.

Finally, Section 10.3 examined the question of "accuracy" given the encouraging evidence found in the academic and commercial literature reviews. Specifically, the question was, "Is any performance-based FFD test sufficiently accurate to be of practical use in the work place?" Statistical analysis of one selected test indicated the following:

- One performance-based FFD test was shown to be accurate in identifying high levels of alcohol impairment.
- Before employing any performance-based test it is useful to estimate test accuracy by using data from well-controlled experiments. These data may be

obtained from vendors or by conducting one's own studies.

### 10.4.2: Implementing Performance-Based Tests in Industry

Conclusions about test accuracy, together with the earlier literature review results, indicate that further evaluation of performance-based testing as an FFD assessment tool may be warranted. Before such tests are implemented in any industry, it is important to consider some of the factors that have not been addressed fully by this research. These factors include: 1) the process for demonstrating the link between performance-based tests and job performance; 2) procedures for administering such tests; and 3) the implications of and procedures for next steps after an employee has failed a performance-based test. Each of these factors is examined in greater detail below.

#### 10.4.2.1: Linking Performance-Based Tests to Job Performance Requirements

Linking performance-based tests to job performance requirements is a process that involves identifying the jobs or occupations involved in performance-based testing, collecting or reviewing job analysis data to identify the abilities and skills required for performance in each job, and conducting a criterion-related validity study to demonstrate the relationship between performance-based test scores and job performance. Ideally, each industry should establish a direct link between performance on a specific performance-based test and performance in a work situation. This particular model is currently being used to identify and evaluate the utility of administering performance-based tests in the transit industry (Bittner et al., 1993).

Another issue to consider in selecting tests for a performance-based test battery is compliance with the *Americans with Disabilities Act of 1990*, or ADA (see *Americans with Disabilities Act*, 1990). Conducting a formal study to identify the abilities and skills required for effective job performance and using this information to select performance-based tests (i.e., procedures described above) should result in testing procedures that are in compliance with ADA. It is necessary to consider ADA in selecting and implementing a performance-based test battery to avoid the situation in which performance-based tests systematically define disabled employees as impaired, when, in fact, they are not impaired and can perform the job.

#### **10.4.2.2: Procedures for Implementing a Performance-Based Test Battery**

Once a test or test battery has been identified for performance-based testing, it is important to define the procedures related to administering the test. For example, the performance-based test battery can be administered on a daily basis (e.g., each time an employee appears for work), or the tests may be administered randomly, for cause, or following an incident or accident.

The organization should provide an explicit rationale for using the tests and designing administration procedures. This rationale can include assumptions about errors on the job (e.g., small errors that could have substantial impact on public safety), behaviors or conditions the performance-based tests will detect, and costs and benefits associated with different administration procedures. Costs can be defined in terms of materials and equipment required for testing, employee time required to establish baseline performance, and employee time required to complete the test or test battery on the designated schedule. Benefits are defined by the savings associated with avoiding serious human errors caused by impaired employees and programs or procedures in place to assist employees in altering and improving their lifestyles.

#### **10.4.2.3: Implications of a Test Failure**

A final consideration in implementing performance-based tests involves defining the procedures for the next steps after an employee has been defined as impaired. As noted earlier, performance-based tests detect impairment, but do not determine the cause of impairment. Hence, when impairment is ascertained using performance-based test scores, an employer may either take action to reduce the impact of the impairment (e.g., reassign the employee or send the employee home) or may perform further tests to determine the cause of impairment (e.g., medical evaluation, a breath test for alcohol, and/or urine testing

for drugs). In some cases it will not be possible to identify the cause of a performance-based test failure. For example, some performance-based tests have been shown to be sensitive to intake of chemicals from the environment and hence could be caused by the workplace itself. In this case, the reasons for the impairment may remain unknown and the solution (modifying the employee's work conditions) may not be apparent. Hence, performance-based test failures should not result in negative consequences for workers without determining the cause of the impairment.

#### **10.4.2.4: Conclusions**

At the present time, performance-based tests are available for implementation in the work place. These tests are reliable, as well as sensitive to drug use, environmental exposure to chemicals, and problems associated with sleep loss and stress. Further, an analysis of a specific selected test presented in this chapter indicates that such tests can be highly accurate in detecting significant impairment, such as that caused by a BAC of 0.15. One problem associated with performance-based tests is the lack of quantitative criterion-related validity information for any of the performance-based tests summarized in the literature. Further studies linking performance-based test results to impairment are needed.

Information supplied by vendors suggests that some performance-based tests are currently being used in the work place. At the present time, specific information about implementation procedures and procedures for handling failures is not available. Before performance-based tests can be effectively used in further industry applications, it is important to collect information about the procedures that other industries are using or have tried. It is also important to understand the merits of performance-based tests versus urinalysis in terms of employee acceptability, avoidance of incidents and accidents in the workplace, and organizational outcomes.

## Performance-Based Tests

## Endnotes

1. Jody Toquam, Ph.D., Battelle Seattle Research Center  
Alvah Bittner, Ph.D., Battelle Seattle Research Center
2. See respectively Gilliland and Schelegel (1993) and Bittner et al. (1993).
3. When batteries of tests (or measures) are used, some information regarding the nature of impairment may be contained in the pattern of impairments.
4. The terms test reliability and test validity appear to be interchangeable. Although both indicate qualities of a test, each focuses on different characteristic of a test. Reliability represents consistency or stability of test information (test scores). That is, a highly reliable test will produce the same or nearly the same information for a given person if administered two or more times. Validity indicates whether a test can be used to predict a criterion which might consist of job performance. In this case, a test is valid if it is shown to correlate significantly with a measure of job performance (e.g., performance ratings by a supervisor, number of errors in a production line).
5. Although it is true that values for the correlation coefficient range from -1.00 to +1.00, reliability estimates fall within the range of 0.00 to +1.00. Negative reliability estimates would suggest severe problems with a test.
6. The Spearman-Brown formula was designed to estimate the change in reliability for a given a test that is increased or decreased in length. The formula used to estimate the new reliability is

$$r_{kk} = \frac{kr_{11}}{1 + (k-1)r_{11}}$$

Where  $r_{kk}$  = estimated reliability;  $k$  is the factor by which test length is changed [e.g., double in length ( $k=2$ ) or reduced by half ( $k=.5$ )];  $r_{11}$  is the current known reliability estimate.

For example, a 20-minute test with a known reliability of ( $r_{11}$ ) .88, may be reduced by half ( $k=.5$ ) yielding a 10-minute test. Using the Spearman-Brown formula, the estimated reliability ( $r_{kk}$ ) of the reduced test is .79.

7. Sensitivity of a test is defined as the ability to detect impairment regardless of the source of that impairment. This is distinguished from test specificity in which the specific source of the impairment can be identified (e.g. breath analysis to detect BAC; urinalysis to detect drug metabolites).
8. The common decision rule specifies that an individual's current score must be three standard deviations above or below his or her baseline mean score to demonstrate impairment. As an example, assume that Subject A has completed a test on 15 previous occasions. Overall, Subject A has obtained a mean score of 75 (out of 90 items) and a standard deviation of 8; this represents Subject A's baseline performance. On occasion 16, Subject A completes the same test and obtains a score of 51. This score is much lower than the baseline score (mean = 75). Impairment is determined by using a formula that includes Subject A's baseline performance (75), standard deviation of performance (8), and current test score (51). The formula is as follows:

$$(Baseline Score - Current Score) / Standard Deviation$$

This translates into:  $(75-51)/8 = 24/8 = 3$  standard deviations.

Thus, Subject A's current test performance is very different from the baseline (3 standard deviations below the baseline performance), and we would conclude that Subject A is impaired.

9. Although it is often expected that impairment will produce lower than normal test scores, it is possible for some drugs to produce scores that are higher than normal (e.g., the effect of amphetamines on reaction time scores).
10. Criterion-related validity is described as concurrent validity (tests and job performance scores are administered to subjects at the same time) or predictive validity (tests are first administered to subjects and later, job performance scores are obtained from the same group of subjects). Although the predictive validity method is often preferred, the concurrent validity method is acceptable (Barrett, Phillips and Alexander, 1981).
11. Procedures for validating a test include conducting or reviewing current job analyses to identify the critical job duties and to identify the abilities and skills required to perform the critical job duties. Performance-based tests designed to measure the requisite skills and abilities are identified or developed. Further, results from the job analysis are used to construct measures of job performance. Performance-based tests are administered to a sample of subjects, and job performance scores are also collected for these same subjects. The correlation between test scores and performance scores is computed to provide a quantitative estimate of criterion-related validity. If the correlation between the two sets of scores is statistically significant, then it is appropriate to conclude that the test assesses critical requirements of the job.

12. These errors and accuracy will be described in greater quantitative detail later in Section 10.3.
13. Some portions of these sections were adapted from Bittner et al. (1993).
14. Tests measuring spatial ability, numerical ability, mechanical ability, and reasoning have been shown to be linked to performance of nuclear power plant operators (Dunnette et al., 1982).
15. Eye Dynamics, Inc., is a company formed by two other companies, OculoKinetics and DDSI. The gaze nystagmus test, originally marketed by OculoKinetics, is associated with a large body of research. Appendix G refers to many of the studies conducted using this performance-based test.
16. Test accuracy is here quantitatively defined in terms of the signal detection theory-based metric "d".
17. BAC equivalent impairments were chosen as a standard because alcohol has been used to demonstrate the empirical accuracy of performance-based tests, and its performance effects are better understood than those for any other substance.
18. Equation 2 was derived from the ratio of Model III mean-square expected values (Winer et al., 1991, p. 267), with BAC treatment effects  $y$  assumed to be linear (i.e.,  $y = b - ax$  where  $b$  and  $a$  are constants, and measurement error variance is inversely proportional to testing time; Bittner et al., 1985).
19. The study was designed to assess BAC levels including 0.000, 0.025, 0.050, 0.075, 0.100, and 0.150. In this study, performance was assessed at 0.000 on two occasions, resulting in seven BAC treatment levels.

## References

*Americans with Disabilities Act (ADA) of 1990*, Public Law 101-336, 101st Congress.

American Management Association. (1993). *Workplace Drug Testing and Drug Abuse Policies: Summary of Key Findings*. New York, NY: Author.

Banisch, J. (1992). Letter to A.C. Bittner, Jr., November 16, 1992.

Barrett, G.V., Philips, J.S. and Alexander, R.A. (1981). Concurrent and predictive validity designs: A critical reanalysis. *Journal of Applied Psychology*, 66: 1-6.

Bittner, A.C., Jr., Kennedy, R.S. and McCauley, M.E. (1980). Time estimation: Repeated measures testing and drug effects. *Proceedings of the Seventh Psychology in the DoD Symposium*, pp. 44-450. Colorado Springs, CO: USAF Academy.

Bittner, A.C., Jr., Smith, M., Kennedy, R.S., Staley, C.F., and Harbeson, M.M. (1985). Automated portable test (APT) system: Overview and prospects. *Behavioral Research Methods, Instruments and Computers*, 19(2): 217-221.

Bittner, A.C., Carter, R.C., Kennedy, R.S., Harbeson, M.M., and Krause, M. (1986). Performance evaluation tests for environmental research (PETER): Evaluation of 114 measures. *Perceptual and Motor Skills*, 63: 683-708.

Bittner, A.C., Jr., Toquam, J. L., Forslund, C. and Fleming, T. (1993). *TCRP Project Fitness-for-Duty Testing in the Transit Workplace: Task 2, A Literature Review*. Seattle, WA: Battelle Human Affairs Research Centers.

DDSI. (1992). *EPS-100 Performance System Validation Report*. Torrance, CA: Eye Dynamics, Inc.

Durbin, N., Moore, C., Grant, T., Fleming, T., Hunt, P., Martin, R., Murphy, S., Hauth, J., Wilson, R., Bittner, A., Bramwell, A., Macaulay, J., Olson, J., Terrill, E. and Toquam, J. (1991). *Fitness for Duty in the Nuclear Power Industry: A Review of the First Year of Performance and an Update of the Technical Issues* (NUREG/CR-5784) Seattle, WA: Battelle Human Affairs Research Centers.

Dunnette, M.D., Rosse, R.L., Houston, J.L., Hough, L.M., Toquam, J.L., Lammlein, S.E., King, K.W., Bosshardt, M.J. and Keyes, M.A. (1982). *Development and Validation of an Industry-wide Electric Power Plant Operator Selection System*. Edison Electric Institute.

Elgin, T. (1971). Psychophysics I: Discrimination and Detection. In J.W. King and L.A. Riggs (Eds.), *Experimental Psychology* (3rd ed.), pp. 13-46. San Francisco, CA: Holt Rinehart and Winston.

Englund, C.E., Reeves, D.L., Shingledecker, C.A., Thorne, D.R., Wilson, K.P. and Hegge, F.W. (1987). *Unified Tri-Service Cognitive Performance Assessment Battery (UTC-PAB): I: Design and Specification of the Battery*. (Report No. 87-10). San Diego, CA: Naval Health Research Center.

Essex Corporation. (1992). *Marketing materials*.

Ghiselli, E.E., Campbell, J.P. and Zedeck, S. (1981). *Measurement Theory for the Behavioral Sciences*. San Francisco: W.H. Freeman and Company.

Gilliland, K. and Schelelegel, R.E. (1993). Readiness to Perform Testing: A Critical Analysis of the Concept and Current Practices. Washington, DC: DOT/FAA/AM-93.

Hamilton, J. (June 3, 1991). A video game that tells if employees are fit for work. *Business Week*, 36.

ISCAN Eye Dynamics Research Laboratory. (1992). *ISCAN Eye Movement Monitoring Systems Brochure*. Author.

Jensen, A.R. (1980). *Bias in Mental Testing*. New York: The Free Press.

Kennedy, R.S., Turnage, J.J. and Dunlap, W.P. (1992). The use of dose equivalency as a risk assessment index in behavioral neurotoxicology. *Neurotoxicology and Teratology*, 14: 167-175.

McHenry, J.J. and Rose, S.R. (1988). *Validity and Potential Usefulness of Psychomotor Ability Tests for Personnel Selection and Classification: A Review and Integration of the Literature* (ARI Research Note 88-13). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Moore, C., Barnes, V., Hauth, J., Wilson, R., Fawcett-Long, J., Toquam, J., Baker, K., Wieringa, D., Olson, J. and Christensen, J. (1989). *Fitness for Duty in the Nuclear Power Industry: A Review of Technical Issues (NUREG/CR-5227, Supplement 1)*. Seattle, WA: Battelle Human Affairs Research Centers.

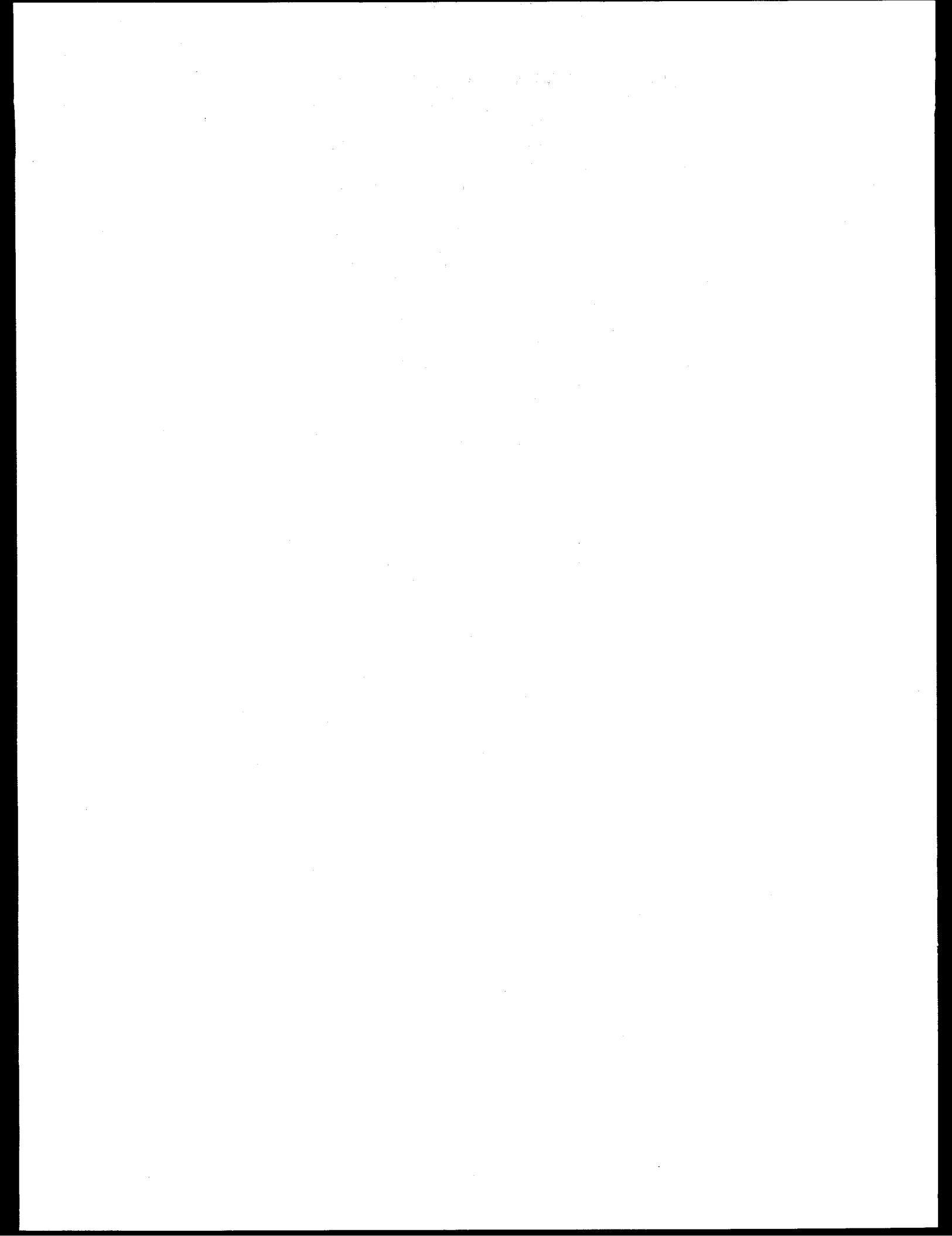
National Institute on Drug Abuse (NIDA). (1991). National Household Survey on Drug Abuse: Population Estimates 1991 (revised November 20, 1992). Washington, DC: U.S. Department of Health and Human Services.

Nova Technology, Inc. (1992). *Technical Description: Alcohol and Drug Analyzer*. Beverly Hills, CA: Author.

SEDI Corporation. (1991-1992). *Marketing Materials for Simulator Evaluation of Drug Impairment (SEDI) Tasks*. Chapel Hill, NC: Profile Associates.

Waldorf, R. (1991). *EPS-100 Performance System Validation Report*. OculoKinetics.

Winer, B.J., Brown, D.R., and Michels, K.M. (1991). *Statistical principles in experimental design* (3rd ed.). San Francisco, CA: McGraw-Hill, Inc.



## APPENDIX A: DEPARTMENT OF HEALTH AND HUMAN SERVICES CHANGES

The sections below present a brief summary of the major changes regarding the Department of Health and Human Services (DHHS) Mandatory Guidelines for Federal Workplace Drug Testing Programs, as published in the *Federal Register*, June 9, 1994 (59 FR 29908). For more details, see the above-mentioned *Federal Register* notice.

### A.1: General Definitions

Several changes concern the addition or deletion of various definitions throughout the Mandatory Guidelines for Federal Workplace Drug Testing Programs. The definitions in the Mandatory Guidelines are intended to conform to the definitions established by the National Committee for Clinical Laboratory Standards (NCCLS). These changes to definitions are as follows:

- A definition for *certifying scientists* was added.

This definition was added to ensure that all test results are reviewed by an individual who has the appropriate training and experience in laboratory practices for reviewing such test results. The new definition specifically states that a certifying scientist is an individual who understands chain of custody.

- Definitions for *specimen chain of custody form* and *laboratory chain of custody form* were added.

These definitions were added to clarify the difference between an external form initiated at the collection site and internal forms used for maintaining chain of custody on specimens and aliquots in the laboratory.

- The definition for a *permanent record book* was deleted.

This definition was deleted because all information that was formerly required in the permanent record book is now indicated on each specimen chain of custody form.

- The definitions for *calibrators, controls, and standards* were added.

These definitions were added to define more clearly the requirements of laboratory testing.

- Definitions for *donor, specimen, sample, and quality control sample* were revised to ensure that these definitions are consistent with those proposed by the NCCLS.

### A.2: Role Clarifications

Four new sections were added to clarify the role of the Secretary of Health and Human Services. In addition, two sections were amended to require federal agencies to seek approval for testing for additional drugs. These changes are as follows:

- A section was added to clarify the existing authority of the Secretary of Health and Human Services.

This section reiterates and clarifies that the Secretary can take whatever action is necessary to ensure that DHHS-certified laboratories continue to satisfy all provisions of the Mandatory Guidelines for Federal Workplace Drug Testing Programs and to ensure the full reliability of drug testing.

- Two new sections were added to provide that the Secretary of Health and Human Services will publish announcements in the *Federal Register* when laboratories are certified, when laboratory certifications have been suspended, or when laboratory certifications have been revoked.

These two new provisions coincide with existing DHHS policy. The Secretary will not publish notices in the *Federal Register* when a laboratory's certification is not yet suspended but is proposed to be revoked because, unlike suspended laboratories, these laboratories do not pose an immediate threat to the public health and welfare and therefore may continue federal employee drug testing until such time that the revocation becomes effective. In the meantime, these laboratories have the opportunity to request internal review of the decision to revoke their licenses.

- A section was added to provide that the written notice of suspension that is sent to a laboratory, as well as the reviewing official's written decision upholding or denying the suspension or proposed revocation, will be made available to the public upon request.

- Two sections were amended to require federal agencies covered by Executive Order 12564 to submit in writing a proposed performance test program when seeking approval of the Secretary of Health and Human Services for testing of other drugs.

### A.3: Specimen Collection

Six changes concerning the collection of urine specimens will likely make drug-testing regulations easier for licensee collection sites to enforce. These changes are as follows:

- All references to *record book* or *permanent record book* were replaced with *specimen chain of custody form*.

The specimen chain of custody form now consists of several parts or "batches" and contains the same information that was normally found in the permanent record book. One copy of this form stays at the collection site; thus, collection sites will no longer need to keep a permanent record book.

- The required amount of urine to be collected from individuals has decreased from 60 mL to 30 mL.

This change was made because the collection of a minimum of 60 mL of urine has led to difficulties in a substantial number of cases in which the donor is required to wait and provide additional urine, so that the final specimen is the accumulation of intermittent collections. The Secretary of Health and Human Services believes that decreasing the urine volume to 30 mL is adequate for the laboratory to complete the required testing and satisfy other program requirements.

- Federal agencies, in contracting with collection sites, are allowed to choose an individual, other than a collection site employee, to observe the collection of a specimen whenever there is reason to believe that the donor may have altered or substituted the specimen. The observing individual must still be of the same gender as the donor, however.

This change is based on the Department of Health and Human Services' belief that it is not always possible, under all contracts, to have a collection site employee of the same gender observe the specimen collection. The Department believes this determination is best left to the individual federal agencies.

- A new section was added to describe the collection procedure for split specimens. A minimum of 45 mL of urine is required when using a split specimen collection procedure. Among other things, all requirements for specimen collection are to be followed for both specimens (or Bottle A and Bottle B), including the requirement that both the Bottle A

and Bottle B specimens be sent to the laboratory at the same time and that a copy of the chain of custody form accompany each bottle processed under split specimen procedures. The temperature of Bottle A shall be measured and a minimum of 15 mL of urine shall be poured into Bottle B. If the test of Bottle A is verified by the Medical Review Officer (MRO) as positive, the MRO must report the result to the agency without waiting for the donor to request that the Bottle B specimen be tested.

Allowing federal agencies to collect split specimens brings the DHHS guidelines into concurrence with Department of Transportation (DOT) rules and is based on the DHHS Secretary's belief that this practice would not compromise the drug testing program, provided that both samples are handled with identical security, confidentiality, and chain of custody safeguards. The procedures are intended to provide for these safeguards.

- The time limit for requesting a test of the split sample is specified as 72 hours after the employee is actually notified of a positive test result.

The MRO is to honor such a request if it is made within 72 hours of the employee having actual notice that he or she tested positive. Personnel action required by the federal agency as a result of an MRO-verified positive drug test, such as removal from a safety-sensitive position, may proceed whether Bottle B is or is not tested. If the result of the Bottle B analysis is negative, the MRO will void the results of Bottle A and the donor will re-enter the random testing subject pool as if the test had not been conducted. The MRO will notify the federal agency when a failure to reconfirm has occurred and the agency will contact the Secretary. The Department of Health and Human Services believes that these procedures provide for timely testing of Bottle B and provide for public safety.

- The temperature range specified for urine samples was changed to read whole numbers from 32°-38°C, or 90°-100°F.

This change is based on practical reasons related to the normal divisions indicated on various types of temperature measuring devices. The Secretary of Health and Human Services believes that the slightly increased acceptable temperature range will have no significant impact on the opportunity for an individual to defeat the temperature measurement requirement of the collection procedure.

## A.4: Medical Review Officers

Of the changes concerning the responsibilities of the MRO, two changes create new restrictions for MROs, while two changes make the MRO's duties somewhat easier. These changes are as follows:

- A new section was added and another section was amended to restrict the types of arrangements that can exist between the MRO and the DHHS laboratory. The MRO is restricted from being an employee or an agent of, or having any financial interest in, the DHHS-certified laboratory for which the MRO reviews test results. Likewise, the laboratory is restricted from having any involvement with the MRO that might be a conflict of interest. The MRO's responsibility to report a laboratory's testing deficiencies to the government, however, remain unchanged.

These restrictions will assist in eliminating any conflict of interest between the MRO and the laboratory that may affect the impartiality and objectivity of the MRO in reporting testing deficiencies or errors to appropriate agency officials. These changes are based on the concern that in the retesting process, an MRO with a significant business relationship with a DHHS laboratory may be reluctant to report a false positive to the appropriate authority (*MRO Alert*, 1994).

- A new section was added to clarify that the MRO must submit final drug test results to the federal agency in writing and must ensure the confidentiality of these results.

This provision is simply a clarification of existing practice. Requiring final results to be in writing assists in preventing reporting errors. The Department of Health and Human Services also believes that the MRO must transmit drug test results in a manner designed to ensure confidentiality of the information, since this is sensitive information.

- MROs may report several results on the same correspondence sent to the federal agency.
- One section was simplified to require the MRO to send results only to the designated person in the federal agency rather than to both the agency's Employee Assistance Program and the agency's management official.

The Department believes that the federal agency should have the discretion to determine who should receive test results.

## A.5: Certified Laboratories

Twenty changes concern the role of DHHS-certified laboratories. Each of these changes either decreases laboratory costs or makes complying with DHHS regulations easier. The reasoning behind efforts to reduce costs to laboratories is that reduced costs to laboratories may, in turn, reduce the cost of federal agency testing programs. These changes are as follows:

- Two sections were added to clarify that specimens which test negative on initial immunoassay may be pooled for use in the laboratory's internal quality control program rather than discarded.

This change is being proposed to reduce costs for the laboratories.

- The initial cutoff test level for marijuana metabolites was changed from 100 ng/mL to 50 ng/mL.

This change reflects the advances in technology of immunoassay tests for marijuana metabolites that have occurred since the original Mandatory Guidelines were adopted in 1988. Schwartz et al. (1991) and Armbruster et al. (1993) provide more details concerning these advances in immunoassay technology.

- One section was amended to require that a specimen that tests positive in the initial screen for methamphetamine must contain a concentration equal to or greater than 200 ng/mL of the metabolite amphetamine in the confirmatory test. If this criterion is not met, the specimen must be reported as negative for methamphetamine.

This change would codify current protocol that laboratories use to solve the problem of false positives for methphetamines caused by sympathomimetic amines (including, but not limited to ephedrine, pseudoephedrine, and phenylpropanolamine) found in some over-the-counter and prescription medications. This reporting criterion has been in effect as a temporary requirement since December 22, 1990.

- Another change clarifies that aliquots and laboratory chain-of-custody forms are to be used by laboratory personnel for conducting initial and confirmatory tests while the original specimen (primary or split)

## Appendix A

and specimen chain-of-custody form remain in secure storage.

This amendment clarifies what is already expected of laboratories.

- A new section was added to require agencies to purchase only blind quality control materials that (1) have been certified by immunoassay and gas chromatography/mass spectrometry (GC/MS), and (2) have stability data which verifies their performance over time.

This requirement is included because there is no current uniform criteria for the preparation, certification, and stability of urine samples provided by numerous vendors for use in agencies' blind performance testing programs. Requiring the certification data assists in preventing the use of materials that may be unacceptable as blind quality control specimens.

- Another change allows a laboratory to designate a certifying scientist who is only qualified to certify results that are negative on the initial test. The individual who certifies negative results is expected, for example, to understand all aspects of immunoassay tests, chain of custody, and quality assurance and to examine negative test results in light of that knowledge.

This change will assist in decreasing the cost of testing without compromising the reliability of drug testing. In keeping with the Mandatory Guidelines, however, individuals who certify positive test results are expected to understand all aspects of laboratory procedures and must examine the positive results in light of the immunoassay test, GC/MS results, quality assurance, chain of custody, and all other relevant factors. The Department of Health and Human Services believes that these requirements will ensure that quality testing is maintained.

- One section was amended to permit laboratories to perform multiple initial (screening) tests for the same drug or drug class. For example, laboratories will be allowed to use an initial test performed by one immunoassay technique and then forward all presumptive positives for a second test by a different immunoassay technique to minimize possible presumptive positives due to the presence of structural analogues in the specimen.

This procedure should decrease the cost of drug testing while maintaining the full reliability of drug testing because all tests performed must meet all Mandatory Guideline cutoffs and quality control requirements.

According to the Substance Abuse and Mental Health Services Administration (SAMHSA), amphetamines are the primary reason for this change. Therefore this amendment will probably be limited by the Nuclear Regulatory Commission (NRC) and DHHS laboratories to amphetamines.

- A section was amended to allow a certified laboratory to report one quantitative result differing by more than 50 percent from the target value within any two consecutive cycles of performance testing.

This change allows a certified laboratory to make one administrative or technical error on a performance testing sample and be given an opportunity to take internal corrective action to ensure that this kind of error will not reoccur before an external investigation becomes necessary.

- A section was amended to change the proficiency testing cycles for certified laboratories from six cycles per year to four cycles per year, that is, on a quarterly basis.

Experience in this and other performance testing programs indicates that four cycles per year is sufficient to assess a laboratory's ability to forensically test and report results for proficiency testing specimens.

- The Department of Health and Human Services changed the laboratory quality control requirements for confirmation tests. During the initial 90-day period of any new contract, each agency is required to submit blind performance test samples to contracting laboratories in the amount of at least 20 percent of the total number of specimens submitted (up to a maximum of 200 blind samples). Thereafter, the Department has reduced the percentage of blind samples or unmarked quality control samples that a federal agency is required to maintain and send to a laboratory from 10 percent to 3 percent (with a maximum of 100 blind samples) submitted per quarter. There is no prohibition, however, for an agency to use a higher percentage or a higher number of blind samples to be submitted with donor specimens.

This change could significantly reduce the costs associated with maintaining a blind sample program

without affecting the ability to monitor a laboratory's performance. In addition, the 3 percent requirement coincides with the requirement adopted by the DOT for their regulated industry programs.

- One section was changed to clarify that rather than have DHHS conduct an initial investigation of any unsatisfactory blind performance testing results, the contracting agency shall conduct an initial investigation and submit its findings to the Department.

This section was changed because the Department believes that the agency must gather all pertinent information and investigate the reason for unsatisfactory blind performance test results before DHHS is contracted to continue the investigation and to ensure that the laboratory has taken corrective action.

- One change states that laboratories may conduct dilution/adulteration testing to determine the validity of a sample. Such testing, however, is optional. Laboratories may use their best judgment to select tests to be conducted for possible dilution/adulteration as well as the cutoff levels for such tests.
- The last sentence in section 2.4(g)(1), "The results (positive and negative) for all specimens submitted at the same time to the laboratory shall be reported back to the Medical Review Officer at the same time," has been deleted.

This requirement is no longer necessary or practical since individual specimen chain of custody forms, rather than a batch chain of custody form containing information on more than one urine specimen, are used for all federal employee urine specimens.

- One section was amended to allow, but not require, a federal agency to carry out pre-award inspections and evaluation of the procedural aspects of the laboratory's drug testing operation.

This change allows an agency to determine whether a pre-award inspection is necessary.

- A sentence in one section was amended to clarify that a certified laboratory must inform all clients when procedures followed for those clients do not conform to the standards specified in the Mandatory Guidelines or when any action is taken which suspends or revokes the laboratory's certification.

This change is meant to ensure that a certified laboratory does not use its certification to promote itself as DHHS-certified if, in fact, it uses procedures that do not comply with the Mandatory Guidelines. It is essential that laboratory clients are aware of any departure from the Mandatory Guidelines or any problems with the laboratory's certification.

- One section was revised to clarify how a laboratory who has had its certification revoked may seek recertification.

The Department of Health and Human Services revised one section to provide that, unless otherwise provided by the Secretary of Health and Human Services in the notice of suspension or proposed revocation or the reviewing official's decision, a laboratory which has had its certification revoked and seeks to be recertified must meet the criteria of section 3.12(b), as well as all other requirements of the Mandatory Guidelines, including the successful participation in three cycles of performance testing and a laboratory inspection. Once recertified, the laboratory must undergo a second inspection within three months, after which biannual inspections will be required to maintain certification, as well as participation in the quarterly performance testing program.

- One section was revised to clarify that inspectors are to perform inspections consistent with the guidance provided by the Secretary of Health and Human Services. This revision also clarifies that laboratories are required to follow good forensic laboratory practice in all aspects of their drug testing operations and are required to be in compliance with the Mandatory Guidelines.

It is the laboratory's responsibility to correct all deficiencies identified during the inspection and to have the knowledge, skill, and expertise to correct deficiencies consistent with good forensic laboratory practice.

- One section was clarified to read that a laboratory must satisfy all pertinent provisions of the Mandatory Guidelines in order to maintain certification.

This clarification was made because the original requirement only addressed satisfying the provisions in order to *qualify* for certification.

- A new subpart D was added which sets forth more detailed procedures for the internal review of a suspension or proposed revocation of a certified laboratory.

## Appendix A

It is the Department of Health and Human Services' view that the following procedures will provide a timely and fair review of suspensions or proposed revocations. In general, these procedures describe how laboratories may request an informal review of the immediate suspension and the proposed revocation and how the review will generally be conducted. The presiding official will review documents and briefs that are submitted and may provide for a hearing at which time each party may present witnesses as agreed upon in a prehearing conference and may question the opposing party's witnesses.

More specifically, the procedures provide that a laboratory has 30 days from the written notice of suspension or proposed revocation or 3 days from notification for an expedited review of the suspension to request a review. The National Laboratory Certification Program within the Division of Workplace Programs, Substance Abuse and Mental Health Services Administration, must prove, by a preponderance of the evidence, that its decision to suspend or propose revocation is appropriate. If the reviewing official upholds the suspension and proposed revocation, the revocation will become effective immediately and, if the suspension and proposed revocation are denied, the suspension will be lifted immediately.

The procedures also provide for an abeyance agreement. That is, the parties may agree to hold the informal review procedures in abeyance for a reasonable time while the laboratory attempts to regain compliance with the Mandatory Guidelines. For example, if a laboratory receives notice of an immediate suspension and a proposed revocation and prefers to remedy the deficiencies rather than proceed immediately with an informal review, the parties involved may agree, upon mutually acceptable conditions, to extend the time periods for requesting review of the suspension or proposed revocation. If the dispute has been resolved, the request for review will be dismissed.

- The Department has stated that the written notice of suspension sent to a laboratory, as well as the reviewing official's written decision upholding or denying the laboratory's suspension or proposed revocation, are to be made available to the public upon request.

This provision ensures that the public has access to the documents containing the basis for DHHS actions.

## References

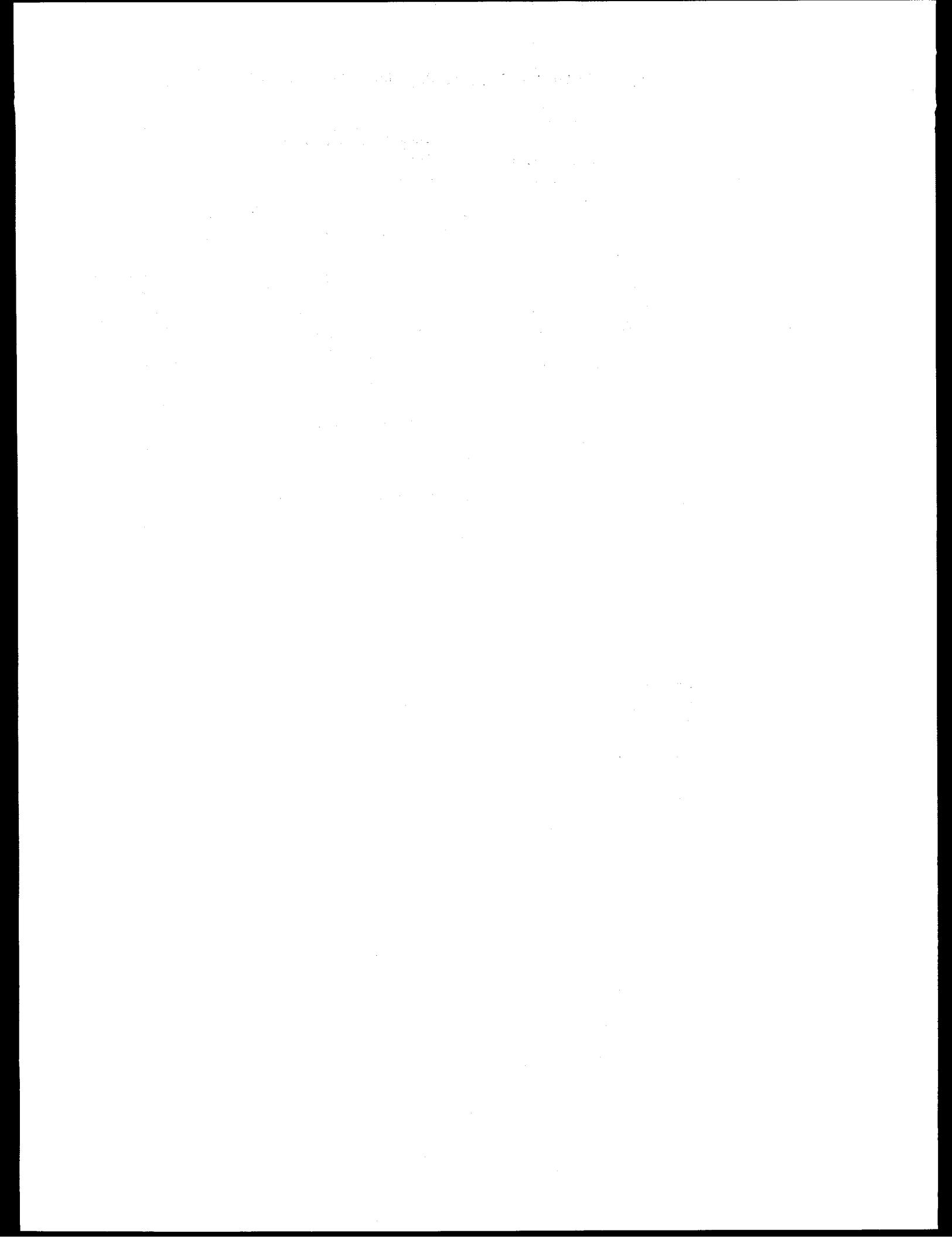
59 FR 29908. (June 9, 1994). Department of Health and Human Services, Substance Abuse and Mental Health Services Administration. "Mandatory Guidelines for Federal Workplace Drug Testing Programs." *Federal Register*.

Armbruster, D.A., Schwarzhoff, R.H., Hubster, E.C. and Liserio, M.K. (1993, October). Enzyme immunoassay, kinetic microparticle immunoassay, radioimmunoassay, and fluorescence polarization immunoassay compared for

drugs-of-abuse screening. *Clinical Chemistry*, 39(10), 2137-2146.

*MRO Alert*. (1994, June). Revised mandatory guidelines for federal workplace drug testing programs, 5(5):2-7.

Schwartz, J.G., Zollars, P.R., Okorodudu, A.O., Carnahan, J.J., Wallace, J.E. and Briggs, J.E. (1991, March). Accuracy of common drug screen tests. *American Journal of Emergency Medicine*, 9(2), 166-170.



## APPENDIX B: NON-INSTRUMENTED SCREENING TEST DEVICES

Non-instrumented devices have become available for immunoassay screening of urine samples for drugs of abuse. These testing devices show great promise as quick and easy methods for testing on-site. They have been shown to be relatively reliable and are being used in hospital emergency rooms. The Substance Abuse and Mental Health Services Administration (SAMHSA) has been tasked by Congress to evaluate these devices.

Under the current Nuclear Regulatory Commission (NRC) Fitness-for-Duty (FFD) rule (10 CFR, Part 26), licensees are not precluded from using these testing devices. There are indications that they are being used and/or will be used by NRC licensees in their FFD programs. There are several potential issues that have been raised concerning the use of these devices:

- Although manufacturers state that the accuracy and specificity of these tests (i.e., the expectations of false positives and false negatives) are comparable to other immunoassay techniques (e.g., EMIT<sup>®</sup>, RIA) their accuracy and specificity has not been thoroughly established by independent agencies.
- The potential for subversion using these devices has not been addressed, for example, the impact of adulterants and dilute specimens on these devices is not known.
- Quality assurance procedures may be different. For example, procedures are needed so licensees could validate the accuracy and reliability of these screening mechanisms around a cutoff level.
- Three of the six devices mentioned in this appendix require multiple tests in order to test for all five of the drugs in the SAMHSA panel. Some devices require as many as five separate tests. This requirement could negatively impact the collection accuracy as storing, handling, using, and records keeping become more complex and each added step introduces a new opportunity for error. Chain-of-custody forms also may be more burdensome.
- Screening tests using instrumented devices (RIA, EMIT<sup>®</sup>, etc.) are performed separately from collection, providing a built-in safeguard against individuals involved in the testing process knowing the identity of a donor. With non-instrumented testing devices this built-in safeguard may be lost. The protection of individual privacy may be a problem with non-instrumented testing devices, if, for example, the individual collecting the specimen also immediately performed the test. Although the NRC's FFD rule requirement that trained personnel

supervised by a professional perform any on-site testing should protect against this occurring, the ease of use of these devices could result in unintended violation of the rule and/or an individual's rights.

There are also reasons that non-instrumented testing devices may not result in expected cost savings:

- Much of the expected potential monetary savings resulting from using these kits derives from the ability to use less experienced laboratory employees to administer the tests. Although non-instrumented devices do not require such a level of training (device literature expound on the ease of use and short training time required by those who will use the kits) rule requirements would preclude using less expensive non-professionals for on-site testing.
- Shelf life may be an issue.

This Appendix addresses several currently available test kits. To establish uniformity each product will be described using the following format:

### I. General Information

### II. Device Characteristics

- Wash/Fixer: Does this test require that separate wash and fixer solutions be applied in turn?
- Built-In Control: Does each device include procedures which establish that the test is working?
- Storage: Does this device require a special storage environment?
- Color Change Dependence: Does accurate reading of this device require the ability to distinguish shades and hues of color?
- Approximate Cost per Test: What is the market price of this device per testing episode? Note: these figures are subject to change by the manufacturer and usually depend on the quantity of devices purchased.
- Availability: In what quantities are the screening devices available?

### III. Comparison of Testing Results

Reported testing results from the non-instrumented devices have been provided by some manufacturers.

## Appendix B

In these reports comparisons of the results from the devices are made to results from other testing methods, primarily RIA, EMIT®, and GC/MS. These reports are generally based on small studies sponsored by the manufacturer. The results are reported here for informational purposes only. Often manufacturer claims of accuracy, ease of use, etc. have not undergone independent study. In these instances only the manufacturer's claims can be reported. Any information on testing results from an independent study will be noted; otherwise this information should be considered as provided by the manufacturer.

### IV. Potential Advantages Compared with Other On-Site Testing Devices

### V. Differences Compared with Other On-Site Testing Devices

### VI. Other Comments

Where possible, the results of independent studies will be cited.

**Abuscreen ONTRAK®****I. General Information**

a. Trade Name:	Abuscreen ONTRAK®
b. Manufacturer:	Roche Diagnostic Systems
c. Address:	1080 U.S. Hwy. 202 Branchburg, NJ 09976-1760 1-800-777-1000
d. Representative:	Rod Sprague or Al Arnold 503-492-3312 Oregon/WA/Alaska 1-800-526-1247 1-800-777-1000

**II. Device Characteristics**

a. Wash/Fixer	Yes
b. Built-in Control	No
c. Storage	No refrigeration required
d. Color Change Dependence	No, agglutination method
e. Approximate Cost/Test	\$3.50
f. Availability	100 tests - \$385; 50 tests for \$225

**III. Comparison of Testing Results**

Abuscreen ONTRAK® results were compared with GC/MS results to assess sensitivity. Agreement between the results are as follows:

Amphetamines	88.2 %	Barbiturates	100 %
Cocaine	100 %	Opiates	99.1 %
THC	98.6 %		

Abuscreen ONTRAK® results were compared with RIA results to assess specificity. Agreement for amphetamines, barbiturates, cocaine, opiates, and THC was 100%.

Samples assessed as positive by Abuscreen ONTRAK® were confirmed by GC/MS with 100% agreement.

According to a non-scientific study by the State of Tennessee, the Roche test kit was the most accurate overall of four products tested.

**IV. Potential Advantages Compared with Other On-Site Testing Devices**

The Abuscreen ONTRAK® system does not rely on color change but rather on the presence or absence of "clumping" of fluid in the viewing window. Clumping indicates a negative test result. *ONTRAK is accurate, the cost is competitive [\$3.50/test card], and tests can be run in tandem (Schwartz et al., 1993).*

**V. Differences Compared with Other On-Site Testing Devices**

A separate kit is needed for each tested drug. OnTrak kits are currently available for the detection of cocaine, THC, methadone, amphetamines, barbiturates, PCP, and benzodiazepines. There is no built-in control per test kit. Positive reference control solutions must be purchased separately. The control solutions are multianalyte control tests which control for each of the eight drugs. Test results must be read within 5 to 10 minutes after reagents are mixed or results may be read as "negative" (clumping dissipates with time). *Some training seems to be required to conduct these tests. Nonmedical volunteers who ran the tests had an unacceptably high level of false-negative results for marijuana and cocaine. Roche Diagnostics now offers a video and regional training instructors (Schwartz et al., 1993).*

**VI. Other Comments**

A Florida Dept. of Corrections study (1993) testing for cocaine metabolite and Carboxy-THC reported that the OnTrak system was the most accurate when compared with EZ-SCREEN® and microLINE®.

From the *Schwartz et al., 1993*:

"In terms of accuracy, this test performed virtually perfectly when conducted by a laboratory technologist, comparing favorably with GC/MS. When performed by office nurses, it was slightly less accurate."

## Appendix B

### ONTRAK TesTcup®

#### I. General Information

a. Trade Name: ONTRAK TesTcup®  
b. Manufacturer: Roche Diagnostic Systems  
c. Address: 1080 U.S. Hwy. 202  
Branchburg,  
NJ09976-1760  
d. Representative: Al Arnold  
Phone: 1-800-526-1247

#### II. Device Characteristics

a. Wash/Fixer No  
b. Built-in Control Yes  
c. Storage No refrigeration required prior to test. After test is run, sample may be frozen for prolonged storage.  
d. Color Change Dependence Yes  
e. Approximate Cost/Test \$3.50  
f. Availability 25 cups in a box.  
Each cup tests for a panel of three drugs (marijuana, cocaine, morphine).

#### III. Comparison of Testing Results

ONTRAK TesTcup® was evaluated using specimens screened by an automated immunoassay and confirmed positive by GC/MS. Agreement was reported by the manufacturer as follows:

Cocaine	100%
TCH	100%
Morphine	100%

The precision of ONTRAK TesTcup® was determined by testing various concentrations of drug standards for primary drugs or metabolites detected by the system. There is a greater than ninety-five percent (95%) confidence level that a positive result will be attained with drugs at 150% of its concentration. Results can be interpreted for up to 30 minutes after the test is run.

#### IV. Potential Advantages Compared with Other On-Site Testing Devices

The ONTRAK TesTcup® is a one-step testing procedure which simultaneously tests for a panel of three drugs (marijuana, cocaine, and morphine). The ONTRAK TesTcup® has a built-in control to verify quality.

#### V. Differences Compared with Other On-Site Testing Devices

The ONTRAK TesTcup® serves as a test device for detection of drug or drug metabolite as well as a urine collection and storage container. Fresh urine samples are delivered directly into the cup and do not require any special handling or pretreatment.

The test is color dependent. A blue band or any color in the test result window indicates a negative result. A positive result is indicated by the lack of a blue band or any color in the test result window.

#### VI. Other Comments

The ONTRAK TesTcup® is extremely easy to use and results are available in under five minutes. It is recommended that specimens be tested shortly after collection. However, samples may be stored in the ONTRAK TesTcup® at room temperature (65-85°F) and tested within 8 hours of collection. Specimens can be frozen for prolonged storage.

**MicroLINE® Screen****I. General Information**

a. Trade Name:	microLINE® Screen
b. Manufacturer:	Drug Screening Systems, Inc. (DSSI)
c. Address:	Box 579 Blackwood, NJ 08012 1-800-247-3784
d. Representative:	Patty Lydic
Phone:	1-800-247-3784

**II. Device Characteristics**

a. Wash/Fixer	No
b. Built-in Control	Yes
c. Storage	Refrigeration of kit prior to use
d. Color Change Dependent	No
e. Approximate Cost/Test	\$4
f. Availability:	Test panels are sold in packages of 20

**III. Comparison of Testing Results**

The manufacturer reported the following agreement between microLINE® and RIA/EMIT® testing results:

Amphetamines	100 %	Cocaine	98.0 %
Opiates	98.8 %	THC	97.0 %
PCP	100 %		

The manufacturer reports that independent clinical testing performed by the Toxicology Laboratory at the hospital of the University of Pennsylvania clearly shows that DSSI screening systems achieve performance equal to or better than laboratory screening tests. Furthermore, DSSI antibodies are highly specific to the drug of abuse being tested or

its metabolite, and little or no cross reacting substances will be detected.

**IV. Potential Advantages Compared with Other On-Site Testing Devices**

This test system requires no mixing or adding of solutions. Six separate drug tests can be accomplished with one kit. Tests are available for 8 different drugs (barbiturates, amphetamines, benzodiazepines, cocaine, methamphetamines, heroin/opiates, THC). The presence or absence of a second dark line in the viewing window gives the test result. A single line on the left side of the viewing window indicates a positive result. This is the "reference" line and it indicates that the test has run properly. Two lines (one on the left, one on the right) indicate a negative test result. Color change interpretation is unnecessary. Each test has a built-in control for quality assurance. If no lines appear, the test did not run and should be repeated with a different plate.

**V. Differences Compared with Other On-Site Testing Devices**

Test kits must be refrigerated then allowed to come to room temperature (30-45 minutes) prior to use. The results must be read 10 minutes after the test is run.

**VI. Other Comments**

The requirement for refrigeration could be a handicap should this test kit be used outside a controlled environment. A large, refrigerated storage area would be necessary to support an extensive testing program. Usually tests for amphetamines and benzodiazepines are sold as separate single panels and the remaining 6 drugs are sold as one kit (panel of 6 drugs). However, only a 6-drug combination of the 8 available drugs is possible. DSSI also markets a microLINE® Screen kit that tests urine only for cocaine and marijuana, the two most common drugs of abuse. In addition, the company distributes a simple "test tube" type blood alcohol concentration (BAC) tester using specialized crystals which indicate by a color change the presence of alcohol when exhaled breath contacts the crystals. This test is available for 4 different percentages (.02, .04, .08, 1.0). The product's retail price is \$1.75 per test.

## Appendix B

### Triage

#### I. General Information

a. Trade Name: Triage  
b. Manufacturer: Biosite Diagnostics  
c. Address: 11030 Roselle Street  
San Diego, CA  
92121  
d. Representative: Gail Spotts  
Phone: 800-745-8026  
Ext 3120

#### II. Device Characteristics

a. Wash/Fixer Yes  
b. Built-in Control Yes  
c. Storage Room temperature  
d. Color Change Dependent Yes, any change = positive  
e. Approximate Cost/Test 7-drug panel \$26.20 per device; 5-drug panel \$20.00 per device (volume discounts are available)  
f. Availability Test kits come in packages of 25

#### III. Comparison of Testing Results

The manufacturer reports the following agreement between results from the Triage Panel and GC/MS:

Amphetamines	97.0 %	THC	99.0 %
Opiates	97.0 %	Barbiturates	100.0 %
PCP	100.0 %	Opiates	97.0 %
Cocaine	99.0 %	Benzo-diazepines	97.0 %

According to Wu, et al (1993), the agreement between Triage and GC/MS for positive samples for

931 samples ranged from 93 to 100%. For negative samples the agreement was between 95 and 100%. "The majority of the urine samples studied . . . produced identical results to a commercial-instrument-based immunoassay."

#### IV. Differences Compared with Other On-Site Testing Devices

One Triage kit can test for all five of the SAMHSA drugs. A kit consisting of a panel of 7 drugs is also available. The panel includes tests for the 5 SAMHSA drugs plus benzodiazepines and barbiturates. The kit may be stored at room temperature. A built-in control within each kit gives a positive indication of the kit's functioning. Any color change is read as positive, eliminating the need to judge degree of color change.

#### V. Differences compared with other on-site test kits

Triage requires the kit user to "draw-off" the urine from the "Reaction Cup" and transfer it to the "Detection Area." This step could be mishandled, resulting in decreased accuracy. The built-in control, however, could alert the user to this possibility. The device requires an incubation period of 10 minutes followed by a chemical wash. Results must be read within 5 minutes after applying the wash solution.

#### VI. Other Comments

Biosite Diagnostics offers test kits with combinations of tests. Customizing the test panel to perform only those tests required by the NRC or adding tests to the panel to meet licensee needs could be arranged with the manufacturer.

**EZ-SCREEN®****I. General Information**

a. Trade Name:	EZ-SCREEN®
b. Manufacturer:	EDITEK, Inc.
c. Address:	1238 Anthony Road Burlington, NC 27215
d. Representative: Phone:	Mary DeChellis 1-800-237-7352

**II. Device Characteristics**

a. Wash/Fixer	Yes
b. Built-in Control	Yes
c. Storage	Refrigeration for kit and reagents
d. Color Change Dependent	Yes
e. Approximate Cost/Test	\$9
f. Availability	20 in a box (bulk test kits are available on request)

**III. Comparison of Testing Results**

The manufacturer reports the following agreement between results from EZ-SCREEN® and GC/MS:

Amphetamines	97.5 %	Cocaine	97.3 %
Opiates	98.3 %	THC	94.7 %
PCP	99 %		

In addition, EZ-SCREEN® correctly identified:

48 of 51 positive samples (94.1%) for cocaine metabolite;

49 of 50 positive samples (98.0%) for amphetamine;

49 of 50 positive samples (98.0%) for PCP;

59 of 60 positive samples (98.0%) for opiate (there was one *false negative*); and

44 of 48 positive samples (92.7%) for THC.

**IV. Potential Advantages Compared with Other On-Site Testing Devices**

The EZ-SCREEN® test has both a negative and a positive control to verify reagent precision. A single EZ-SCREEN® device tests for three drugs (marijuana, cocaine, and amphetamines). The manufacturer just began marketing a new EZ-SCREEN® profile which includes 5 drugs in its testing panel (marijuana, cocaine, opiates, amphetamines, and PCP).

**V. Differences Compared with Other On-Site Testing Devices**

Reagents must be brought to room temperature before using (~1 hr). The test process requires approximately 17 steps. One vial of liquid is shaken vigorously while another is mixed gently. Any excess liquid must be removed from around the test cells to avoid cross contamination. The negative control should be grey-blue in color, the positive control should be lighter than the negative control. A positive test result should display a color equal to or lighter than the adjacent positive control site color. A negative test result should be darker than the adjacent positive control site. It is important for the interpreter of the results of this test to be able to distinguish between shades of grey-blue. A separate test kit is required for each of the five drugs.

**VI. Other Comments**

The number of steps required in this procedure together with the requirement that the reagent be refrigerated may preclude field use under less than ideal conditions. Error margin may increase when a test depends on time and temperature, and involves multiple steps. As in all color-dependent test reactions, the ability to differentiate shades of color is important. The amphetamine test does not indicate a specific drug nor distinguish between *l* and *d* isomers.

## Appendix B

### **VERDICT**

#### I. General Information

a. Trade Name:	<b>VERDICT</b>
b. Manufacturer:	EDITEK, Inc.
c. Address:	1238 Anthony Road Burlington, NC 27215
d. Representative: Phone:	Mary DeChellis 1-800-237-7352

#### II. Description of test type/method

a. Wash/Fixer	No
b. Built-in Control	Yes (blue line in validation window = valid test)
c. Storage	Room temperature
d. Color Change Dependent	No (blue line = negative, no line in result window = positive)
e. Approximate Cost/Test	\$4 (quantity discounts available)
f. Availability	50 per box

#### III. Comparison of testing results

The manufacturer reported the following agreement between VERDICT testing results and GC/MS testing results:

Amphetamines	97.5%	Cocaine	97.3%
Opiates	99.0%	THC	94.7%
PCP	99.0%		

Of 60 positive samples, VERDICT correctly identified 59 (98% accurate).

#### IV. Potential Advantages Compared with Other On-Site Testing Devices

VERDICT does not require fixer, rinse, etc. The test is very simple and considerably faster than most. The device may be kept at ambient temperature (not to exceed 77 degrees F) for three months, though refrigeration prolongs the usable shelf life to two years.

#### V. Differences Compared with Other On-Site Testing Devices

A double test is available which tests for both marijuana and cocaine at one time. The remaining drugs require a separate Verdict testing device for each individual test.

#### VI. Other Comments

The manufacturer notes that the presence of any blue vertical line in the validation and/or results window is sufficient for interpretation of results. However, because of the qualitative nature of the test and because visual interpretation can be subjective, there may be samples that produce results that are not positive, but that are not visually distinct enough to be interpreted as negative. It is suggested that these samples be retested by some other technique.

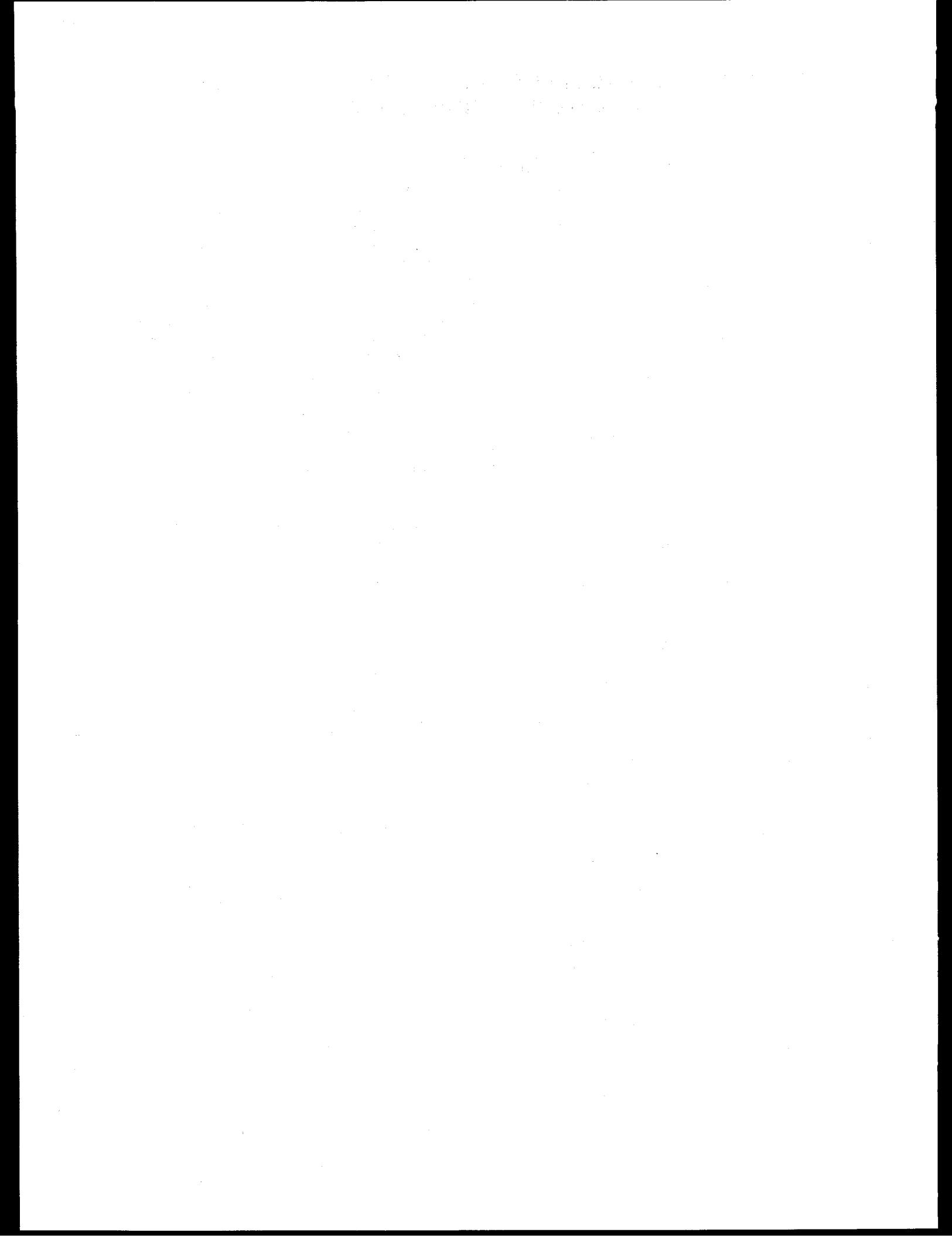
## References

10 CFR Part 26. (January 1, 1990). U.S. Nuclear Regulatory Commission. "Fitness-for-duty programs." *U.S. Code of Federal Regulations*.

Wu, A.H.B. and Wing, S.S.; Johnson, K.; Calies, J. and Shu, D.X; Dunn, W.E. and Wong, S.H.V. (1993) "Evaluation of the Triage System for Emergency Drugs-

of-Abuse Testing in Urine." *Journal of Analytical Toxicology*. Vol. 17, July-August pp. 241-245.

Schwartz, R.H.; Clark, H.W.; Meek, P.S. (September, 1993). Laboratory tests for rapid screening of drugs of abuse in the workplace: A Review. *Journal of Addictive Diseases*, September, 1993. Vol. 12 No. 2, pp. 43-56.



## APPENDIX C: REQUIREMENTS OF THE OCCUPATIONAL EXPOSURE TO BLOODBORNE PATHOGENS RULE

The sections below present an overview of the requirements regarding compliance with the Occupational Exposure to Bloodborne Pathogens standard. For more details, Nuclear Regulatory Commission (NRC) licensees and contracted clinical facilities should refer to 29 CFR Part 1910.1030.

### C.1: Exposure Control

Employers are required to establish a written Exposure Control Plan if they have personnel who have a reasonable chance of being exposed to bloodborne pathogens, including the hepatitis B virus (HBV) and the human immunodeficiency virus (HIV), during the normal performance of their job duties. The Exposure Control Plan should be accessible to employees and should include the following:

- An exposure determination; and
- Schedules and methods of implementation for methods of compliance, hepatitis B vaccination and post-exposure evaluation and follow-up, communication of hazards to employees, and recordkeeping.

### C.2: Methods of Compliance

In general, employees should observe universal precautions when working with blood and other potentially infectious materials; that is, they should treat all human blood and certain human body fluids as if they were known to be infectious for HIV, HBV, and other bloodborne pathogens. In addition, employers are required to provide engineering and work practice controls to eliminate or minimize employee exposure; provide personal protective equipment to employees, if occupational exposure remains after implementation of the engineering and work practice controls; and, at all times, ensure that the worksite is maintained in a clean and sanitary condition.

Engineering and work practice controls include the following:

- Provide readily accessible handwashing facilities to employees, or if providing handwashing facilities is not feasible, provide access to antiseptic hand cleansers and clean cloth/paper towels or antiseptic towelettes.

- Ensure that employees wash their hands after removing gloves or other personal protective equipment.
- Ensure that employees wash their hands and any other skin with soap and water, or flush mucous membranes with water after such body areas contact blood or other potentially infectious materials.
- Ensure that employees do not bend, recap, remove, break, or shear contaminated needles and other contaminated sharps, unless the employer can demonstrate that no alternative is feasible or that such action is required by a specific procedure.
- Ensure that, as soon as possible after use, contaminated reusable sharps are placed in puncture resistant, leakproof, labeled or color-coded containers until the sharps can be properly reprocessed.
- Prohibit employees from eating, drinking, smoking, applying cosmetics or lip balm, or handling contact lenses in work areas where there is a reasonable likelihood of occupational exposure.
- Prohibit employees from keeping food and drink in refrigerators, freezers, shelves, cabinets, or on countertops or benchtops where blood and other potentially infectious materials are present.
- Ensure that all procedures involving blood or other potentially infectious materials are performed in a manner that minimizes splashing, spraying, spattering, and generation of droplets of these substances.
- Prohibit mouth pipetting/suctioning of blood or other potentially infectious materials.
- Ensure that specimens of blood or other potentially infectious materials are placed in leakproof containers during collection, handling, processing, storage, transport, or shipping.
- Ensure that equipment, which may become contaminated with blood or other potentially infectious materials, is examined prior to servicing or shipping and decontaminated as necessary, unless the employer can demonstrate that decontamination of the equipment is not feasible.

Personal protective equipment provisions include the following:

## Appendix C

- Ensure that the employer provides, at no cost to employees, appropriate personal protective equipment, including but not limited to, gloves, gowns, laboratory coats, face shields or masks and eye protection, and mouthpieces, resuscitation bags, pocket masks, or other ventilation devices.
- Ensure that employees use appropriate personal protective equipment.
- Ensure that personal protective equipment in the appropriate sizes and hypoallergenic versions is readily accessible at the worksite or is issued to employees.
- Ensure that the employer cleans, launders, and disposes of personal protective equipment at no cost to the employee.
- Ensure that the employer repairs or replaces personal protective equipment as needed to maintain its effectiveness, at no cost to the employee.
- Ensure that employees remove a garment that is penetrated by blood or other potentially infectious materials as soon as feasible.
- Ensure that all personal protective equipment is removed prior to employees leaving the work area.
- Ensure that when employees remove personal protective equipment, they place it in an appropriately designated area or container for storage, washing, decontamination or disposal.
- Ensure that employees wear gloves when it can be reasonably anticipated that the employee may have hand contact with blood, other potentially infectious materials, mucous membranes, and non-intact skin; when performing vascular access procedures; and when handling or touching contaminated items or surfaces.
- Ensure that employees wear masks in combination with eye protection devices, such as goggles or glasses with solid side shields, or chin-length face shields, whenever splashes, spray, spatter, or droplets of blood or other potentially infectious materials may be generated and eye, nose, or mouth contamination can be reasonably anticipated.
- Ensure that employees wear appropriate protective clothing such as, but not limited to gowns, aprons, lab coats, clinic jackets, or similar outer garments, depending on the task and degree of exposure anticipated.
- Ensure that employees wear surgical caps or hoods and/or shoe covers or boots in instances when gross contamination can reasonably be anticipated.

Housekeeping provisions include the following:

- Ensure that employers maintain the worksite in a clean and sanitary condition. To this end, the employer should determine and implement an appropriate written schedule for cleaning and decontaminating the worksite area.
- Ensure that all equipment and environmental and working surfaces are cleaned and decontaminated after contact with blood or other potentially infectious materials.
- Ensure that contaminated work surfaces are decontaminated with an appropriate disinfectant.
- Ensure that protective coverings, such as plastic wrap, aluminum foil, or imperviously-backed absorbent paper used to cover equipment and environmental surfaces, are removed and replaced as soon as feasible or at the end of a work shift if these coverings become overtly contaminated.
- Ensure that all bins, pails, cans, and similar receptacles intended for reuse are inspected and decontaminated on a regularly scheduled basis and cleaned and decontaminated as soon as feasible upon visible contamination.
- Ensure that employees use a brush and dust pan, tongs, or forceps to pick up broken glassware which may be contaminated.
- Ensure that reusable sharps that are contaminated with blood or other potentially infectious materials are not stored or processed in a manner that requires employees to reach by hand into the containers where these sharps have been placed.
- Ensure that employees discard, as soon as feasible, contaminated sharps in containers that are closable, puncture resistant, leakproof, and labeled or color coded.

- Ensure that containers for contaminated sharps are 1) easily accessible to employees who use or find sharps, 2) maintained upright throughout use, and 3) replaced routinely and not allowed to overfill.
- Ensure that when employees move containers of contaminated sharps, the containers are 1) closed immediately prior to removal or replacement to prevent spillage or protrusion of contents during handling, storage, transport, or shipping; and 2) placed in a secondary container that is closable, leakproof, and labeled or color coded, if leakage of the primary container is possible.
- Ensure that reusable containers are not opened, emptied, or cleaned manually or in any other manner which would expose employees to the risk of percutaneous injury.
- Ensure that other regulated waste is placed in closable, leakproof, labeled or color-coded containers that are closed prior to removal to prevent spillage or protrusion of contents during handling, storage, transport, or shipping.
- Ensure that if outside contamination of the regulated waste container occurs, it is placed in another closable, leakproof, labeled or color-coded container that is closed prior to removal to prevent spillage or protrusion of contents during handling, storage, transport, or shipping.
- Ensure that all regulated waste is disposed of in accordance with applicable regulations of the United States, states and territories, and political subdivisions of states and territories.
- Ensure that contaminated laundry is handled as little as possible with a minimum of agitation.
- Ensure that contaminated laundry is bagged or put in a container at the location where it was used and is not sorted or rinsed in the location of use.
- Ensure that contaminated laundry is placed and transported in labeled or color-coded bags or containers.
- Ensure that, whenever contaminated laundry is wet and presents a reasonable likelihood of soaking through or leaking from the bag or container, the laundry is placed and transported in bags or containers which prevent soak-through or leakage of fluids.
- Ensure that employees who have contact with contaminated laundry wear protective gloves and other appropriate personal protective equipment.

### **C.3: Hepatitis B Vaccination and Post-Exposure Evaluation and Follow-Up**

In general, employers are required to make available, at no cost to the employee, the hepatitis B vaccine and vaccination series to all employees who have occupational exposure to blood or other potentially infectious materials. Employers must also provide, at no cost to the employee, post-exposure evaluation and follow-up to all employees who have had an exposure incident. For specific information regarding providing hepatitis B vaccination and post-exposure evaluation and follow-up for employees, see 29 CFR Part 1910.1030.

### **C.4: Communication of Hazards to Employees**

Provisions for communication of hazards to employees include labels and signs and information and training. The following requirements pertain to labels and signs:

- Ensure that warning labels are affixed to containers of regulated waste, refrigerators and freezers containing blood or other potentially infectious material; and other containers used to store, transport, or ship blood or other potentially infectious materials.
- Ensure that labels include the BIOHAZARD label. These labels should be fluorescent orange or orange-red or predominantly so, with lettering and symbols in a contrasting color.
- Ensure that the labels are affixed as close as is feasible to the container by string, wire, adhesive, or other method that prevents their loss or unintentional removal.

Exceptions to using labels and signs as noted above include the following situations:

- When red bags or red containers are substituted for labels.

## Appendix C

- When containers of blood, blood components, or blood products that are labeled as to their contents and have been released for transfusion or other clinical use.
- When individual containers of blood or other potentially infectious materials that are placed in a labeled container during storage, transport, shipment or disposal.
- When regulated waste has not been contaminated.

The following requirements pertain to information and training:

- Ensure that employers provide appropriate and accessible training to all employees with occupational exposure to blood or other potentially infectious materials. This training must be provided at no cost to the employee and during working hours when an employee is initially assigned to a task where occupational exposure may take place, or within 90 days after the effective date of the Occupational Safety and Health Administration (OSHA) standard, and at least annually thereafter.
- Ensure that employers provide additional training when changes such as modification of tasks or procedures or institution of new tasks or procedures affect the employee's occupational exposure. The additional training may be limited to addressing the new exposures created.
- Ensure that the training program is conducted by a knowledgeable person and contains at a minimum the following elements:
  - An accessible copy of the OSHA standard text and explanation of the standard.
  - A general explanation of the epidemiology and symptoms of bloodborne diseases.
  - An explanation of how bloodborne pathogens are transmitted.
  - An explanation of the employer's exposure control plan and the means by which the employee can obtain a written copy of the plan.
  - An explanation of the appropriate methods for recognizing tasks and other activities that may

involve exposure to blood and other potentially infectious materials.

- An explanation of the use and limitations of methods that will prevent or reduce exposure including appropriate engineering controls, work practices, and personal protective equipment.
- Information on the types, proper use, location, removal, handling, decontamination, and disposal of personal protective equipment.
- An explanation of the basis for selection of personal protective equipment.
- Information on the hepatitis B vaccine, including information on its efficacy, safety, method of administration, the benefits of being vaccinated, and information that the vaccine and vaccination will be offered free of charge.
- Information on the appropriate actions to take and persons to contact in an emergency involving blood or other potentially infectious materials.
- An explanation of the procedure to follow if an exposure incident occurs, including the method of reporting the incident and the medical follow-up that will be made available.
- Information on the post-exposure evaluation and follow-up that the employer is required to provide for the employee following an exposure incident.
- An explanation of the signs and labels and/or color coding required in the OSHA standard.
- An opportunity for interactive questions and answers with the person conducting the training session.

## C.5: Recordkeeping

Provisions for recordkeeping pertain to medical records, training records, availability of records, and transfer of records. The following requirements pertain to medical records:

- Ensure that the employer establishes and maintains an accurate, confidential record for each employee with occupational exposure to blood or other potentially

infectious materials. This record should contain medical information, including but not limited to information on the employee's hepatitis B vaccine status and ability to receive the vaccination, and copies of all medical examination, testing, and follow-up procedure results.

- Ensure that the employer maintains this medical record for the duration of an employee's employment, plus 30 years.

The following requirements pertain to training records for the OSHA bloodborne pathogen standard:

- Ensure that the employer maintains training records for three years. These training records should contain information such as the dates of the training sessions; the contents or a summary of the training sessions; the names and qualifications of the trainers; and the names and job titles of all people attending the training sessions.

The following requirements pertain to the availability of records for the OSHA standard:

- Ensure that all medical records maintained in accordance with the OSHA bloodborne pathogen standard are provided upon request for examination or copying to the subject employee, to anyone having written consent of the subject employee, to the Director, and to the Assistant Secretary.
- Ensure that all training records maintained in accordance with the OSHA bloodborne pathogen standard are provided upon request for examination or copying to employees, to employee representatives, to the Director, and to the Assistant Secretary.

Finally, the requirements pertaining to the transfer of records for the OSHA standard stipulate that the employer complies with the requirements of 29 CFR 1910.20 (h) involving the transfer of records. In addition, if the employer ceases to do business and there is no successor employer to receive and retain records, the employer is required to notify and transmit, three months in advance, all medical and training records maintained under the OSHA bloodborne pathogen standard to the Director.

## **References**

29 CFR Part 1910.1030. (July 1, 1992). Occupational Safety and Health Administration. "Bloodborne pathogens, 29 CFR Part 1910.1030."

## APPENDIX D: VIOLATIONS, NON-CITED VIOLATIONS, AND UNRESOLVED ITEMS FROM THE NRC INSPECTION REPORTS

This section presents specific information about the types of violations, non-cited violations, and unresolved items that the Nuclear Regulatory Commission (NRC) found in conducting initial inspections, special reactive inspections, and follow-up inspections of licensees' fitness-for-duty (FFD) programs from the program's initiation in 1990 through the early spring of 1995. The results of the inspections conducted prior to March 1991 were previously reported in NUREG/CR-5784, *Fitness for Duty in the Nuclear Power Industry* (Durbin et al., 1991). The summary provided here combines the results of these earlier inspections with those of the more recent inspections.

Temporary Instruction (TI) #2515/106 requires that each site undergo an initial inspection to evaluate its compliance with the FFD rule. Initial inspections were conducted by NRC regional inspectors to evaluate implementation of licensee fitness-for-duty programs, to provide information needed to assess the effectiveness of the FFD rule, and to determine whether further FFD rulemaking is needed. Initial inspections were sometimes preceded or followed by special reactive inspections when serious problems were identified or when a number of issues remained unresolved.

In addition to feedback on program strengths, inspectors noted violations, non-cited violations, and unresolved issues that were not being identified or adequately addressed by some licensees. Violations are serious infractions of the rule or minor infractions that were not being identified or adequately addressed by the licensees. Non-cited violations are minor violations of 10 CFR Part 26 that were identified by the licensee and for which the licensee was taking appropriate corrective action. Unresolved items are matters about which more information is required and which are resolved subsequent to the inspection. Unresolved items also include a licensee's failure to adhere to generally accepted practices that have not been made legally binding by the Commission but are expected to be implemented by the licensee. In many cases, the unresolved items involve contradicting interpretations of the FFD rule; this is not unusual during the initial implementation of complex regulations.

The violations, non-cited violations, and unresolved items noted in these inspection reports were distributed across many subject areas of the rule. The following sections describe each of those areas. This information may be useful to NRC licensees in improving their FFD programs.

### D.1: Written Policy and Procedures

Section 26.20 of the rule contains specific requirements concerning written policies and procedures to support fitness-for-duty programs. The development of policies and procedures for a new program is a process that requires substantial effort—and one where revisions and additions are expected as the program matures. In addition, due to the nature of the inspection effort (inspectors reviewed the policies and procedures prior to each inspection), this is an area of program implementation that received particularly careful scrutiny. Hence, it is not surprising that a substantial number of problems in the area of policy and procedures were identified by licensees or during the inspection process. Most of the problems fell into the category of unresolved items. It should also be noted that lack of complete written procedures did not necessarily mean that the program was not functioning correctly.

A violation in the area of policies and procedures occurred when a licensee's ambiguous procedures governing MRO responsibilities resulted in improper handling of a positive test result. One non-cited violation occurred in the area of chain of custody; while processing a simulated blood specimen during the inspection, the on-site testing laboratory technician had temporary custody of the specimen without signing the chain-of-custody form. A second non-cited violation occurred due to licensee failure to maintain a copy of an official policy statement on the use of illegal drugs and the sanctions for such use.

Several unresolved items occurred in the area of written policies and procedures. These items were all very different in nature. One site had two unresolved items in this area: they did not have a verification procedure for identifying couriers who transferred specimens to the certified laboratory, and their procedure for call-ins inappropriately exempted workers from random testing for drugs. At a second site, the MRO job description did not require the MRO to conduct in-person interviews with individuals who tested positive for drugs. At another site, the licensee's chain-of-custody forms provided to the MRO did not include the testing laboratory's internal chain-of-custody forms. A second chain-of-custody issue arose when an inspector determined, based on a review of the MRO's desk (detailed) procedures, that the procedures were lacking in the essential elements to ensure that chain-of-custody issues were properly addressed. As a result of these inadequate procedures, an MRO did not investigate a missing chain-of-custody form and take

## Appendix D

prompt corrective action. The licensee committed to take corrective action in this area.

Another unresolved item occurred as a result of guidance on for-cause testing in one licensee's corporate nuclear security guideline. This unresolved item had three aspects. The first part of the unresolved item concerned the licensee's policy to test for alcohol only when alcohol and not drug use is suspected, rather than to test for drugs and alcohol during any for-cause testing, as required by 10 CFR 26.24(6.c) and Section 2.1(a) of Appendix A. The second part of the unresolved item concerned the licensee's policy of having the MRO determine if drug and/or alcohol testing is required if an employee is observed to be impaired or displaying aberrant or atypical behavior. However, it is the NRC's position that the MRO interpret and evaluate test results, not determine if tests are needed. Finally, the third part of the unresolved item concerned the licensee's policy of allowing an individual to be taken home and not tested if the MRO cannot be contacted. Because this procedure as written would allow a worker to be exempted from a required for-cause test, inspectors identified this procedure as a potential violation of 10 CFR 26.24(a)3, but included the issue as part of the unresolved item.

Another unresolved item concerned a licensee's written FFD procedures which were not sufficiently detailed to provide the MRO with adequate instructions for timely reporting of confirmed positive drug test results. The FFD procedures did not specify time limitations for reporting results to FFD personnel and did not identify FFD contacts during off-hours and holidays. Only normal daytime work phone numbers were provided. As a consequence, the MRO delayed notifying FFD personnel for two days after he confirmed that the test results were positive.

### D.2: Policy Communication and Awareness Training

As stated in Section 26.21 of the rule, licensees are required to provide all workers with training in general policy awareness and must provide additional training in policy awareness for supervisors and escorts. Inspectors noted program strengths in the area of policy awareness training but also noted some deficiencies.

There were no violations in this area. An unresolved item occurred in this area for a licensee that had insufficient requalification training and procedures in place. The licensee had incorporated their fitness-for-

duty training into their general employee training and believed that this met the requirement for requalification training. However, the inspector's review found that this training did not cover all of the areas required in the rule for refresher training. In addition, three unresolved items resulted from licensees who did not have sufficient records for their training programs.

### D.3: Supervisor and Escort Training

Section 26.22 of the rule requires that managers and supervisors be provided with appropriate training to ensure they understand their roles and responsibilities, techniques for recognizing drugs and drug use, behavioral observation techniques for detecting degradation in performance, and procedures for initiating appropriate corrective action. Several violations, non-cited violations, and unresolved items resulted from inadequacies in supervisor or escort training.

One licensee received a violation for failing to provide 20 newly appointed supervisors with appropriate FFD training before or within three months of being assigned responsibilities for implementing the FFD program. At a second site, the licensee failed to provide procedures for supervisory training. In a third case, the licensee failed to administer both initial and refresher training for supervisory personnel. Four violations resulted from licensees' failure to provide initial FFD supervisory training to contract supervisors prior to authorizing their unescorted access to the site, as required in 10 CFR 26.22(c).

In addition, several licensees received non-cited violations when they failed to provide behavioral observation training for employees with supervisory and managerial responsibilities. At one site, 50 supervisors were found to have not been trained within three months of assignment in their behavioral observation responsibilities. At another site, the licensee failed to train contractor supervisors in their roles and responsibilities in implementing a fitness-for-duty program.

Two unresolved items concerned supervisor training issues. One unresolved item concerned a licensee's lack of a good mechanism to track behavioral observation training for newly promoted supervisors to ensure that they are trained within three months of appointment. The licensee committed to do a 100% audit of supervisor training records. The other unresolved item occurred when an Operations Shift Supervisor (OSS) was told by a reactor operator that he had tested positive on a random drug test. The OSS did not attempt to contact either his

management or security/FFD personnel for guidance or consultation regarding the information he was provided by the operator. Had the OSS communicated such information to the appropriate licensee personnel, it is likely that the operator's access would have been removed in a more timely manner.

#### **D.4: Contractors and Vendors**

Section 26.23 of the rule outlines requirements for contractor adherence to licensee FFD policies. One violation occurred in this area because several of a licensee's agreements with contractors did not contain the required fitness-for-duty provisions.

#### **D.5: Chemical Testing**

Section 26.24 of the rule sets forth the required chemical testing program elements that licensees must adhere to. This includes requirements for pre-access, random, for-cause, and follow-up testing. The section of the rule also requires that licensees conform to specific collection procedures, use certified laboratories, and test for particular drugs at minimum cutoff levels as described in Appendix A of the rule. Inspectors identified problems in the area of random and pre-access testing, certified laboratories, and general deficiencies in the specimen collection process.

Two violations resulted from licensee failure to include certain groups of workers in the random testing pool: at one site, the licensee failed to include some personnel from their Emergency Operations Facility in the random testing program; at another site, a licensee's random testing program failed to make a tested individual immediately eligible for another random test. Non-cited violations occurred where several licensees had insufficient random testing during weekends, holidays, and/or backshifts. Another licensee had inappropriately suspended random testing while trying to process a large number of pre-access tests.

One licensee had two violations for inadequate licensee policies regarding pre-access testing. The first violation was for failure to test for alcohol during pre-access testing. The second was for failure to conduct some pre-access drug tests within 60 days prior to the initial granting of unescorted access.

A second area of chemical testing that appeared problematic concerned licensee failure to contract with a DHHS-certified laboratory. This failure resulted in one

violation. In addition, three non-cited violations occurred due to inadequate licensee evaluation of DHHS-certified laboratories. Of the non-cited violations, one occurred because a licensee contracted with a laboratory that was not DHHS-certified. The second resulted from a licensee failure to conduct a pre-award audit of the DHHS laboratory prior to awarding the laboratory a contract for services. In the final instance, a licensee failed to promptly notify the NRC of a false positive blind performance test and instead waited until their 30-day report to alert the NRC of the incident.

Unresolved issues related to laboratory testing included one instance where a certified laboratory did not use GC/MS testing for samples that exceeded the cutoff limits, and another instance where a certified laboratory did not test for 6-monoacetylmorphine (6-MAM)<sup>1</sup> as part of a confirmatory test for presumptive positives for morphine. In the latter case, neither the licensee's MRO nor the FFD program manager were aware that the test was not being performed. Even though the test was not performed, the MRO attested to performing a medical examination on each individual who tested positive for opiates when there was not a valid medical explanation. The licensee agreed to investigate why the 6-MAM analysis had not been conducted by its DHHS-certified laboratory, to take corrective action, and to have the laboratory conduct the 6-MAM test for those samples that were positive without valid medical explanations.

The proper collection and processing of specimens was an issue at several sites. One licensee inappropriately required a blood specimen in addition to urine specimens for some types of tests. Another licensee failed to provide behavioral observation for all personnel subject to testing. At another site, a collection site person failed to require that an observed specimen be given or suggest that an oral temperature be taken to substantiate the temperature of a urine specimen that was found to be outside the prescribed temperature range, per Section 2.4 of Appendix A. One licensee received a violation and another licensee received a non-cited violation when each licensee failed to collect a second breath specimen from individuals after a reading of 0.000 was obtained from the first breath specimen. An unresolved item concerned a licensee that had an inappropriate requirement for a blood specimen. Another unresolved item occurred as a result of a licensee's past practice of not submitting blind performance test samples for the complete panel of drugs under the Substance Abuse and Mental Health Services Administration (SAMHSA) guidelines.

## Appendix D

At two separate sites, violations occurred as a result of licensees allowing FFD staff members to collect specimens from FFD supervisors and allowing co-workers to collect specimens for other persons directly involved in administrating the FFD program. This practice was found to be contrary to 10 CFR Part 26, Appendix A, paragraph 2.3(1), which reads that, "Supervisors, co-workers, and relatives of the individual being tested shall not perform any collection, assessment, or evaluation procedures."

One licensee received a violation due to failure to ensure that the designated collection site was secure from unauthorized personnel at all times. This violation was resolved shortly after the inspection. A final unresolved item occurred when an inspection report noted an inadequate procedure for notifying employees to report to testing. In this case, the licensee permitted supervisors to postpone a test until later in the day if the test posed an inconvenience, but did not limit the amount of time between notification and testing, presenting a potential for abuse.

In addition, there was one non-cited violation for an administrative error that resulted in a worker with a presumptive pre-access positive drug test being granted unescorted access.

### **D.6: Employee Assistance Programs (EAP)**

Section 26.25 of the rule requires that licensees maintain an employee assistance program (EAP) to achieve early intervention and to offer confidential assistance. Employee assistance program services may be provided by utility employees or by contract. Many program strengths were identified by inspectors including strong employee awareness and use of the programs and creation of peer support groups. Nevertheless, problems were identified at some sites.

One unresolved item concerning EAPs resulted from a licensee whose FFD policy treated worker self-referrals to the EAP as a positive drug test. In some cases, the self-referrals were equated to a positive test even when an actual test was not performed. Section 26.25 of the rule requires that EAPs be designed to achieve early intervention and to provide for confidential assistance. This practice is not conducive to achieving early intervention and does not provide for confidential assistance. Another unresolved item occurred when a licensee failed to require the EAP to disclose employees

who are determined to be a threat to themselves or to others.

### **D.7: Management Actions and Sanctions**

Minimum sanctions for FFD violations are specified in Section 26.27 of the rule. These sanctions require that impaired workers be removed from activities that require unescorted access to protected areas. The rule also requires that, following a first confirmed positive test result for illegal drugs, a worker must be immediately suspended from activities that require unescorted access for a minimum of 14 days while an assessment is made and plans for treatment follow-up and future employment are developed. A second confirmed positive test results in suspension from unescorted access for a minimum of 3 years. Sanctions are less specific regarding positive tests for alcohol. Licensees must, however, provide sanctions for confirmed misuse of alcohol and legal drugs that are sufficient to deter abuse. This section of the rule further specifies that management must obtain a written statement from the individual and, as applicable, complete a suitable inquiry to determine if that individual had in the past tested positive for drugs or alcohol or been denied unescorted access.

One violation and one non-cited violation resulted from licensee failure to conduct suitable inquiries for workers requiring unescorted access. The violation occurred because a licensee equated the written statement from the individual concerning his or her history of substance abuse as sufficient inquiry, and did not conduct any further suitable inquiry activities. The non-cited violation occurred when the licensee failed to obtain suitable inquiry information for 33 contractor employees prior to granting them access. The employees had this inappropriate access for nearly two months. A third problem concerned the perception by employees that a presumptive positive drug screen would result in negative management action.

### **D.8: Appeals**

Section 26.28 of the rule requires that licensees establish a procedure for both licensee and contractor employees to appeal a positive alcohol or drug test. The appeals process was not the subject of any violations or non-cited violations. There were, however, two unresolved items that pertained to this part of the rule.

One unresolved item concerned a licensee's practice of separating every sample into two split-samples and marking one of the two as an "appeal" sample to be sent to the licensee's DHHS-certified laboratory if there was reason to suspect adulteration or dilution of the sample, or if the MRO determined a test to be confirmed positive. The inspectors determined that under the licensee's program, the appeal sample would not be sent to a second off-site laboratory, nor did the licensee's FFD procedures specify how a suspect individual was to be made aware of appeal rights for a confirmed positive test result. In another case, a licensee failed to provide workers with access to lab records regarding test results.

## D.9: Protection of Information

Section 26.29 of the rule states requirements for the protection of fitness-for-duty information. Inspectors noted two unresolved items concerning inappropriate personnel access to unconfirmed positive test results. At one licensee site, the supervisor of security received notification of positive drug test results prior to MRO confirmation. At the second site, the licensee's procedures inappropriately required that laboratory test results be sent to the FFD administrator.

## D.10: Recordkeeping Requirements

Section 26.71 of the rule requires licensees to keep records of confirmed positive test results, and to collect and compile program performance data every six months. One violation occurred when a licensee failed to maintain a permanent record book. In another case an inspector determined that a licensee's permanent record books were not well maintained: the licensee could not provide permanent record books for the first four months of program implementation; collection site personnel were not signing the currently existing permanent record books next to the identifying information; and one individual who provided a sample was never recorded in the permanent record book.

One unresolved item occurred when an inspector was unable to reconcile a licensee's records completely against testing data that were provided. The MRO-confirmed positives and the blind performance data were found to be in error. The licensee agreed to revise and resubmit the performance data it had submitted to the Commission for the first year of program implementation.

## D.11: Reporting Requirements

Section 26.73 of the rule addresses reporting requirements, and requires licensees to inform the Commission of significant fitness-for-duty events.

One licensee received a non-cited violation as a result of failure to receive test results within five working days from the Health and Human Services-certified laboratory with which it had contracted. The same licensee received another non-cited violation for failure to have its Medical Review Officer (MRO) review and report FFD test results to licensee management within 10 days, as required by 10 CFR 26.24(e). Both of these issues were found during the licensee's self-audit and were resolved by contracting with a new laboratory and a new MRO. In another case, a licensee failed to promptly report to the NRC that a supervisor had tested positive for alcohol.

An unresolved item concerned a licensee's failure to report to the NRC a confirmed positive test of a contract foreman performing supervisory duties. There was some question about whether the contract foreman's duties categorized him as a supervisor, as that term relates to the reporting requirements of 10 CFR Part 26.

Section 2.7(e)(2), "Preliminary Initial Test," in 10 CFR Part 26, Appendix A, states that licensees may specify more stringent cut-off levels for illegal drugs and alcohol, but that the results of screening tests must be reported for both the NRC levels and the licensee's levels in such cases. One unresolved item occurred when a licensee failed to report the results of onsite screening of samples analyzed at the NRC guideline cutoff levels. This occurred because screening was in fact not being performed at the NRC cutoff levels. However, the licensee did report the results of its DHHS-certified laboratory's analyses at both the licensee's and the NRC's cutoff levels. The inspector determined that this issue required further evaluation.

## D.12: Audits

Section 26.80 of the rule requires licensees to audit their fitness-for-duty programs nominally every 12 months. In addition, licensees retain responsibility for the effectiveness of contractor and vendor programs and the implementation of appropriate corrective action. Inspectors noted many strengths in licensee's audit

## Appendix D

programs and their ability to identify program weaknesses. Only two inspection reports noted that

licensees had failed to implement an effective FFD audit program.

## Endnotes

1. 6-monoacetylmorphine (6-MAM) is now referred to as 6-acetylmorphine (6-AM). The term "6-MAM" was in use when this inspection was conducted.

## Appendix D

### References

10 CFR Part 26. (January 1, 1990). U.S. Nuclear Regulatory Commission. "Fitness-for-duty programs." *U.S. Code of Federal Regulations*.

Durbin, N., Moore, C., Grant, T., Fleming, T., Hunt, P., Martin, R., Murphy, S., Hauth, J., Wilson, R., Bittner, A.,

Bramwell, A., Macaulay, J., Olson, J., Terrill, E., and Toquam, J. (1991b). *Fitness for Duty in the Nuclear Power Industry: A Review of the First Year of Program Performance and an Update of the Technical Issues (NUREG/CR-5784)*. Washington, DC: Nuclear Regulatory Commission.

## APPENDIX E: METHODOLOGICAL STRENGTHS AND LIMITATIONS OF SELECTED MAJOR DRUG USE SURVEYS

### E.1: Methodological Strengths and Limitations of NHSDA

As acknowledged in the National Household Surveys on Drug Abuse (NHSDA) reports (SAMHSA\_NHSDA91, 1993:6; SAMHSA\_NHSDA92, 1993:11-15) and an independent NIDA publication (Turner, Lessler and Gfroerer, 1992), NHSDA has certain properties that could limit its accuracy or generalizability. First, because it relies on self-reported behavior, the survey depends on the memory and veracity of respondents. Given the current stigmatizing atmosphere regarding drug use, self-reporting undoubtedly results in an undercounting of drug use, particularly illegal drug use. However, the NHSDA survey is designed to maximize the respondent's anonymity and privacy, in an attempt to minimize such undercounting:

Interviewers attempt to conduct interviews in a private place, away from other household members . . . the answers to sensitive questions (such as those on illicit drug use) are recorded by the respondent and not seen or reviewed by the interviewer. [The answer sheets are then] placed by the respondent in an envelope, which is sealed and mailed back to the contractor . . . with no respondent name or address information (SAMHSA\_NHSDA92, 1993:11).

Second, because NHSDA relies on voluntary participation in a survey asking sensitive questions, its results may not fully represent the population sampled. The NHSDA survey has a fairly high refusal rate (varying from 16-23 percent; Hser et al., 1992); in addition, participating respondents can refuse to answer specific questions. Third, NHSDA is a cross-sectional, rather than longitudinal, survey, meaning that it surveys a new set of respondents each year rather than following the same individuals over time. Fourth, it excludes a small portion (about 2%) of the U.S. population; namely, persons "institutionalized" (e.g., in prisons, nursing homes, residential treatment centers, and college dormitories), homeless, or in active military service. While these subgroups are relatively small, some of them are known to contain significant numbers of current or former drug users. While their omission does not greatly affect the estimates of usage of more prevalent drugs (e.g., alcohol), it probably does affect the estimates for less prevalent drugs (e.g., heroin).

A recent report by the General Accounting Office (GAO/PEMD-93-18, 1993) assessing the methods of data

collection used in several major federal surveys of drug use was critical of NHSDA for both its reliance on self-reports and its omission of institutionalized and homeless persons, an omission which may result in an underrepresentation of groups most at risk from illegal drugs, especially users of heroin and cocaine.

In addition, Hser et al. (1992) point out that surveys such as NHSDA, while surveying each respondent's recency and frequency of use of various drugs, fail to assess two important patterns in the respondent's drug use: 1) concurrent use of drugs and 2) trends of escalation or deescalation in drug use. Discovering an individual's concurrent use highlights particularly dangerous forms of drug abuse (e.g., alcohol and barbiturates; heroin and cocaine), while noting an individual's trend toward greater (or lesser) use of drugs provides useful information on the dangers of drugs whose use typically shows rapid escalation after initiation (e.g., heroin, crack cocaine).

Despite these numerous limitations, NHSDA findings still provide reasonable, though somewhat conservative, estimates of drug use prevalence in the United States. A Peer Review Committee established recently by SAMHSA's Office of Applied Studies (to analyze unexpected NHSDA findings regarding drug use among African-Americans) concluded that "the design and procedures for sampling, weighting, editing, and imputing of [NHSDA] survey results are statistically sound" (SAMHSA\_NHSDA92, 1993:15). In addition, as is noted in Section 3.2.2.1, recent NHSDA findings on trends in drug use tend to resemble those of another self-report survey, the Monitoring the Future (MTF) Survey of college students and young adults, suggesting again that these self-report surveys produce reasonable, though sometimes conservative, estimates of drug use prevalence in the sampled populations.

### E.2: Methodological Strengths and Limitations of MTF

Like NHSDA, the Monitoring the Future (MTF) survey relies on self-reporting by respondents, and is thus susceptible to many of the same methodological limitations. In addition, unlike NHSDA, Monitoring the Future is restricted to young adults, reducing the generalizability of its findings. Finally, selective case attrition in the form of high school dropouts in the original sample introduces bias in the subsequent College Students and Young Persons surveys (Hser et al., 1992).

### **E.3: Methodological Strengths and Limitations of DAWN**

By collecting data on drug abuse-related emergency room (ER) episodes, DAWN avoids some of the problems encountered in self-report surveys. The drug users in DAWN data cannot refuse participation (unless they avoid emergency care), cannot be as selective in which types of drug use to reveal, and need not remember which drugs they used.

However, DAWN also has limitations, probably more consequential than those of the self-report surveys. First, DAWN's sample, limited to ER episodes, is far less representative of the U.S. population than that of NHSDA, or even MTF. Second, the most frequent self-reported motive for drug use in 1991 episodes was suicide (SAMHSA\_DAWN\_92, 1992), suggesting that DAWN data are reflective of more problems than just drug abuse (e.g., psychological distress). (Still, such data suggest the types of drugs available, and perhaps

previously used, by the individual.) Third, alcohol is reported only when present in combination with other drugs; thus, the category "alcohol-in-combination." As a result, alcohol use is undoubtedly undercounted in DAWN data (though perhaps not significantly, since the presence of alcohol and even a common over-the-counter drug such as aspirin would result in an "alcohol-in-combination" [as well as "aspirin"] mention). Fourth, up to four drug categories, in addition to "alcohol-in-combination," can be reported in an episode. Consequently, DAWN data include mentions of drugs not necessarily contributing to the medical emergency. Fifth, identification of specific drugs is based on self-reported usage and usage identified by medical personnel, usually through on-site assessment and not laboratory tests (Hser et al., 1992). Consequently, identification of specific drugs may be imprecise or incomplete, and identification methods and precision may vary across data collection sites. Finally, DAWN data are episode-based, not person-based, so that repeat visits by the same individual bias the results.

## References

General Accounting Office. 1993 (June). *Drug Use Measurement: Strengths, Limitations, and Recommendations for Improvement* (GAO/PEMD-93-18). Washington, D.C.: GAO.

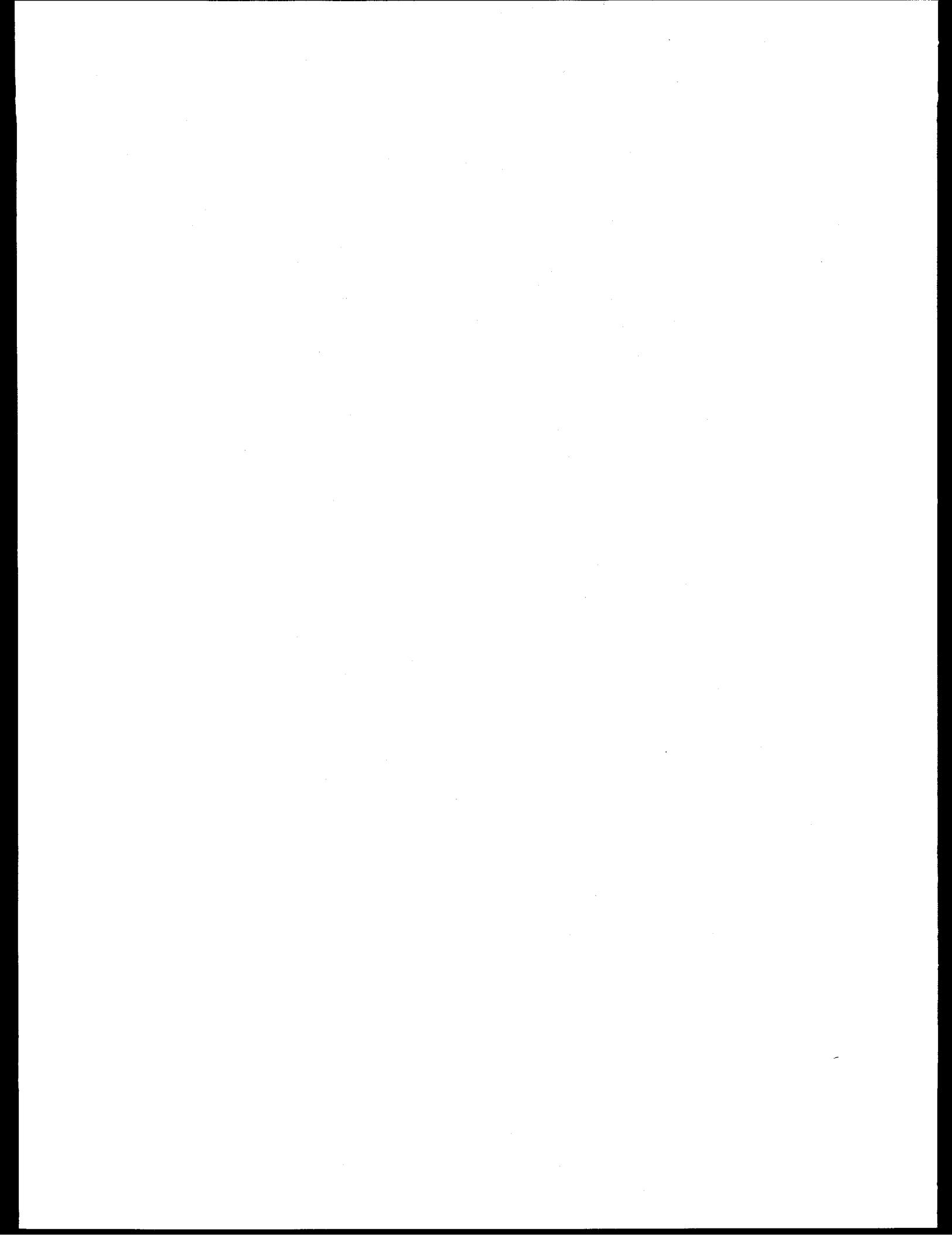
Hser, Y., Anglin, M.D., Wickens, T.D., Brecht, M., and Homer, J. 1992 (May). *Techniques for the Estimation of Illicit Drug-Use Prevalence: An Overview of Relevant Issues*. Washington, D.C.: National Institute of Justice.

SAMHSA\_DAWN\_92. 1993 (September). *Estimates from the Drug Abuse Warning Network: 1992 Estimates of Drug-Related Emergency Room Episodes*. SAMHSA, Office of Applied Studies, Drug Abuse Warning Network.

SAMHSA\_NHSDA91. 1993. *National Household Survey on Drug Abuse: Main Findings 1991*. SAMHSA, Office of Applied Studies, National Household Survey on Drug Abuse.

SAMHSA\_NHSDA92. 1993 (June). *Preliminary Estimates from the 1992 National Household Survey on Drug Abuse: Selected Excerpts*. SAMHSA, Office of Applied Studies, National Household Survey on Drug Abuse.

Turner, C.F., Lessler, J.T. and Gfroerer, J.C. (Eds.). 1992. *Survey Measurement of Drug Use: Methodological Studies*. Washington, D.C.: U.S. Department of Health and Human Services.



## APPENDIX F: SUBVERSION ISSUES RELEVANT TO THE PROPOSED REVISIONS TO 10 CFR PART 26<sup>1</sup>

Subversion in drug testing programs has become a key issue. Because successful subversion can not be identified, the nature and extent of subversion must be extrapolated from evidence of unsuccessful subversion attempts. While this provides little in terms of hard evidence of the total extent of the problem, it does provide a basis for an understanding of the level of subversion that exists. This paper addresses the nature and extent of efforts to subvert drug testing generally and within the nuclear industry in particular.

### 1. What is the nature of the subversion problem generally?

**What are the key types of subversion and how are they prevented?** There are a number of ways that drug testing can be successfully subverted.

- Test avoidance. This is generally prevented through administrative procedures (e.g., not providing advance notification of random testing).
- Clearing the system of drug metabolites prior to testing. This is done by delaying the test and/or flushing the system with water. Sometimes this is augmented by ingesting substances to promote rapid flushing. Flushing can be prevented by minimizing the time between notification and testing. In addition, flushing can be deterred by decreasing its effectiveness (e.g., by testing at the DHHS laboratory level of detection (LOD), see discussion below).
- Providing a surrogate specimen. This can be done in two ways. One way is to provide a surrogate (substitute) specimen at the collection site. This is prevented by assuring that the specimen is at body temperature and that the individual does not have the opportunity to obtain and bring a surrogate specimen to the collection facility. The other way is to substitute a clean specimen during storage, transportation to the laboratory, or at the DHHS laboratory. This is prevented by rigorous chain of custody and security procedures and by assuring the reliability of all personnel involved in the testing process.
- Diluting or adulterating specimens prior to testing. This is accomplished by reducing the concentration of metabolites in the specimen below the DHHS cutoff level (dilution) or by adding substances to the specimen that interfere with the chemical analysis (adulteration). These are prevented by minimizing the time and opportunity to subvert, eliminating access to water for dilution (e.g., putting bluing agents in toilets), by preventing individuals from bringing water or adulterants with them to the testing area, and/or by testing specimens for the presence of adulterants. Testing specimens for the presence of adulterants can be done in a variety of ways, from visual examination for color or clarity to sophisticated tests for commercially available adulterants such as UrinAid®.

All of these techniques are identified routinely at testing facilities. Examples of each of these types of subversion are found in the compilation of reported instances of subversion in the nuclear industry (see Attachment A).

**What evidence exists about the extent of the subversion generally?** Although there is little hard evidence of successful subversion (if it is successful, we don't know about it) there are clear indications that it is pervasive. For instance:

- There is a striking consensus among individuals working in drug testing—both those working in testing programs for addicts and those working in testing programs in military and employment settings—that subversion is a serious problem.
- Laboratory analyses of specimens from the Department of Transportation program show that there was a significantly higher proportion of dilute specimens among individuals covered by a program (7.6% of specimens of those in the program were dilute) than among those not covered by a testing program (3% of the control group were dilute). Since the perceived reason for the increase in dilute specimens among those covered by a program is attempted flushing, this provides some evidence of attempted subversion. This is further supported by the results of level of detection (LOD)<sup>2</sup> testing of these dilute specimens. At LOD testing levels, 9.9% of the dilute specimens were positive—a higher positive rate than in the DOT testing program generally. At DHHS cutoff levels, 1.2% of the dilute specimens tested positive for drugs—a lower rate than in the DOT testing

## Appendix F

program generally. Since diluting specimens subverts testing by reducing the concentration of metabolites in the specimen (thus making it test negative at the cutoff level) those specimens testing positive at LOD but negative at DHHS cutoff levels were potential successful subversion events. The DHHS Drug Testing Advisory Board was advised, based on the results of this study, that identifying dilute specimens and subjecting them to LOD testing would double the positive rate, and that using LOD testing on all specimens would quadruple the positive rate.

- The prevalence of dilution and adulteration was reported by Dr. Gary Hemphill of MEDTOX Laboratories. The number of suspect specimens identified by his laboratory has steadily risen from about 3 percent a few years ago to a current rate of about 5 percent of specimens received for testing. Dr. Hemphill also reported that, although adulteration is a technique frequently detected by the laboratory, flushing is the technique used the most and is successful at producing the most false negatives. Other laboratories that use a process to determine validity also report 5 percent or more of specimens are suspect.

### ***What are the consequences of leaving subversion undetected?***

- Successful subversion substantially decreases drug testing program effectiveness. Testing that can be subverted fails to detect and fails to deter.
- Basic subversion techniques such as dilution through flushing that are advocated widely and publicly should be expected to be used by anyone worried about a positive drug test. They are effective—particularly at DHHS cutoff levels. Hence, as individuals become more adept at these techniques, the only positive test results will be from the uninformed and/or the unlucky.
- Successful subversion erodes support for testing programs. Being tested is not a positive experience for individuals who are not drug users. The only reward is the assurance of working in a drug free workplace (which is reported by workers as a plus). Being tested but knowing that users are not identified (either in reality or in perception) erodes any support.
- Subversion techniques are well known to individuals using illegal drugs. Sophisticated techniques are more likely to be employed by those who use illegal drugs chronically (i.e., addicts); less sophisticated techniques—such as flushing—are likely to be employed by all users (both addicted and “recreational” users).

It is therefore critical that at least the easiest forms of subversion be prevented, that subversion attempts be identified and acted upon, and that the NRC demonstrate that these activities will not be tolerated.

### **2. What is the nature of the subversion problem within the nuclear industry and how is it being addressed?**

***What evidence exists about the extent of the subversion among individuals covered by 10 CFR Part 26?*** While the prevalence of subversion by the nuclear powerplant workforce is unknown, the staff does have information about certain detected acts of subversion. There are indications that there are a substantial number of workers who are evading detection of their substance abuse through subversion.

- In a recent meeting with several licensee fitness-for-duty program representatives, the staff inquired about licensee experience with testing subversion in their programs. Virtually all of the licensee representatives acknowledged that they rather routinely discover acts of attempted testing subversion. This conversation indicated that instances of subversion are by no means unique or isolated occurrences in licensee programs. On the contrary, the licensee FFD program staff present seemed to consider subversion to be an all too common occurrence that threatens overall FFD program effectiveness.
- During a telephone conference call sponsored by Bensinger and Dupont (about one-third of licensees participated), 3 licensees stated that their efforts to identify suspect specimens had produced a dramatic increase in the number of workers being identified as substance abusers.

- San Onofre has reported that 55 percent of its positives were from dilute specimens in 1993. In that year, San Onofre reported 71 positives from a total of 7,458 tests or a positive rate of 0.95 percent, a rate substantially above the industry average of 0.62 percent. If San Onofre had not taken special measures, that licensee's FFD staff report that only 32 positives would have been detected. San Onofre's experience is another indication that requiring those licensees that do not go beyond the minimum anti-subversion measures to do so would lead to detection of a substantial number of substance abusers who are currently evading detection. In 1995, San Onofre reported that the percentage of its positive test results from dilute specimens was 30 percent.
- There have been a significant number of incidents of subversion reported by the industry. These are presented in Attachment A.

Note that relatively few licensees currently take anti-subversion measures beyond the minimums required by the rule and few licensees now report such events to the Commission. Hence, the instances of subversion reported in Attachment A represent only a partial list of the subversion events that have been detected during the past five years. Also, because successful subversion is not detected, these cases of known subversion are only a subset of the total subversion events. Many workers caught in attempting subversion have admitted that they have successfully subverted the testing process in the past.

***How is subversion in the nuclear industry being addressed?*** The current rule requires some efforts to prevent all types of subversion. These include requiring that random testing be unpredictable and unannounced (e.g., to prevent avoidance of testing), that specimen temperature be checked (e.g., to prevent substitution of specimens), and that specified minimum precautions be taken to ensure specimens are not adulterated or diluted. The proposed rule revisions incorporate some changes to requirements to make them more effective and/or less costly based on industry experience. For example,

- “unpredictable and unannounced” is clarified to assure that individuals working nights and weekends are at risk of being tested,
- the range of acceptable temperatures is narrowed based on the experience of San Onofre, Washington Nuclear, and other licensees that a narrower temperature range is effective and does not create testing problems, and
- the consequences of identifying a dilute specimen have been revised in light of industry experience and improved technology in order to reduce the impact on the industry and on individuals and to increase the effectiveness of detecting drug abuse. Under the revised regulation the original dilute specimen would be tested at LOD—substantially reducing the need for second, observed collections as well as deterring drug users from attempting to subvert the testing process through flushing.

These, and other, changes to the revised regulation would not address all attempts at subversion, it would increase the probability of detecting drug use and would provide a strong message to the industry that the NRC is serious about the program.

## Appendix F

### ATTACHMENT A: ATTEMPTS TO AVOID DETECTION OF SUBSTANCE ABUSE IN THE NUCLEAR POWER INDUSTRY

Since licensees are not required to report instances of subversion by the general workforce, evidence of subversion is limited to such acts by supervisors and licensed operators, flavored with occasional verbal reports on selected incidents involving the general workforce. Obviously, there is no data on successful acts of subversion.

10/9/90 FITZPATRICK LICENSED OPERATOR

Specimen provided for random testing had low temperature. SRO refused to provide second specimen. SRO was removed for 14 days and referred to EAP. SRO admitted during investigation by OI in April 1991 that, after being notified he was selected for random testing he retrieved a surrogate sample from his locker, a practice he had started in August 1990. On his way to the collection facility, he went to a men's room to heat the specimen. SRO admitted to use of cocaine since 1977 and occasional use of "speed."

SRO had tested positive for cocaine in August 1988.

NOTE: NRC investigator obtained admissions of substance abuse that licensee's medical staff was unable to obtain (56 FR 22020). ASLB involvement ended when SRO tested positive on 1/21/92 during NRC-mandated followup test.

8/8/91 CALVERT CLIFFS SUPERVISOR

Local sheriff received tip that marijuana was being grown in the owner-controlled area. Patch was located and placed under surveillance.

Security supervisor showed up and was arrested. Since he had been tested several times with negative results, the licensee investigated. Supervisor admitted submitting surrogate samples obtained from his 13 year old son.

Comment: Irony of case was that tip was on another marijuana patch.

11/22/91 SAN ONOFRE LICENSED OPERATOR

Delayed appearance at collection facility, after 2 hours, specimen provided with low specific gravity and temperature (licensee uses narrower range with low of 96°F). Second specimen was collected 4 hours later. Quantitative results of confirmation tests requested because of suspicious circumstances. Second specimen was determined positive for marijuana and cocaine (just below cutoff levels but showed signs of dilution). Specimens were evidence of in vivo dilution (flushing).

Operator had provided specimens in August 1991 that were low in S.G. and borderline temperature.

12/10/91 VOGTLER FFD PROGRAM STAFF

Contract medical support employee returned to collection site after hours and substituted a surrogate specimen. Detected because of differences in specific gravity between the specimen tested by the laboratory and the employee's specimen which had been tested on site immediately after collection.

5/20/92 SAN ONOFRE LICENSED OPERATOR

Provided specimens less than required volume. Quantitative results of confirmation tests requested because of suspicious circumstances. Specimen was determined positive for cocaine (just below cutoff levels).

Comment: Providing low volume specimens buys time to increase probability of successful flushing.

During 1992 BEAVER VALLEY CONTRACT ENGINEER

Specimen provided for pre-access test had low specific gravity. A second specimen collected under observation was positive. Licensee determined that individual had a fake bladder strapped to the inside of his thigh. By holding legs together, he was able to warm surrogate specimen to 92° F.

During 1993 NINE MILE POINT EMPLOYEE

Specimen provided for random testing had been adulterated with clorox.

3/9/93 INDIAN POINT 3 SENIOR NUCLEAR PRODUCTION TECHNICIAN

Specimen provided for random testing had low temperature. Licensee determined that employee had attempted to submit surrogate sample. OI Report 5-93-014 contained employee's admission that he had successfully submitted bogus (surrogate) urine samples an unknown number of times before being caught.

Note: This has become a precedent-setting enforcement case. OE wants to issue an order to the individual but is willing to accept the licensee's termination of employment provided proposed changes to the rule adequately address subversion of the testing process.

11/23/93 INDIAN POINT 3 LICENSED OPERATOR

Specimen provided for random testing had low temperature. SRO provided second (witnessed) specimen, subsequently determined to contain marijuana and cocaine. SRO resigned. First specimen was a surrogate sample.

NOTE: SRO had tested positive for marijuana on 8/9/90, but had 4 negative test results since.

6/20/94 BROWNS FERRY SUPERVISOR

Specimen provided for random testing had questionable specific gravity and creatine. Further analysis determined that specimen was not human urine and was highly diluted.

Note: Licensee did not report this event under §26.73 because examples in rule did not include subversion.

## Appendix F

1994 & 1995 NINE MILE POINT

CONTRACTORS

Specimen provided for pre-access test was 102° F. Second specimen collected under observation was positive. Licensee concluded initial specimen was a surrogate sample.

Three or four cases (one as recently as 9/6/95 and another in February 95) in which specimens provided for pre-access test had low specific gravity. Second specimens collected under observation were positive for THC. Licensee concluded these cases were the result of flushing.

Note: Licensee believes flushing is a common practice that should be addressed by the NRC and all licensees.

12/94

LIMERICK

EMPLOYEES AND SUPERVISORS

Licensee received allegation that a number of employees, including supervisors, were using drugs. By February 1995, the licensee's investigation had substantiated the allegations; this was only accomplished through the tenacious efforts of a skilled investigator. Investigation determined that 3 supervisors, 9 PECO employees, 2 contractor employees, and 1 former PECO employee were using drugs. (Three supervisors and 8 PECO employees admitted using drugs.) Three (1 supervisor) admitted to hydrating prior to a test to subvert the testing process; all others claimed knowledge of subversion techniques, especially flushing, but denied such practices because random testing was so infrequent that they have never been selected shortly after use. It should be noted that the group apparently adopted the practice of drinking copious amounts of water after drug use; their specimens were frequently low in specific gravity. Investigation also determined that 1 employee sold drugs and 3 distributed drugs.

Note: Licensee considered raising the random testing rate, but recognized that would not solve the problem. The time between notification and collection was reduced to 1 hour. Specimens with low S.G. are now tested to LOD, and several positives have resulted. The FFD program staff is now more vigilant for indications of subversion.

Comment: Based on statements made during the interviews, the investigator concluded that the hydration technique is common knowledge and not limited to PECO's sites. The subjects conveyed confidence that they could avoid detection through hydration.

During 1994

COMMONWEALTH EDISON

CONTRACTOR

Specimen provided for pre-access test was cold. Second specimen collected under observation. Both specimens were positive for opiates and cocaine.

Comment: Licensee concluded that the contractor had obtained the first specimen from himself during a time when he believed he was free of drugs. Individual spent a very brief period of time providing initial specimen, an indication of possible subversion monitored by COMED.

3/95

WASHINGTON NUCLEAR 2

CONTRACTORS

After discovering 4 incidents of substitution (submitting surrogate sample), when specimen temperatures were out of range, the licensee decided to narrow the allowable temperature range by raising the lower temperature to 94° F. The licensee screened the files of 929 personnel tested recently and identified 69 who had provided specimens with temperatures below the new standard. Of the 63 retested with limited notification time (6 were unavailable), there were 2 more attempted substitutions, 2 confirmed positive test results, and 1 refusal to test. In addition to narrowing the temperature range, the licensee now requires the selected person to proceed to the collection facility immediately upon notification.

8/95

INDIAN POINT 3CONTRACTOR

Specimen provided for random test was below 90° F. Second specimen was collected; both were positive for THC. Licensee concluded initial specimen was a surrogate sample.

**OTHER VERBAL, NON-DOCUMENTED EXAMPLES**

Other examples of verbal reports from licensees to NRC staff on selected attempts of subversion, based upon staff memory, include:

- One specimen was found to contain antifreeze
- Several specimens were found to contain soap
- Several specimens were found to contain Mountain Dew
- Employee submitted a surrogate sample obtained from son and kept in truck.
- Contract worker submitted a surrogate sample obtained from fellow worker. Because of suspicious nature of collection, licensee collected a second specimen under direct observation. First specimen was positive; second was negative.
- Surrogate sample obtained from a horse.

Some other adulterants discussed at meetings of the DHHS Drug Testing Advisory board include: Aspirin to lower rates of absorbency during testing, vitamin B to provide a yellower dilute specimen, oxidizing agents such as bleach, vinegar, Visine® eye drops, etc.

## Appendix F

### Endnotes

1. This appendix was prepared by NRC staff and contains a listing of event reports relevant to subversion. The event reports were provided to the NRC by licensees.
2. DHHS certified laboratories are capable of reliably identifying the presence of drugs at much lower levels of concentration than are specified in the DHHS guidelines (and by 10 CFR Part 26). It is our understanding that DHHS requires that they have a level of detection that is no higher than 50% of the specified cutoff level, and that most DHHS certified laboratories have a substantially lower LODs. Specimens found to contain drugs at LOD or above provide technically acceptable evidence of exposure to drugs.

## APPENDIX G: SUMMARIES FOR PERFORMANCE-BASED TESTS IDENTIFIED IN THE PUBLISHED LITERATURE

This appendix presents summaries for tests identified in the published literature. The table is arranged by the ability construct measured by the test. The following information appears for each test:

- **Measure/Construct:** Contains the type of measure or name of vendor-designed test apparatus and the factor(s) measured (e.g., cognitive skill, attention, short-term memory).
- **Method of administration:** Provides a description of how the test was administered (e.g., computer-administered, paper-and-pencil, apparatus).
- **Administration time/scoring:** Includes the total test administration time which includes practice and instructions, number of items per test, and length of each test.
- **Use of measure:** Records whether the test was used in an experimental or field study and describes the setting.
- **Predictive/Construct Validity:** Summarizes supporting information about the test such as correlates with other measures (construct) and correlates with job performance (predictive).
- **Sensitivity:** Indicates whether subjects are fit for duty or not based on the effect of some experimental treatment (e.g., alcohol use, fatigue, drug use).
- **Reliability:** Provides information on the method of statistical reliability (e.g., internal consistency, alpha, stability, reliability efficiency, interrater reliability, stabilization time) and an estimate of reliability.
- **References:** Lists the authors from which the information was gathered.

**Cognitive Task: Attention/Divided Attention**

**Appendix G**

Measure/Construct	Method of Administration/Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Altered/Divided attention, dual task	Computer 5 minutes	Experimental test using occasional drug and alcohol users	Not reported	For alcohol: $F=7.41$ $p<0.01$ , amphetamines: $F=5.21$ $p<0.05$ , and for diazepam $F=5.15$ $p<0.05$ . Sensitive to marijuana, but no F value given.	Not reported	Kelly, Fottin, Emurian and Fischman (1993)
Alpha-Numeric Visual Vigilance Task/Sustained visual attention	Computer 30 minutes/Ipsative*	Experimental test using military personnel	This task is closely related to the continuous performance task in measuring sustained attention performance	Authors infer sensitivity to sleep loss and reference Townsend and Johnson (1979)	Not directly demonstrated	Perez, Mastline, Ramsey and Urban (1987)
Dichotic Listening Task/Auditory selective attention	Computer and headphone 36 27-second test trials and 6 practice trials for a total test of approximately 20 minutes/Ipsative*	Experimental test using military personnel	High degree of predictive validity for success in flight training, especially in flight performance. Also significant correlation with a rod and frame task, embedded figures task, a choice-reaction time task, and a complex reaction time task.	Not directly demonstrated	Not reported	Perez, Mastline, Ramsey and Urban (1987)
Divided Attention Task/Attention, psychomotor	Computerized driving simulator Not reported/ Ipsative*	Experimental test involving truck drivers	Validity was assumed because of the similarity between a driving simulator and real-life driving	Sensitive to impairment due to fatigue	Not reported	Allen, Stein and Miller (1990)
Divided Attention Task/Divided attention, psychomotor	Computer 12 minutes per trial/ Ipsative*	Experimental test involving moderate alcohol and marijuana users	Not reported	Impairment due to alcohol and marijuana	Not reported	Burns and Moskowitz (1981)
FFM Mini-Mental Test/Attention	Paper and pencil 15 to 25 minutes/ Normative*	Experimental test using pilots from flying schools and volunteers from an outpatient mental health clinic	Not reported	FFM showed sensitivity of 0.55 in detecting medically classified impairment (e.g., head trauma)	Not reported	Stokes, Bainich and Elledge (1991)

\* Determined by reviewers.

Cognitive Task: Attention/Divided Attention (Continued)

Appendix G

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Memory Search-Tracking Combination/ Short-term working memory, psychomotor	Computer	90 seconds per trial + 30-second break between trials for total of 120 seconds per trial/Ipsative*	Experimental test using military personnel	Not reported	Not reported	Authors infer test-retest reliability using single task reliability data	Perez, Masline, Ramsey and Urban (1987)
Peripheral Vision Test/Perception, attention	Apparatus	114, 20-second test trials separated by 5- second intervals and 2, 5-minute rest periods for a total test time of about 1 hour/ Ipsative*	Experimental test involving occasional marijuana users	Not directly demonstrated	Significant for all marijuana doses tested	Not reported	Moskowitz, Sharma and McGlothlin (1972)
Tracking: Dual Critical/Divided attention, psychomotor	See Damros, Bittner, Kennedy and Harbeson, 1981	1 minute/Ipsative*	Experimental test using Navy enlisted men	Not reported	Total stabilization time=100 minutes; reliability efficiency (3 minutes)=0.50	Not reported	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Visual Probability Monitoring Task/ Attention	Computer	3 minutes per trial/ Ipsative*	Experimental test using military personnel	Not directly demonstrated	Authors cite Shingledecker et al., 1983, and infer validity for tapping resources associated with visual information processing	Authors cite research by Shingledecker et al., 1983, which indicates that there is very little practice effect; the subject's performance is relatively stable at the start of testing	Perez, Masline, Ramsey and Urban (1987)

\*Determined by reviewers.

## Appendix G

### Cognitive Task: Time Estimation

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Interval Production Task/Time estimation, psychomotor, attention	Computer	Practice time: 15 minutes. Task time: 3 minutes/Ipsative*	Experimental test using military personnel	The literature indicates that performance on the tapping task in a dual-task condition is diagnostic of the motor output load imposed by the primary task.	This task has sensitivity in dual-task experiments to primary tasks that impose demands on motor output performance.	Not reported	Perez, Masline, Ramsey and Urban (1987)
Time Wall/Ability to estimate time	Computer	10 trials or 5 minutes, whichever occurs first/Ipsative*	Experimental test using military personnel	Authors cite sensitivity to atropine based on study by Seppala and Visakorpi, 1983	Not reported	Not directly demonstrated	Perez, Masline, Ramsey and Urban (1987)

\*Determined by reviewers.

## Cognitive Task: Perceptual Speed\*

## Appendix G

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Code Substitution/ Perceptual speed, associative memory	Paper and pencil	2 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=16 minutes; reliability efficiency (3 minutes)=0.84	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Code Substitution/ Perceptual speed, associative memory	Computer	Practice time: 30 seconds. Trial time: 60 seconds. Total task time, less practice: 60 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to impairment from alcohol	Not reported	Kennedy, Dunlap, Turnage, Rugotzke, Wilkes and Smith (1991)
Code Substitution/ Perceptual speed, associative memory	Computer	Practice time: 30 seconds. Trial time: 90 seconds. Total task time, less practice: 90 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to alcohol. F(6,102) ratios for impact of alcohol on performance=20.74; p<.000	Not reported	Kennedy, Turnage and Dunlap (1992)
Code Substitution/ Perceptual speed, associative memory	Computer	Not reported	Experimental (see Kennedy, Turnage and Dunlap, 1993)	Not reported	Sensitive to alcohol F=26.2 p<.01	Not reported	Kennedy, Turnage and Dunlap (1993)
Code Substitution/ Perceptual speed, short-term memory	Computer	Practice time: 10 seconds. Trial time: 60 seconds. Total task time, less practice: 900 seconds/Ipsative*	Experimental test using students from an introductory psychology class	Intermode correlations between computer-administered subtests and WAIS-R, Verbal, Performance, and Full-Scale IQ (N=25) showed correlations with performance (0.44, p<0.01)	Not reported	Stability r=0.78; 3-minute reliability efficiency=0.86	Kennedy, Wilkes, Dunlap and Kuntz (1987)

\*Some measures of short-term memory are actually measures of perceptual speed (Guilford and Hoepfner, 1971). Although we distinguish between the two constructs, it is important to note that they may be measures of the same underlying cognitive ability.

\*Determined by reviewers.

## Appendix G

### Cognitive Task: Perceptual Speed\* (Continued)

Measure/Construct	Method of Administration/Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Code Substitution/ Perceptual speed, short-term memory	Paper and pencil	Practice time: 15 seconds. Trial time: 60 seconds. Total task time, less practice: 60 seconds/ Ipsiative*	Intermode correlations between paper-and-pencil and computer-administered subtests (N=25) showed correlations with Pattern Comparison (0.58, $p<0.05$ ), Manikin (0.43, $p<0.01$ ), Code Substitution (0.77, $p<0.05$ )	Not reported	Stability $r=0.81$ ; 3-minute reliability efficiency=0.86	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Code Substitution/ Perceptual speed, associative memory	Computer	Not reported	Authors infer that this task measures perceptual speed and associative learning ability from research by Cohen, 1957a, 1957b	Sensitive to organic factors, such as brain damage. No tests have been conducted to determine sensitivity to environmental stressors	Authors infer test-retest reliability from research by Demer, Abome and Castore (1950)	Perez, Masline, Ramsey and Urban (1987)
Code Substitution/ Perceptual speed, associative memory	Computer	Not reported	Experimental test using military personnel	Sensitive to alcohol, chemotherapy, antihistamines, and Halon	Six metrology studies, range for stability=0.55 to 0.85	Turnage, Kennedy, Smith, Baltzley and Lane (1992)
Letter Classification: Category/ Perceptual speed	See Posner and Mitchell, 1967	11 minutes/Ipsiative*	Experimental test using Navy enlisted men	Not reported	Total stabilization time=121 minutes; reliability efficiency (3 minutes)=0.69	Bittner, Carter, Kennedy, Harbeson and Krause (1986)

\*Some measures of short-term memory are actually measures of perceptual speed (Guilford and Hoepfner, 1971). Although we distinguish between the two constructs, it is important to note that they may be measures of the same underlying cognitive ability.

\*Determined by reviewers.

Cognitive Task: Perceptual Speed<sup>#</sup> (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Letter Classification: Name/Perceptual speed	See Posner and Mitchell, 1967	12 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=84 minutes; reliability efficiency (3 minutes)=0.55	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Letter Classification: Physical Match/Perceptual speed	See Posner and Mitchell, 1967	12 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=108 minutes; reliability efficiency (3 minutes)=0.52	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Letter Search: Time per Correct Item/Perceptual speed	See Neisser, Novick and Lazar, 1963	3 minutes/Ipsative *	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=27 minutes; reliability efficiency (3 minutes)=0.87	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Linguistic Processing Task/Perceptual speed	Computer	3 minutes per trial and 15 to 30 minutes of practice/Ipsative*	Experimental test using military personnel	This task assesses ability to encode information at different levels of processing	Not directly demonstrated	Not reported	Perez, Masline, Ramsey and Urban (1987)
Number Comparison/Perceptual speed	Paper and pencil	3 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=27 minutes; reliability efficiency (3 minutes)=0.84	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Object Naming/Perceptual speed	Apparatus	50-60 minutes/ Ipsative*	Experimental research using social drinkers	Not directly demonstrated	Both alcohol and word frequency produced statistically significant effects on response latency	Not reported	Moskowitz and Roth (1971)
Pattern Comparison/Perceptual speed	Paper and pencil (see Klein and Armitage, 1979)	2 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time= 18 minutes; reliability efficiency (3 minutes)=0.93	Bittner, Carter, Kennedy, Harbeson and Krause (1986)

<sup>#</sup>Some measures of short-term memory are actually measures of perceptual speed (Guilford and Hoepfner, 1971). Although we distinguish between the two constructs, it is important to note that they may be measures of the same underlying cognitive ability.

\*Determined by reviewers.

## Appendix G

### Cognitive Task: Perceptual Speed<sup>#</sup> (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Pattern Comparison/Perceptual speed	Computer	Practice time: 30 seconds. Trial time: 60 seconds. Total task time, less practice: 60 seconds/ Ipsative*	Experimental test using male university students	Not reported	Sensitive to impairment from alcohol	Not reported	Kennedy, Dunlap, Turnage, Rugotzke, Wilkes and Smith (1991)
Pattern Comparison/Perceptual speed	Computer	Practice time: 30 seconds. Trial time: 90 seconds. Total task time, less practice: 90 seconds/ Ipsative*	Experimental test using male university students	Not reported	Sensitive to alcohol. $F(6,102)$ ratios for impact of alcohol on performance=13.41; probability=0.000	Not reported	Kennedy, Turnage and Dunlap (1992)
Pattern Comparison/Perceptual speed	Paper and pencil	Practice time: 15 seconds. Trial time: 75 seconds. Total task time, less practice: 75 seconds/ Ipsative*	Experimental test using students from an introductory psychology class	Intermode correlations between paper-and-pencil and computer-administered subtests (N=25) showed correlations with short-term memory (0.49, $p<0.05$ ), pattern comparison (0.84, $p<0.05$ ), code substitution (0.55, $p<0.05$ ), and grammatical reasoning (0.48, $p<0.01$ )	Not reported	Stability $\alpha=0.84$ ; 3-minute reliability efficiency=0.90	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Pattern Comparison (Simultaneous)/Perceptual speed	Computer	Not reported	Experimental test using military personnel	Test scores correlate with spatial tasks such as Digit Symbol Substitution (0.71), Block Design (0.59), Picture Arrangement (0.29), and Object Assembly (0.27). See Kennedy, Dunlap, Ones, Lane and Wilkes (1985).	Not directly demonstrated	Authors infer repeated measures reliability based on task reliability of 0.93 reported in Klein and Armitage (1979)	Perez, Masline, Ramsey and Urban (1987)

\*Some measures of short-term memory are actually measures of perceptual speed (Guilford and Hoepfner, 1971). Although we distinguish between the two constructs, it is important to note that they may be measures of the same underlying cognitive ability.

\*Determined by reviewers.

Cognitive Task: Perceptual Speed<sup>#</sup> (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Type of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Pattern Comparison (Simultaneous)/ Perceptual speed	Computer	Not reported	Experimental (see Turnage, Kennedy, Smith, Baltzley and Lane, 1992)	Correlates with American College Testing (0.65, $p<0.01$ ), Wonderlic (0.68, $p<0.01$ ), Armed Services Vocational Aptitude Battery (0.81, $p<0.01$ ), WAIS Verbal (0.63, $p<0.01$ ), WAIS Performance (0.56, $p<0.01$ ), and WAIS Composite (0.64, $p<0.01$ ). Average $r=0.66$	Sensitive for alcohol, altitude, motion sickness drugs, chemotherapy, antihistamines, and Halon	Six metrology studies. Range for stability=0.60 to 0.94	Turnage, Kennedy, Smith, Baltzley and Lane (1992)
Pattern Comparison (Successive)/ Perceptual speed, short-term spatial memory	Computer	25 seconds per trial/ Ipsiative*	Experimental test using military personnel	Test scores correlate with spatial tasks such as Digit Symbol Substitution (0.71), Block Design (0.59), Picture Arrangement (0.29), and Object Assembly (0.27). See Kennedy, Dunlap, Ones, Lane and Wilkes (1985).	Not directly demonstrated	Authors infer repeated measures reliability based on task reliability of 0.93 reported in Klein and Armitage (1979)	Perez, Masine, Ramsey and Urban (1987)
Pattern Recognition/ Perceptual Speed	Paper and pencil	Practice time: 15 seconds. Trial time: 75 seconds. Total task time, less practice: 150 seconds/ Ipsiative*	Experimental test using students from an introductory psychology class	Intermode correlations between paper-and-pencil and computer-administered subtests ( $N=25$ ) showed correlations with nonpreferred-hand tap (0.43, $p<0.01$ ), two-handed tap (0.54, $p<0.05$ ), pattern comparison (0.71, $p<0.05$ ), and code substitution (0.61, $p<0.05$ )	Not reported	Stability $r=0.90$ ; 3-minute reliability efficiency=0.93	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Pattern Recognition: Time Per Correct Item/Perceptual speed	See Fitts, Weinstein, Rappaport, Anderson and Leonard, 1956	2 minutes/Ipsiative*	Experimental test using Navy enlisted men	Not reported	Total stabilization time=20 minutes; reliability efficiency (3 minutes)=0.76	Bittner, Carter, Kennedy, Harbeson and Krause (1986)	

<sup>#</sup>Some measures of short-term memory are actually measures of perceptual speed (Guilford and Hoepfner, 1971). Although we distinguish between the two constructs, it is important to note that they may be measures of the same underlying cognitive ability.

\*Determined by reviewers.

## Appendix G

### Cognitive Task: Perceptual Speed<sup>#</sup> (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Perceptual Speed/Perceptual speed	Paper and pencil. See Ekstrom, French, Harmon and Dermen (1976)	2.5 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=23 minutes; reliability efficiency (3 minutes)=0.86	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Random Field Number Search: Time Per Correct Item/Perceptual speed	Not reported	5 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=35 minutes; reliability efficiency (3 minutes)=0.55	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Reading Speed/Perceptual Speed	See Baron and McKillop, 1975	8 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=16 minutes; reliability efficiency (3 minutes)=0.66	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Search For Types In Prose: Median Detection Time/Perceptual speed	Not reported	6 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=54 minutes; reliability efficiency (3 minutes)=0.65	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Visual Scanning/Perceptual speed	Computer	40 trials or 5 minutes, whichever occurs first/Ipsative*	Experimental test using military personnel	Authors report that this task taps a person's speed for making rapid visual discriminations	Not directly demonstrated	Test-retest reliabilities obtained on this task risk being confounded by practice effects	Perez, Masline, Ramsey and Urban (1987)

<sup>#</sup>Some measures of short-term memory are actually measures of perceptual speed (Guilford and Hoepfner, 1971). Although we distinguish between the two constructs, it is important to note that they may be measures of the same underlying cognitive ability.

\*Determined by reviewers.

## Cognitive Task: Memory

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
American Medical Association Mini-Mental Test/Short-term memory, attention	Paper and pencil	Under 5 minutes/ Normative*	Experimental test using pilots from flying schools and volunteers from an outpatient mental health clinic	Not reported	AMA showed sensitivity at 0.69 in detecting medically classified impairment (e.g., head trauma)	Not reported	Stokes, Banich and Ellledge (1991)
Associative Memory: Number Correct/ Associative memory	Paper and pencil	2.5 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported		Total stabilization time=20 minutes; reliability efficiency (3 minutes)=0.65	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Continuous Recognition Task/Short-term memory	Computer	3 minutes per trial/ Ipsative*	Experimental test using military personnel	Results from factor analyses indicate that scores on this test load on a memory factor. See Hunter, 1975.	Not directly demonstrated	Not directly demonstrated	Perez, Masline, Ramsey and Urban (1987)
Item Order Test/ Short-term memory	Computer	10 seconds per trial/ Ipsative*	Experimental test using military personnel	Not directly demonstrated	Not directly demonstrated	Not directly demonstrated	Perez, Masline, Ramsey and Urban (1987)
Matching to Sample/Short-term memory, spatial ability	Computer	30 2.5-minute trials for a total test time of 75 minutes/Ipsative*	Experimental test using military personnel	Not directly demonstrated	Not directly demonstrated	Not directly demonstrated	Perez, Masline, Ramsey and Urban (1987)
Memory Search Tasks (Variation of Sternberg Memory Task)/ Short-term working memory	Computer	Not reported	Experimental test using military personnel	Diagnostic of retrieval and comparison of items in short-term memory processes. Also diagnostic of time required for stimulus encoding and response formulation	Authors cite studies that report sensitivity for drugs, dual-tasks, and event-related potentials (ERPs)	Authors infer test-retest reliability based on Bittner, Carter, Kennedy, Harbeson and Krause, 1986	Perez, Masline, Ramsey and Urban (1987)
Number Recall/ Memory	Computer	10 trials, total task time 3 minutes	Experimental using occasional drug and alcohol users.	Not reported	$F=10.85$ $p<0.005$ for diazepam. Sensitive to marijuana, but no F value given.	Not reported	Kelly, Folen, Emurian and Fischman (1993)
Rate of Information Processing Task/ Memory	Computer	Not reported	Experimental test using moderate alcohol and marijuana users	Not reported	Sensitive to alcohol and marijuana	Not reported	Burns and Moskowitz (1981)

\*Determined by reviewers.

## Appendix G

### Cognitive Task: Memory (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Short-Term Memory/Short-term memory	Computer	Practice time: 30 seconds. Trial time: 60 seconds. Total task time, less practice: 60 seconds/ Ipsative*	Experimental test using male university students	Not reported	Sensitive to impairment from alcohol	Not reported	Kennedy, Dunlap, Turnage, Rugoizke, Wilkes and Smith (1991)
Short-Term Memory/Short-term memory	Computer	Practice time: 30 seconds. Trial time: 90 seconds. Total task time, less practice: 90 seconds/ Ipsative*	Experimental test using male university students	Not reported	Sensitive to alcohol. F(6,102) ratios for impact of alcohol on performance=8.60; p<0.00	Not reported	Kennedy, Turnage and Dunlap (1992)
Short-Term Memory/Short-term memory	Computer	Practice time: 10 seconds. Trial time: 60 seconds. Total task time, less practice: 900 seconds/ Ipsative*	Experimental test using students from an introductory psychological class	Not reported	Intermode correlations between computer-administered subtests and WAIS-R, Verbal, Performance, and Full-Scale IQ (N=25) showed no correlations	Stability r=0.75; 3-minute reliability efficiency=0.84	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Short-Term Memory/Short-term scanning rate	Computer	Not reported	Experimental. See Turnage, Kennedy, Smith, Baltzley and Lane, 1992	Not reported		Not reported	Turnage, Kennedy, Smith, Baltzley and Lane (1992)
Sternberg Item Recognition: Positive Set 1/ Short-term memory scan	See Sternberg, 1966; Sternberg, 1975	3 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported		Not reported	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Sternberg Item Recognition: Positive Set 4/ Short-term memory scan	See Sternberg, 1966; Sternberg, 1975	3 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported		Not reported	Total stabilization time=18 minutes; reliability efficiency (3 minutes)=0.70
Visual Short-Term Memory Task/Short-term memory	Computer	Not reported	Experimental test using moderate alcohol and marijuana users	Not reported		Not reported	Total stabilization time=15 minutes; reliability efficiency (3 minutes)=0.80
							Burns and Moskowitz (1981)

\*Determined by reviewers.

## Cognitive Task: Perception

## Appendix G

Measure/ Construct	Method of Administration	Administration Time/ Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	References
<b>Flexibility of Closure/ Perception</b>	Paper and pencil	3 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=9 minutes; reliability efficiency (3 minutes)=0.88	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
<b>Speed of Closure/ Perception</b>	Paper and pencil	2.5 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=28 minutes; reliability efficiency (3 minutes)=0.80	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
<b>Stroop: Black and White Words/Perception</b>	See Jensen and Rohwer, 1966	0.5 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=1.5 minutes; reliability efficiency (3 minutes)=0.96	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
<b>Stroop: Color Blocks/Perception, response interference</b>	Not reported	0.5 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=3.5 minutes; reliability efficiency (3 minutes)=0.98	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
<b>Stroop: Color Words/Perception, response interference</b>	Not reported	0.5 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=1.5 minutes; reliability efficiency (3 minutes)=0.97	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
<b>Stroop Test/ Perception, susceptibility to response interference</b>	Computer	Not reported	Experimental test using military personnel	The Stroop interference effect has been correlated with a wide variety of perceptual, memory, and intelligence tests	Sensitive to stimulants and nicotine (improved performance); see Jensen and Rohwer, 1966; Wenes and Warburton, 1978; depressants, LSD, scopolamine, and atropine (decreased performance); see Jensen and Rohwer, 1966; Calloway and Band, 1958; Ostfeld and Ariquette, 1962. Also sensitive to age and psychiatric disturbance; see Jensen and Rohwer, 1966.	Not directly demonstrated	Perez, Masline, Ramsey and Urban (1987)

\*Determined by reviewers.

## Appendix G

### Cognitive Task: Spatial Ability

Measure/ Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Manikin/Spatial ability	Computer	Practice time: 30 seconds. Trial time: 60 seconds. Total task time, less practice: 60 seconds/ Ipsative*	Experimental test using male university students	Not reported	Sensitive to impairment from alcohol	Not reported	Kennedy, Dunlap, Turnage, Rugotzke, Wilkes and Smith (1991)
Manikin/Spatial ability	Computer	Practice time: 30 seconds. Trial time: 90 seconds. Total task time, less practice: 90 seconds/ Ipsative*	Experimental test using male university students	Not reported	Sensitive to alcohol. $F(6,102)$ ratios for impact of alcohol on performance=16.87; $p<.00$	Not reported	Kennedy, Turnage and Dunlap (1992)
Manikin/Spatial ability	Computer	Not reported	Experimental (see Kennedy, Turnage and Dunlap, 1993)	Not reported	Sensitive to alcohol $F=16.9$ $p<$	Not reported	Kennedy, Turnage and Dunlap (1993)
Manikin/Spatial ability	Computer	Practice time: 10 seconds. Trial time: 60 seconds. Total task time, less practice: 900 seconds/ Ipsative*	Experimental test using students from an introductory psychology class	Intermode correlations between computer- administered subtests and WAIS-R, Verbal, Performance, and Full-Scale IQ (N=25) showed correlations with performance (0.50, $p<0.01$ )	Not reported	Stability $r=0.88$ ; 3- minute reliability efficiency=0.93	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Manikin/Spatial ability	Computer	Between 64 and 96 trials are administered for each test. Total test time is approximately 4 minutes/Ipsative*	Experimental test using military personnel	Appears to measure spatial transformation (see Carter and Wolstad, 1985)	Not directly demonstrated	No reliability data reported. Authors infer reliability from the manikin test developed by Carter and Wolstad, 1985	Perez, Masline, Ramsey and Urban (1987)

\*Determined by reviewers.

## Cognitive Task: Spatial Ability (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
<b>Manikin/Spatial ability</b>	Computer	Not reported	Experimental. See Turnage, Kennedy, Smith, Baltzley and Lane, 1992	Correlate with American College Testing (0.50, p<0.01), Wonderlic (0.66, p<0.01), Armed Services Vocational Aptitude Battery (0.62, p<0.01), WAIS Verbal (0.42, p<0.05), WAIS Performance (0.39, p<0.05), and WAIS Composite (0.44, p<0.05). Average r=0.51.	Sensitive for alcohol, motion sickness drugs, chemotherapy, antihistamines, and Halon	Five metrology studies. Range for stability=0.70 to 0.97	Turnage, Kennedy, Smith, Baltzley and Lane (1992)
<b>Manikin Test: Cognitive Latency/Spatial ability</b>	See Egan, 1978	7 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=14 minutes; reliability efficiency (3 minutes)=0.79	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
<b>Matrix Rotation Task/Spatial ability</b>	Computer	20-minute practice time; 20 90-second trials for total task time, less practice of 30 minutes/Ipsative*	Experimental test using military personnel	Not directly demonstrated	Not directly demonstrated	No reliability data reported. Authors infer reliability from the Fitts Histogram, as reported in Klein and Armitage, 1979	Perez, Masline, Ramsey and Urban (1987)
<b>Multiple Performance Test Battery/Spatial ability</b>	Computer	150 trials total task time 10 minutes	Experimental using occasional drug and alcohol users	Not reported	For diazepam F=4.29 p<0.05. Sensitive to alcohol and marijuana, but no F value given.	Not reported	Kelly, Foltin, Emurian and Fischman (1993)
<b>Spatial Processing Task/ Spatial ability</b>	Computer	3 minutes per trial/ Ipsative*	Experimental test using military personnel	Authors infer validity based on the Fitts Histogram (see Kennedy et al., 1985)	Not directly demonstrated	Authors infer test-retest reliability based on the Fitts Histogram (see Kennedy et al., 1985)	Perez, Masline, Ramsey and Urban (1987)

\*Determined by reviewers.

## Appendix G

### Cognitive Task: Reasoning

Measure/ Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Grammatical Reasoning/ Reasoning	Paper and pencil	1.5 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=18 minutes; reliability efficiency (3 minutes)=0.93	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Grammatical Reasoning/ Reasoning	Computer	Practice time: 30 seconds. Trial time: 60 seconds. Total task time, less practice: 60 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to impairment from alcohol	Not reported	Kennedy, Dunlap, Turnage, Rugotzke, Wilkes and Smith (1991)
Grammatical Reasoning/ Reasoning	Computer	Practice time: 30 seconds. Trial time: 90 seconds. Total task time, less practice: 90 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to alcohol; F(6,102) ratios for impact of alcohol on performance=8.04; p<0.00	Not reported	Kennedy, Turnage and Dunlap (1992)
Grammatical Reasoning/ Reasoning	Computer	Not reported	Experimental (see Kennedy, Turnage and Dunlap, 1993)	Not reported	Sensitive to alcohol F=6.4	Not reported	Kennedy, Turnage and Dunlap (1993)
Grammatical Reasoning/ Reasoning	Computer	Practice time: 15 seconds. Trial time: 90 seconds. Total task time, less practice: 1350 seconds/Ipsative*	Experimental test involving students from an introductory psychology class	Intermode correlations between computer-administered subtests and WAIS-R, Verbal, Performance, and Full-Scale IQ (N=25) showed correlations with verbal (0.41, p<0.01), performance (0.63, p<0.05), and full- scale (0.66, p<0.05)	Not reported	Stability r=0.85; 3- minute reliability efficiency=0.89	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Grammatical Reasoning/ Reasoning	Paper and pencil	Practice time: 15 seconds. Trial time: 90 seconds. Total task time, less practice: 90 seconds/Ipsative*	Experimental test involving students from an introductory psychology class	Intermode correlations between paper- and-pencil and computer-administered subtests (N=25) showed correlations with preferred-hand tap (0.56, p<0.05), nonpreferred-hand tap (0.58, p<0.05), pattern comparison (0.47, p<0.01), code substitution (0.50, p<0.01), and grammatical reasoning (0.85, p<0.05)	Not reported	Stability r=0.85; 3- minute reliability efficiency=0.89	Kennedy, Wilkes, Dunlap and Kuntz (1987)

\*Determined by reviewers.

## Cognitive Task: Reasoning (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Constrained Validity	Sensitivity	Reliability	Reference
Grammatical Reasoning (Symbolic)/Reasoning	Computer	Two 36-minute training sessions of four 3-minute trials at each level of task difficulty are suggested/Ipsative*	Experimental test using military personnel	Authors infer validity for tapping resources of logical reasoning ability and working memory processing capacity based on the Baddeley and Hatch, 1974, study using the traditional version of this task	Not directly demonstrated	Authors infer test-retest reliability of this task based on findings for paper-and-pencil versions of this task	Perez, Masline, Ramsey and Urban (1987)
Grammatical Reasoning (Traditional)/Reasoning	Computer	15 seconds + 500 msec/trial for a total of 32 stimuli/ Ipsative*	Experimental tests using military personnel	Authors report the task appears to measure "higher mental process" associated with logical reasoning and working memory processing capability	This test has been shown to be sensitive to the effects of sleep loss, environmental stress (e.g., performance under water), road pollutants, and diurnal variations. This task has also been shown to affect performance on a driving task in dual task experiments (Brown et al., 1968)	This test has been shown to be sensitive to the effects of sleep loss, environmental stress (e.g., performance under water), road pollutants, and diurnal variations. This task has also been shown to affect performance on a driving task in dual task experiments (Brown et al., 1968)	Perez, Masline, Ramsey and Urban (1987)
Grammatical Reasoning/Reasoning	Computer	Not reported	Experimental. See Turnage, Kennedy, Smith, Baltzley and Lane, 1992	Correlations with American College Testing (0.52, p<0.01), Wonderlic (0.52, p<0.01), Armed Services Vocational Aptitude Battery (0.53, p<0.01). Average r=0.40	Sensitive to effects of alcohol, motion sickness drugs, chemotherapy, antihistamines, and Halon	Seven metiology studies. Range for stability=0.59 to 0.96	Turnage, Kennedy, Smith, Baltzley and Lane (1992)
Wisconsin Card Sorting Test (WCST)/Reasoning	Computer	Not reported	Experimental test using social drinkers	Reduced inhibition of perseverative tendencies may reflect disruption of more general functions identified with frontal cortex, such as the inhibitory regulation of a wide range of behaviors related to planning, the regulation of behavior, and anticipating future consequences of action (Lhermitte, 1986; Luria, 1973; Mesulam, 1986; Stuss and Benson, 1984)	Significant for alcohol	Not directly demonstrated	Lyvers and Maltzman (1991)

\*Determined by reviewers.

## Appendix G

### Cognitive Task: Fluency

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Graphemic and Phonemic and Analysis: Sense/ Nonsense/Fluency	See Baron and McKillip, 1975	8 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=72 minutes; reliability efficiency (3 minutes)=0.73	Bittnner, Carter, Kennedy, Harbeson and Krause (1986)
Graphemic and Phonemic and Analysis: Sense/ Nonsense/Fluency	See Baron and McKillip, 1975	8 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=40 minutes; reliability efficiency (3 minutes)=0.66	Bittnner, Carter, Kennedy, Harbeson and Krause (1986)
Graphemic and Phonemic and Analysis: Sense/ Nonsense/Fluency	See Baron and McKillip, 1975	8 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=16 minutes; reliability efficiency=0.66 minutes	Bittnner, Carter, Kennedy, Harbeson and Krause (1986)
Word Fluency/ Word fluency	Paper and pencil	3 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time= 1 minute; reliability efficiency (3 minute)=0.79	Bittnner, Carter, Kennedy, Harbeson and Krause (1986)

\*Determined by reviewers.

## Cognitive Task: Number Facility

Measured/Constrained	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Arithmetic Computation/ Number facility	Paper and pencil	10 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=90 minutes; reliability efficiency (3 minutes)=0.83	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Arithmetic: Number Facility/ Number facility	Paper and pencil	10 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=27 minutes; reliability efficiency (3 minutes)=0.83	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Arithmetic: Vertical Addition/Number facility	Paper and pencil	4 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=48 minutes; reliability efficiency (3 minutes)=0.90	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Math Processing/ Number facility	Computer	Practice time: 30 seconds. Trial time: 60 seconds. Total task time, less practice: 60 seconds/ Ipsative*	Experimental test using male university students	Not reported	Sensitive to impairment from alcohol	Not reported	Kennedy, Dunlap, Turnage, Rugotzke, Wilkes and Smith (1991)
Math Processing/ Number facility	Computer	Practice time: 30 seconds. Trial time: 90 seconds. Total task time, less practice: 90 seconds/ Ipsative*	Experimental test using male university students	Not reported	Sensitive to alcohol. F(6,102) ratios for impact of alcohol on performance=13.60; p<.00	Not reported	Kennedy, Turnage and Dunlap (1992)
Mathematical Processing Task/ Number facility	Computer	30 minute practice; experiment ranges from 7 three-minute trials to 30 three-minute trials/Ipsative*	Experimental test using military personnel	Not directly demonstrated	Not directly demonstrated	Not reported	Perez, Masline, Ramsey and Urban (1987)
Two-Column Addition Task/ Number facility	Computer	10 practice trials and 15 seconds per trial Ipsative*	Experimental test using military personnel	Authors infer sensitivity to methyl chloride, jet fuel, atropine, Ditran, scopolamine, and sleep deprivation	Not directly demonstrated	Not directly demonstrated	Perez, Masline, Ramsey and Urban (1987)

\*Determined by reviewers.

## Appendix G

### Cognitive Task: Reaction Time

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Choice Reaction Time: 1 Choice/ Simple reaction time	Paper and pencil. See Donders, 1968; Donders, 1969	5 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=35 minutes; reliability efficiency (3 minutes)=0.58	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Choice Reaction Time: 2 Choice/ Choice reaction time	Paper and pencil. See Donders, 1968; Donders, 1969	5 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=35 minutes; reliability efficiency (3 minutes)=0.51	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Choice Reaction Time: 4 Choice/ Choice reaction time	Paper and pencil. See Donders, 1968; Donders, 1969	5 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=50 minutes; reliability efficiency (3 minutes)=0.80	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Four-Choice Serial Reaction Time/Choice reaction time	Computer	6 minutes per trial/ Ipsative*	Experimental test using military personnel	Authors cite Wilkinson and Houghton, 1975, to support this task's internal validity	Authors cite various studies finding sensitivity to alcohol, toluene, general anesthesia, sustained physical effort, sleep deprivation effects, and dual task paradigms	Authors cite Krause and Bitner, 1982, to support this task's internal reliability	Perez, Masline, Ramsey and Urban (1987)
Inter-Stimulus Intervals in a Psychological Refractory Period Paradigm/ Reaction time	Apparatus	18 sets of 12 10-second trials for a total test time of 1 hour/Ipsative*	Experimental test using 10 male college students and professionals	Authors report differences between randomly presented stimuli and stimuli presented at a constant interval	Not directly demonstrated	Not reported	Burns and Moskowitz (1971)
Reaction Time/ Reaction time	Computer	Practice time: 30 seconds. Trial time: 60 seconds. Total task time, less practice: 60 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to impairment from alcohol	Not reported	Kennedy, Dunlap, Turnage, Ruzicka, Wilkes and Smith (1991)

\*Determined by reviewers.

## Cognitive Task: Reaction Time (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Reaction Time/Reaction time	Computer	Practice time: 30 seconds. Trial time: 90 seconds. Total task time, less practice: 90 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to alcohol. F(6,102) ratios for impact of alcohol on performance=4.38, p<0.001	Not reported	Kennedy, Turnage and Dunlap (1992)
Reaction Time/Reaction time	See Kennedy, Dunlap, Jones, Lane and Wilkes, 1985	Practice time: 15 seconds. Trial time: 60 seconds. Total task time, less practice: 900 seconds/Ipsative*	Experimental test using students from an introductory psychology class	Intermode correlations between computer-administered subtests and WAIS-R, Verbal, Performance, and Full-Scale IQ (N=25) showed no significant correlations	Not reported	Stability r=0.71; 3-minute reliability efficiency=0.80	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Reaction Time (1 Choice, 2 Choice, 4 Choice)/Reaction time	Computer	Not reported	Experimental test. See Turnage, Kennedy, Smith, Baltzley and Lane, 1992	Correlations for 1-choice reaction times with American College Testing (-0.50, p<0.01), Wonderlic (-0.65, p<0.01), Armed Services Vocational Aptitude Battery (-0.66, p<0.01), WAIS Verbal (-0.40, p<0.05), WAIS Performance (-0.40, p<0.05), and WAIS Composite (-0.42, p<0.05). Average r=0.57. No predictive validity cross-correlations were reported for 2- or 4-choice reaction times.	4-choice reaction time test showed sensitivity for alcohol, chemotherapy, antihistamines, and Halon. Sensitivity for 1- or 2-choice reaction time was not reported	Six metology studies. Range for stability=0.56 to 0.86	Turnage, Kennedy, Smith, Baltzley and Lane (1992)
Stimulus-Response/Response time	Not reported	Total test time, including instructions and breaks=2.5 minutes/Ipsative*	Experimental test using social drinkers	Not reported	Significantly affects response latency for alcohol. The F ratio squared for mean response latencies is 3.27 (df=11/102, p<0.01) for the control treatment and 2.99 (df=11/102, p<0.01) for the alcohol treatment	Not reported	Mostkowitz and Burns (1973)

\*Determined by reviewers.

## Appendix G

### Psychomotor Task

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Aiming/Fine eye-hand coordination	See Fleishman and Ellison, 1962	2 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=30 minutes; reliability efficiency (3 minutes)=0.87	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Aiming/Fine eye-hand coordination	Paper and pencil	Practice time: 15 seconds. Trial time: 90 seconds. Total task time, less practice: 180 seconds/Ipsative*	Experimental test using students from an introductory psychology class	Internode correlations between paper-and-pencil and computer-administered subtests (N=25) showed correlations with preferred-hand tap (0.62, p<0.05), nonpreferred-hand tap (0.64, p<0.05), two-hand tap (0.64, p<0.05) and pattern comparison (0.44, p<0.01)	Not reported	Stability r=0.94; 3-minute reliability efficiency=0.97	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Atari Air Combat Maneuvering/ Pursuit tracking	Computer-administered video game	2.25 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=135 minutes; reliability efficiency (3 minutes)=0.63	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Atari Aircraft/ Not reported	Computer-administered video game	2.25 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=126 minutes; reliability efficiency (3 minutes)=0.67	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Compensatory Tracking Task/ Eye-hand coordination	Computer	6 minutes per trial/ Ipsative*	Experimental test using moderate alcohol and marijuana users	Not reported	Test found to be statistically significant for alcohol and marijuana	Not reported	Burns and Moskowitz (1981)
Critical Tracking Task/ Eye-hand coordination	Not reported	Not reported	Experimental test used to validate a statistical method for determining pass/fail levels for convicted drunk drivers	Not reported	Between subject variability is about 9% (range is 6% to 10%)	Negligible trial-to-trial differences. Within subject variability is about 9%	Allen, Stein and Jex (1981)

\*Determined by reviewers.

## Psychomotor Task (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference	
Critical Tracking Task/ Eye-hand coordination	See Jex, McDonnell and Phatak, 1966	1 minute/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=100 minutes; reliability efficiency (3-minutes)=0.60	Bittner, Carter, Kennedy, Harbeson and Krause (1986)	
Critical Tracking Test/ Eye-hand coordination	Computer	Not reported	Experimental test using moderate alcohol and marijuana users	Not reported	The combined alcohol-marijuana treatment was readily identified and produced the highest level of experience intoxication	Not reported	Burns and Moskowitz (1981)	
Critical Tracking Task Based Drunk Driver Warning System (CTT/ DDWS)/Eye-hand coordination	Computer	Not reported	Experimental test involving convicted drunk driving offenders	DDWS performance of subjects was compared with their driving in an interactive car simulator. Results showed 35 to 40 percent of the subjects failed the test at a BAC of 1.0 promil w/v depending on group differences and learning motivation. At 1.5 promil w/v the rejection rate increased to 75 to 80 percent. Rejected subjects at 1.0 promil w/v and above also exhibited impaired simulator driving performance when compared to subjects passing the DDWS test.	Authors cite previous laboratory research, which has shown performance on the CTT to be sensitive to alcohol.	Day failure rates were about what was expected (approximately 2.5%) based on the procedure used to set individual CTT pass scores. Nighttime failure rates were 3 to 7 times greater than this level, and were consistent with the laboratory discriminability results where BAC was controlled for and measured.	Not reported	Stein, Allen and Jex (1987)
Critical Tracking Task (Factor 1000)/ Eye-hand coordination	Computer	About 1 minute/ Ipsative*	Experimental test using truck drivers	Not reported			Not reported	Allen, Silverman and Ikonen (1992), (Performance Factors, Inc.)

\*Determined by reviewers.

## Appendix G

### Psychomotor Task (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Finger Count/Balance, coordination	Observation	Not reported	Experimental test using 62 light drinkers, 86 moderate drinkers, and 90 heavy drinkers	Not reported	Based on scoring by 10 officers, the mean for finger count below 0.10% BAC was 2.94, and the mean for finger count equal to or above 0.10% BAC was 4.95. Based on scoring by 2 observers, the mean for finger count below 0.10% BAC was 2.87, and the mean for finger count equal to or above 0.10% BAC was 4.55	Interrater reliability between 10 police officers and 2 observers were as follows for finger count: test score=0.87 for participants below 0.10% BAC and 0.79 participants equal to or above 0.10% BAC	Burns and Moskowitz (1977)
Finger-to-Nose/Muscle coordination	Observation	Not reported	Experimental test using 62 light, 86 moderate, and 90 heavy drinkers	10 officers' scoring of finger-to-nose test correlated with BAC as 0.421	Based on scoring by 10 officers, the mean for finger-to-nose below 0.10% BAC was 2.54, and the mean for finger-to-nose equal to or above 0.10% BAC was 4.47. Based on scoring by 2 observers, the mean for finger-to-nose below 0.10% BAC was 2.19, and the mean for finger-to-nose equal to or above 0.10% BAC was 4.52	Interrater reliability between 10 police officers and 2 observers were as follows for finger-to-nose: test score=0.60 for participants below 0.10% and 0.72 for participants equal to or above 0.10% BAC	Burns and Moskowitz (1977)
Illinois Screening Tests 1 and 2/ Psychomotor	Paper and pencil	15 to 25 minutes/ Normative*	Experimental test using pilots from flying schools and volunteers from an outpatient mental health clinic	Not reported	IST-1 showed sensitivity of 0.726 and IST-2 showed sensitivity of 0.758	Not reported	Stokes, Banich and Ellidge (1991)
Minnesota Rate Manipulation: Placing/Manual dexterity	See Fleishman and Ellison, 1962	3 to 5 minutes/ Ipsiative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=42 minutes; reliability efficiency (3 minutes)=0.61	Bittner, Carter, Kennedy, Harbeson and Krause (1986)

\*Determined by reviewers.

## Psychomotor Task (Continued)

Measure/ Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Minnesota Rate of Manipulation: Turning/ Manual dexterity	See Fleishman and Ellison, 1962	2 to 4 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=10 minutes; reliability efficiency (3 minutes)=0.64	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Nonpreferred- Hand Tap/ Coordination	Computer	Practice time: 10 seconds. Trial time: 10 seconds. Total task time, less practice: 20 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to impairment from alcohol	Not reported	Kennedy, Dunlap, Turnage, Rugotzke, Wilkes and Smith (1991)
Nonpreferred- Hand Tap/ Coordination	Computer	Practice time: 10 seconds. Trial time: 10 seconds. Total task time, less practice: 300 seconds/Ipsative*	Experimental test using introductory psychology class	Intermode correlations between computer-administered subtests and WAIS-R, Verbal, Performance, and Full-Scale IQ (N=25) showed correlations with verbal (0.48, p<0.01), performance (0.50, p<0.01), and full-scale (0.63, p<0.05)	Not reported	Stability r=0.90; 3- minute reliability efficiency=0.96	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Nonpreferred- Hand Tap/ Coordination	Computer	Practice time: 10 seconds. Trial time: 10 seconds. Total task time, less practice: 20 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to alcohol; F(6,102) ratios for impact of alcohol on performance=7.57; p<0.00	Not reported	Kennedy, Turnage and Dunlap (1992)
Nonpreferred- Hand Tap/Eye- hand coordination	Computer	Not reported	Experimental test (see Turnage, Kennedy, Smith, Baltzley and Lane, 1992)	Predictive validity cross- correlations were not significant for nonpreferred-hand tap	Sensitivity for altitude, motion sickness drugs, and chemotherapy	Six metrology studies, range for stability=0.88 to 0.98	Turnage, Kennedy, Smith, Baltzley and Lane (1992)
One-Leg Stand (NHTSA Standardized 3-Test Battery)/ Balance, coordination	Verbal commands and observations by police officers	"Short period of time" /Normative*	Experimental research using drivers suspected of driving while intoxicated	See Tharp, Burns and Moskowitz, 1981	See Tharp, Burns and Moskowitz, 1977	See Tharp, Burns and Moskowitz, 1981	Burns (1985)

\*Determined by reviewers.

## Appendix G

### Psychomotor Task (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
One-Leg Stand/Balance	Observation	30 seconds per trial/Normalized*	Experimental test using 62 light, 86 moderate, and 90 heavy drinkers	10 officers' scoring of one-leg stand test correlated with BAC as 0.484	Based on scoring by 10 officers, the mean for one-leg stand below 0.10% BAC was 2.01, and the mean for one-leg stand equal to or above 0.10% BAC was 4.61. Based on scoring by 2 observers, the mean for one-leg stand below 0.10% BAC was 2.14, and the mean for one-leg stand equal to or above 0.10% BAC was 4.47.	Interrater reliability between 10 police officers and 2 observers were as follows for one-leg stand: test score=0.77 for participants below 0.10% BAC and 0.81 for participants equal to or above 0.10% BAC	Burns and Moskowitz (1977)
One-Leg Stand/Balance, walking	Verbal instruction and observation	Not reported/Normalized*	Experimental test using light, moderate, and heavy drinkers	Not reported	Sensitive to alcohol; correlation with BAC was 0.48	Not reported	Burns and Moskowitz (1979)
One-Leg Stand/Balance	Not reported	Not reported	Experimental (see Kennedy, Turnage, and Dunlap, 1993)	Not reported	Sensitive to alcohol $F=32.9$	Not reported	Kennedy, Turnage and Dunlap (1993)
Preferred-Hand Tap/Coordination	Computer	Practice time: 10 seconds. Trial time: 10 seconds. Total task time, less performance: 300 seconds/ Ipsiactive*	Experimental test using students from an introductory psychology class	Intermode correlations between computer-administered subtests and WAIS-R, Verbal, Performance, and Full-Scale IQ (N=25) showed correlations with verbal (0.41, $p<0.01$ ), performance (0.42, $p<0.01$ ), and full scale (0.55, $p<0.05$ )	Not reported	Stability $r=0.89$ ; 3-minute reliability efficiency=0.94	Kennedy, Wilkes, Dunlap and Kuniz (1987)
Preferred-Hand Tap/Eye-hand coordination, cognition	Computer	Not reported	Experimental test. See Turnage, Kennedy, Smith, Balfzley and Lane, 1992	No predictive validity cross-correlations were reported for this task	Not reported	Seven metrology studies. Range for stability=0.76 to 0.99	Turnage, Kennedy, Smith, Balfzley and Lane (1992)

\*Determined by reviewers.

## Psychomotor Task (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Purdue Pegboard/Fine finger dexterity	See Tiffin, 1968	2 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=42 minutes; reliability efficiency (3 minutes)=0.90	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Spoke-Speed of arm movement	Paper and pencil	Practice time: 15 seconds. Trial time: 30 seconds. Total task time, less practice: 60 seconds/Ipsative*	Experimental test using students from an introductory psychology class	Intermode correlations between paper-and-pencil and computer-administered subtests (N=25) showed correlations with preferred-hand tap (0.50, p<0.01), nonpreferred-hand tap (0.54, p<0.05), two-hand tap (0.47, p<0.01), and code substitution (0.51, p<0.01)	Not reported	Stability r=0.91; 3-minute reliability efficiency=0.97	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Spoke: Control Task-Speed of arm movement	See Fleishman and Ellison, 1962	0.67 minutes/ Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=1 minute; reliability efficiency (3 minutes)=0.95	Bitner, Carter, Kennedy, Harbeson and Krause (1986)
Tracing Mazes/ Coordination	Paper and pencil	20 seconds per maze, or 1 minute per subject/ Ipsative*	Experimental test using light, moderate, and heavy drinkers	10 officers' scoring of the tracing maze test correlated with BAC as 0.439	Not reported	Not reported	Burns and Moskowitz (1977)
Two-Hand Tap/ Coordination	Computer	Practice time: 10 seconds. Trial time: 10 seconds. Total task time, less practice: 20 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to impairment from alcohol	Not reported	Kennedy, Dunlap, Turnage, Rugotzke, Wilkes and Smith (1991)

\* Determined by reviewers.

## Appendix G

### Psychomotor Task (Continued)

Measure/ Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Two-Hand Tap/ Coordination	Computer	Practice time: 10 seconds. Trial time: 10 seconds. Total task time, less practice: 20 seconds/Ipsative*	Experimental test using male university students	Not reported	Sensitive to alcohol. F(6,102) ratios for impact of alcohol performance=15.65; p<0.00	Not reported	Kennedy, Turnage and Dunlap (1992)
Two-Hand Tap/ Coordination	Computer	Not reported	Experimental (see Kennedy, Turnage and Dunlap, 1993)	Not reported	Sensitive to alcohol F=16.8	Not reported	Kennedy, Turnage and Dunlap (1993)
Two-Hand Tap/ Coordination	Computer	Practice time: 10 seconds. Trial time: 10 seconds. Total task time, less practice: 300 seconds/Ipsative*	Experimental test using students from an introductory psychology class	Intermode correlations between computer-administered subtests and WAIS-R, Verbal, Performance, and Full-Scale IQ (N=25) showed no significant correlations	Not reported	Stability r=0.89; 3- minute reliability efficiency=0.94	Kennedy, Wilkes, Dunlap and Kuntz (1987)
Two-Handed Tap/ Coordination	Computer	Not reported	Experimental test. See Turnage, Kennedy, Smith, Baltzley and Lane, 1992	Predictive validity cross- correlations were not significant for two-handed tapping	Showed sensitivity for altitude and for chemotherapy	Six metronome studies. Range for stability=0.73 to 0.96	Turnage, Kennedy, Smith, Baltzley and Lane (1992)
Unstable Tracking Task/ Eye-hand coordination	Computer: administered with joystick	3 minutes per trial/ Ipsative*	Experimental test using military personnel	Not reported	Positive feedback tracking showed sensitivity to alcohol, carbon monoxide, and seocobarbital	Authors report that the reliability and stability of critical tracking tasks are dependent on the effects of practice, but no reliability data based on average error or number of control lapses per trial have been found.	Perez, Masline, Ramsey and Urban (1987)

\*Determined by reviewers.

## Psychomotor Task (Continued)

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Walk-and-Turn/Balance, coordination	Observation	Not reported	Experimental test using light, moderate, and heavy drinkers	10 officers' scoring of this task correlated with BAC as 0.547	Based on scoring by 10 officers, the mean for walk-and-turn below 0.10% BAC was 2.71, and the mean for walk-and-turn equal to or above 0.10% BAC was 5.75. Based on scoring by 2 observers, the mean for walk-and-turn below 0.10% BAC was 2.92, and the mean for walk-and-turn equal to or above 0.10% BAC was 5.75	Interrater reliability between 10 police officers and 2 observers were as follows for walk-and-turn: test score=0.70 for participants below 0.10% and 0.84 for participants equal to or above 0.10% BAC	Burns and Moskowitz (1977)
Walk-and-Turn/Balance, coordination	Verbal instruction and observation	Not reported	Experimental test using light, moderate, and heavy drinkers	Not reported	Sensitive to alcohol; correlation with BAC was 0.55	Not reported	Burns and Moskowitz (1979)
Walk-and-Turn	Not reported	Not reported	Experimental (see Kennedy, Turnage and Dunlap, 1993)	Not reported	Sensitive to alcohol F=18.5	Not reported	Kennedy, Turnage and Dunlap (1993)
Walk-and-Turn Test (NHTSA Standardized 3-Test Battery)/Balance, coordination	Observation by police officers	"Short period of time"/Normative*	Field test using drivers suspected of driving while intoxicated	See Tharp, Burns and Moskowitz, 1981	See Burns and Moskowitz, 1977	See Tharp, Burns and Moskowitz, 1981	Burns (1985)

\*Determined by reviewers.

## Appendix G

### Physiological Task: Vision and Visual Acuity

Measure/Construct	Method of Administration	Administration Time/Scoring	User of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
<b>Autokinesis/Visual autokinesis</b>	Apparatus	25 minutes per trial/ Ipsative*	Experimental test using moderate alcohol and marijuana users	Not reported	Sensitivity to alcohol not significant; sensitivity to marijuana significant at all three levels	Not reported	Moskowitz, Sharma and Schapero (1972)
<b>Dark Adaptation Test/Light adaptation</b>	Apparatus	1.5 hours at three different test sessions/ Ipsative*	Experimental test using moderate marijuana users	Not reported	Both alcohol and marijuana had the effect of reducing the time required by the subjects to adapt to dark	Not reported	Moskowitz, Sharma and Schapero (1972)
<b>Duction Test/ Lateral and vertical range of fusional vergence eye movements</b>	Apparatus	Not reported	Experimental test using moderate marijuana users	Not reported	Treatment effects for supraduction, infraduction, and adduction were nonsignificant. However, the range within which subjects could see singly was significantly smaller for the abduction condition as a result of drug treatments	Not reported	Moskowitz, Sharma and Schapero (1972)
<b>Peripheral Signal Detection/ Peripheral attention</b>	Apparatus	108 20-second test trials separated by 5-second intervals and 3 5-minute rest periods for a total test time of 1 hour/ Ipsative*	Experimental test involving occasional marijuana users	Not reported	Use of marijuana increases error rate	A subsequent replication of this study failed to demonstrate any influence of marijuana on reaction time	Moskowitz, Sharma and Schapero (1972)
<b>Pupillometry/ Ocular movement</b>	Pupil screen by Applied Science Laboratories, Waltham, MA	20 times per second over a 3 second interval. Total task time 3 minutes.	Experimental using occasional drug and alcohol users	Not reported	For amphetamines $F=7.41$ $p<0.01$ ; for diazepam $F=5.95$ $p<0.05$ ; sensitive to alcohol and marijuana, but no F values given.	Not reported	Kelly, Foltin, Emurian and Fischman (1993)

\*Determined by reviewers.

## Physiological Task: Vision and Visual Acuity (Continued)

Measure/ Construct	Method of Administration Time/Scoring	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Test of Lateral and Vertical phoria/Vertical phoria	Apparatus	Not reported	Experimental test using occasional marijuana users	Not reported	Neither alcohol nor marijuana had significant effect upon vertical phoria, but there were significant effects of the treatments on lateral phoria ( $F=8.25$ , $df=2,5$ , $p<.05$ ). Placebo comparison with alcohol produced a $t=3.40$ , $df=11$ , $p<.01$ , and comparison with marijuana produced a $t=4.65$ , $df=11$ , $p<.01$	Not reported	Moskowitz, Sharma and Schaper (1972)
Visual Acuity Test/Visual acuity	Apparatus	Not reported	Experimental test involving occasional marijuana users	Not reported	Treatments had no effect on visual acuity	Not reported	Moskowitz, Sharma and Schaper (1972)
Visual Contrast Sensitivity: Method of Increasing Contrast/Contrast sensitivity function: 1 cycle per degree	Paper and pencil (see Ginsburg and Evans, 1982)	3 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Total stabilization time=1 minute; reliability efficiency (3 minutes)=0.51	Not reported	Bittner, Carter, Kennedy, Harbeson and Krause (1986)
Visual Contrast Sensitivity: Method of Increasing Contrast/Contrast sensitivity function: 2 cycle per degree	Paper and pencil (see Ginsburg and Evans, 1982)	3 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Total stabilization time=1 minute; reliability efficiency (3 minutes)=0.52	Not reported	Bittner, Carter, Kennedy, Harbeson and Krause (1986)

\*Determined by reviewers.

## Appendix G

### Physiological Task: Vision and Visual Acuity (Continued)

Measure/Construct	Method of Administration	Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Visual Contrast Sensitivity: Method of Increasing Contrast/Contrast sensitivity function: 4 cycle per degree	Paper and pencil (see Ginsburg and Evans, 1982)	3 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=1 minute; reliability efficiency (3 minutes)=0.74	Bittnner, Carter, Kennedy, Harbeson and Krause (1986)
Visual Contrast Sensitivity: Method of Increasing Contrast/Contrast sensitivity function: 8 cycle per degree	Paper and pencil (see Ginsburg and Evans, 1982)	3 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=1 minute; reliability efficiency (3 minutes)=0.75	Bittnner, Carter, Kennedy, Harbeson and Krause (1986)
Visual Contrast Sensitivity: Method of Increasing Contrast/Contrast sensitivity function: 16 cycle per degree	Paper and pencil (see Ginsburg and Evans, 1982)	3 minutes/Ipsative*	Experimental test using Navy enlisted men	Not reported	Not reported	Total stabilization time=1 minute; reliability efficiency (3 minutes)=0.53	Bittnner, Carter, Kennedy, Harbeson and Krause (1986)
Worth Four Dot Test/Binocular vision	Apparatus	Not reported	Experimental test using occasional marijuana users	Not reported	A chi square analysis showed no significance of any treatment for this test	Not reported	Moskowitz, Sharma and Schapiro (1972)

\*Determined by reviewers.

## Physiological Task: Ocular Movement

## Appendix G

Measure/Construct	Method of Administration	Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Alcohol Gaze Nystagmus/Ocular movement	Pen light, pen or finger, and observation	Not reported	Experimental test involving 62 light, 86 moderate, and 90 heavy drinkers	10 officers' scoring of total nystagmus test correlated with BAC as 0.668	Based on scoring by 10 officers, the mean for total nystagmus below 0.10% BAC was 2.30, and the mean for total nystagmus equal to or above 0.10% BAC was 9.76. Based on scoring by 2 observers, the mean for total nystagmus below 0.10% BAC was 2.34, and the mean for total nystagmus equal to or above 0.10% BAC was 10.71.	Interrater reliability between 10 police officers and 2 observers were as follows for total nystagmus: test score=0.88 for participants below 0.10% BAC and 0.78 for participants equal to or above 0.10% BAC	Burns and Moskowitz (1977)
Alcohol Gaze Nystagmus/Ocular movement	Protractor, observation	Not reported	Emergency room patients and drivers suspected of being intoxicated	Not reported	Authors report that test has a sensitivity of 96%. In DWI suspects, there was a correlation of 0.878 between estimated BAC and BAA results	Not reported	Goding and Dobie (1986)
Calibration/ Ocular movement	Electro-nystagmograph (ENG)	Not reported	Not reported	Not reported	Not reported	Not reported	Westerman and Gilbert (1981)
Cold Caloric Test/Ocular movement	Electro-nystagmograph (ENG)	Not reported	Not reported	Not reported	Sensitive to methadone hydrochloride, which showed a depressing effect during cold caloric testing	Not reported	Westerman and Gilbert (1981)
EM/2 Alcohol and Drug Screener/Ocular movement	Computer	8 to 14 minutes for total test time/1positive*	Experimental study using three subject samples: N=16 males administered cocaine, N=13 males administered marijuana, and N=90 males and females administered alcohol	Not reported	Subjects given cocaine showed significant changes in pupillary response. Subjects given alcohol showed impairment for smooth pursuit measures and horizontal gaze nystagmus. There was no significant effect for marijuana.	The test-retest reliability of predrug baseline scores at 1- to 2-week intervals was higher than r=0.80 for most of the measures (full range: r=0.60 to 0.99)	Wilkinson and Burns (1992)

\*Determined by reviewers.

## Appendix G

### Physiological Task: Ocular Movement (Continued)

Measure/Construct	Method of Administration	Administration/Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Gaze Nystagmus/ Ocular movement	Verbal instruction, observation	Not reported/ Normative*	Experimental test using light, moderate, and heavy drivers	Not reported	Sensitive to alcohol; correlation with BAC for alcohol gaze nystagmus was 0.67	Not reported	Burns and Moskowitz (1979)
Gaze Nystagmus/ Ocular movement	Penlight or target object, observation	Not reported/ Normative*	Used among commercial and private drivers suspected of being intoxicated or otherwise impaired	Not reported	Sensitive to impairment from alcohol, barbiturates, inhalants, and phenacyclidine (PCP)	Not reported	Forkiotis (1987)
Gaze Nystagmus/ Ocular movement	Not reported	Not reported	Experimental (see Kennedy, Turnage and Dunlap, 1993)	Not reported	Sensitive to alcohol F=48.1	Not reported	Kennedy, Turnage and Dunlap (1993)
Gaze Nystagmus/ Ocular movement	Penlight, observation	Not reported/ Normative*	Experiment using drivers suspected of intoxication	Not reported	Authors cite studies that have shown alcohol effects on nystagmus and smooth pursuit breakdown. Barbiturates, antihistamines, and phenacyclidine also effect nystagmus, but fatigue does not seem to influence nystagmus.	Not reported	Tharp (1981)
Gaze Nystagmus/ Ocular movement	Fixation light mounted on a protractor	Two 10-hour sessions and two 9-hour sessions/ Normative*	Experiment using 10 moderately heavy drinkers	Not reported	Alcohol significantly decreased the angle of onset by approximately 15 degrees. In addition, a significant alcohol by time interaction was seen after midnight, when alcohol decreased the angle of onset by approximately 20 degrees.	Not reported	Tharp, Moskowitz and Burns (1981)
Gaze Nystagmus/ Ocular movement	Electronystagmograph (ENG)	Not reported/ Normative*	Not reported	Not reported	Not reported	Not reported	Westerman and Gilbert (1981)
Gaze Nystagmus (NHTSA Standardized 3-Test Battery)/ Ocular movement	Observation by police officers *	"Short period of time"/Normative *	Field test involving drivers suspected of being intoxicated	Not reported	Sensitive to impairment by alcohol (see Burns and Moskowitz, 1977)	Not reported	Burns (1985)

\*Determined by reviewers.

## Physiological Task: Ocular Movement (Continued)

Measure/ Construct	Method of Administration Time/Scoring	Use of Measure	Predictive/Construct Validity	Sensitivity	Reliability	Reference
Horizontal Gaze Nystagmus/Ocular movement	Computer Not reported/ Normative*	Experimental research in a marine environment.	Not reported	All three factors were significant for alcohol	Not reported	Burns (1992)
Horizontal Gaze Nystagmus/Ocular movement	Observation Not reported/ Normative*	Field study using drivers stopped for various driving violations during a 5-month period	Over 92% of subjects scoring 4 points or higher on this test registered BACs above 0.10%	Not reported	Not reported	Good and Augsburger (1986)
Optokinetic Testing/Ocular movement	Electro- nystagmo- graph (ENG)	Not reported	Not reported	Diazepam depressed the central system, as seen with optokinetic and pendular tracking tests	Not reported	Westerman and Gilbert (1981)
Pendular Tracking/Ocular movement	Electro- nystagmo- graph (ENG)	Not reported	Not reported	Diazepam depressed the central system, as seen with optokinetic and pendular tracking tests. Ketamine hydrochloride (similar chemically to PCP) produced a pendular nystagmus which approached 500 to 600 beats per minutes, with lack of slow or fast component	Not reported	Westerman and Gilbert (1981)
Positional Alcohol Nystagmus (PAN)/Ocular movement	Electro- nystagmo- graph (ENG)	Not reported/ Normative*	Not reported	This test showed sensitivity to alcohol over 5 minutes, 15 minutes, and 1 hour	Not reported	Westerman and Gilbert (1981)
Sashaying (Positional Maintenance Testing)/Ocular movement	Electro- nystagmo- graph (ENG)	Not reported	Not reported	Sensitive to diazepam, which had a depressor effect on the positional maintenance system and the vestibular end organ	Not reported	Westerman and Gilbert (1981)
Spinning Test/ Ocular movement	Electro- nystagmo- graph (ENG)	Not reported	Not reported	Continued nystagmus with eyes open during the spinning test revealed a defect in the central suppression mechanism brought on by diazepam	Not reported	Westerman and Gilbert (1981)

\*Determined by reviewers.

## Appendix G

### References

Allen, W., Silverman, M. and Itkonen, M. (1992). *Real World Experience in FFD Testing*. Warrendale, PA: SAE International.

Allen, W., Stein, A. and Jex, H. (1981). *Detecting Human Operator Impairment with a Psychomotor Task*. 17th Annual Conference on Manual Control, Los Angeles, CA.

Allen, W., Stein, A. and Miller, J. (1990). *Performance Testing as a Determinant of Fitness-for-Duty*. SAE Technical Paper Series (Aerospace Technology Conference and Exposition: Long Beach, CA; October 1-4, 1990). SAE International--The Engineering Society for Advancing Mobility Land Sea Air and Space.

Bittner, A.C., Carter, R.C., Kennedy, R.S., Harbeson, M.M. and Krause, M. (1986). Performance evaluation tests for environmental research (PETER): Evaluation of 114 measures. *Perceptual and Motor Skills*, 63: 683-708.

Burns, M. (1985). Field sobriety tests: An important component of DUI. *Alcohol, Drugs, and Driving*, 1(3): 21-25.

Burns, M. (1992). *A Study of Field Sobriety Tests in the Marine Environment*. Final Report. Los Angeles, CA: Author.

Burns, M. and Moskowitz, H. (1971). Response time to a first signal as a function of time relationship to a second signal and mode of presentation. *Perceptual and Motor Skills*, 32: 811-816.

Burns, M. and Moskowitz, H. (June 1977). *Psychophysical Tests for DWI Arrest* (Final Report: DOT-HS-5-01242). Washington, DC: Department of Transportation (DOT), National Highway Traffic Safety Administration.

Burns, M. and Moskowitz, H. (1979). Alcohol-impairment tests for DWI arrests. *Transportation Research Record: Driver Performance and Passenger Safety Devices*, 739: 1-4. Transportation Research Board National Research Council.

Burns, M. and Moskowitz, H. (1981). Alcohol, marijuana and skills performance. In L. Goldberg (Ed.), *Alcohol, Drugs, and Traffic Safety*, 3: 954-967. Stockholm, Sweden: Almqvist and Wiksell International.

Forkiotis, C.J. (1987). Optometric expertise: The scientific basis for alcohol gaze nystagmus. *Curriculum II*, 59(7): 1-15.

Goding, G.S. and Dobie, R.A. (July 1986). Gaze nystagmus and blood alcohol. *The Laryngoscope* 96: 713-717.

Good, G.W. and Augsburger, A.R. (1986). Use of horizontal gaze nystagmus as a part of roadside sobriety test. *American Journal of Optometry and Physiological Optics*, 63(6): 467-471.

Guilford, J.P. and Hoepfner, R. (1971). *The Analysis of Intelligence*. New York: McGraw-Hill.

Kelly, T., Folten, R., Emurian, C. and Fischman, M. (1993). Performance-based testing for drugs of abuse: Dose and time profiles of marijuana, amphetamine, alcohol, and diazepam. *Journal of Analytical Toxicology*, 17: 265-273.

Kennedy, R., Dunlap, W., Turnage, J., Rugotzke, G., Wilkes, R. and Smith, M. (July 30, 1991). *Improving Productivity by Dose Equivalency Modeling* (NSF Award No. IDI-9061181). Orlando, FL: Essex Corporation.

Kennedy, R., Turnage, J. and Dunlap, W. (1992). The use of dose equivalency as a risk assessment index in behavioral neurotoxicology. *Neurotoxicology and Teratology*, 14: 167-175.

Kennedy, R., Turnage, J. and Dunlap, W. (1993). Diagnosis of alcohol intoxication: Effectiveness of cognitive and neurovestibular field sobriety tests. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, 964-968.

Kennedy, R.S., Wilkes, R.L., Dunlap, W.P. and Kuntz, L.A. (1987). Development of an automated performance test system for environmental and behavioral toxicology studies. *Perceptual and Motor Skills*, 65: 947-962.

Lyvers, M.F. and Maltzman, I. (1991). Selective effects of alcohol on Wisconsin card sorting test performance. *British Journal of Addiction*, 86: 399-407.

Moskowitz, H. and Burns, M. (1973). Alcohol effects on information processing time with an overlearned task. *Perceptual and Motor Skills*, 37: 835-839.

Moskowitz, H. and Roth, S. (1971). Effect of alcohol on response latency in object naming. *Quarterly Journal of Studies on Alcohol*, 32(4): 969-975.

Moskowitz, H., Sharma, S. and McGlothlin, W. (1972). Effect of marihuana upon peripheral vision as a function of the information processing demands in central vision. *Perceptual and Motor Skills*, 35: 875-882.

Moskowitz, H., Sharma, S. and Schapero, M. (1972). A comparison of the effects of marijuana and alcohol on visual functions. *Current Research in Marijuana*, 35: 129-151.

Perez, W.A., Masline, P.J., Ramsey, E.G. and Urban, K.E. (1987). (UTC-PAB Test Nos. 1-25). *United Tri-Services Cognitive Performance Assessment Battery: Review and Methodology*. (Report No. AAMRL-TR-87-007). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Stein, A.C., Allen, R.W. and Jex, H.R. (1987). The development and use of in-vehicle devices to reduce alcohol and drug impaired driving. *Alcohol, Drugs, and Traffic Safety*. Elsevier Science Publishers B.V. (Biomedical Division), 601-604.

Stokes, A.F., Banich, M.T. and Elledge, V.C. (1991). Testing the tests: An empirical evaluation of screening tests for the detection of cognitive impairment in aviators. *Aviation, Space, and Environmental Medicine*, 62(8): 783-788. Washington, DC: Aerospace Medical Association.

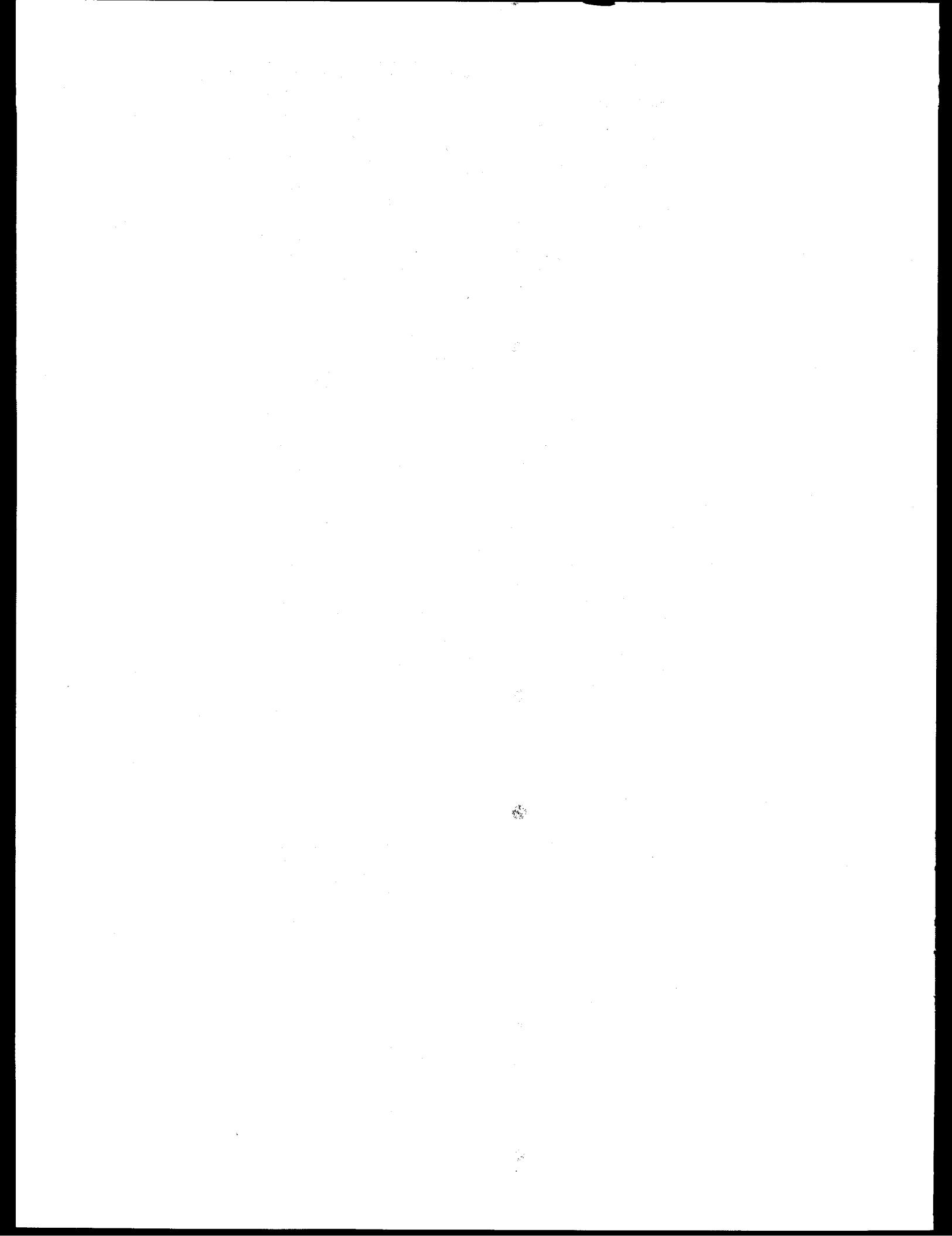
Tharp, V.A. (February 1981). Gaze nystagmus as a roadside sobriety test. *Abstracts and Reviews in Alcohol and Driving*, 5-8.

Tharp, V.K., Moskowitz, H. and Burns, M. (1981). Circadian effects on alcohol gaze nystagmus. *Psychophysiology*, 18(2): 193.

Turnage, J.J., Kennedy, R.S., Smith, M.G., Baltzley, D.R. and Lane, N.E. (1992). Development of microcomputer-based mental acuity tests. *Ergonomics*, 35(10): 1271-1295.

Westerman, S.T. and Gilbert, L.M. (September 1981). A noninvasive method of qualitative and quantitative measurement of drugs. *The Laryngoscope* 91: 1536-1547.

Wilkinson, C. and Burns, M. (March 15, 1992). Phase I validation of EM/ 2 alcohol and drug impairment screener. *SCRI Final Report of Data Analysis for Oculokinetics, Inc.* (SCRI Project EM 86). Los Angeles, CA: Southern California Research Institute.



BIBLIOGRAPHIC DATA SHEET

(See instructions on the reverse)

2. TITLE AND SUBTITLE

Fitness for Duty in the Nuclear Industry: Update of the Technical Issues 1996

5. Edited by

N. Durbin and T. Grant, BSRC  
Contributors

A. Bittner, N. Durbin, I. Field, C. Forslund, T. Grant, J. Hauth, J. Macaulay, C. Moore,  
C. Orians, J. Toquam, M. Silbermagel, R. Wilson, BSRC

B. Baxter<sup>1</sup>, S. Kimme<sup>2</sup>, A. Zebelman<sup>3</sup>

8. PERFORMING ORGANIZATION – NAME AND ADDRESS (If NRC, provide Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address; if contractor, provide name and mailing address.)

Battelle Seattle Research Center  
4000 NE 41st Street  
Seattle, Washington 98105

Pacific Northwest National Laboratory  
Richland, Washington 99352

1. Alcohol and Drug Abuse Institute, University of Washington, Seattle, WA 98105
2. Center for the Disability Policy and Research, University of Washington, Seattle, WA 98105
3. Dynacare Laboratory of Pathology of Seattle, Inc., Seattle, WA 98104

9. SPONSORING ORGANIZATION – NAME AND ADDRESS (If NRC, type "Same as above"; if contractor, provide NRC Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address.)

Division of Reactor Program Management  
Office of Nuclear Reactor Regulation  
U.S. Nuclear Regulatory Commission  
Washington, DC 20555 -0001

10. SUPPLEMENTARY NOTES

L. Bush, NRC Project Manager

11. ABSTRACT (200 words or less)

This report provides an update of information on the technical issues surrounding the creation, implementation, and maintenance of fitness-for-duty (FFD) policies and programs. It has been prepared as a resource for Nuclear Regulatory Commission (NRC) and nuclear power plant personnel who deal with FFD programs. It contains a general overview and update on the technical issues that the NRC considered prior to the publication of its original FFD rule and the revisions to that rule (presented in earlier NUREG/CRs). It also includes chapters that address issues about which there is growing concern and/or about which there have been substantial changes since NUREG/CR-5784 was published. Although this report is intended to support the NRC's rulemaking on fitness for duty, the conclusions of the authors are their own and do not necessarily represent the opinions of the NRC.

12. KEY WORDS/DESCRIPTORS (List words or phrases that will assist researchers in locating the report.)

Fitness for Duty, Substance Abuse, Drugs, Impairment, Urinalysis, Alcohol Abuse, Alcohol Testing, Chemical Testing, Alcohol, Drug Abuse, Drug Testing, Program Performance Reports, Statistics on Drug Abuse, Statistics on Substance Abuse, Medical Review Officer, Drug-Free Workplace, Deterrence, Deterioration, Alternatives to Testing, Drug Detection, Subversion, Performance-Based Assessment Tests, Return to Work

1. REPORT NUMBER  
(Assigned by NRC. Add Vol., Supp., Rev., and Addendum Numbers, if any.)

NUREG/CR-6470  
BSRC-700/96/004  
PNNL-11134

3. DATE REPORT PUBLISHED

MONTH	YEAR
May	1996

4. FIN OR GRANT NUMBER

I2007

6. TYPE OF REPORT

7. PERIOD COVERED (Inclusive Dates)

13. AVAILABILITY STATEMENT

Unlimited

14. SECURITY CLASSIFICATION

(This Page)

Unclassified

(This Report)

Unclassified

15. NUMBER OF PAGES

16. PRICE



Federal Recycling Program