SAND2018-1329C

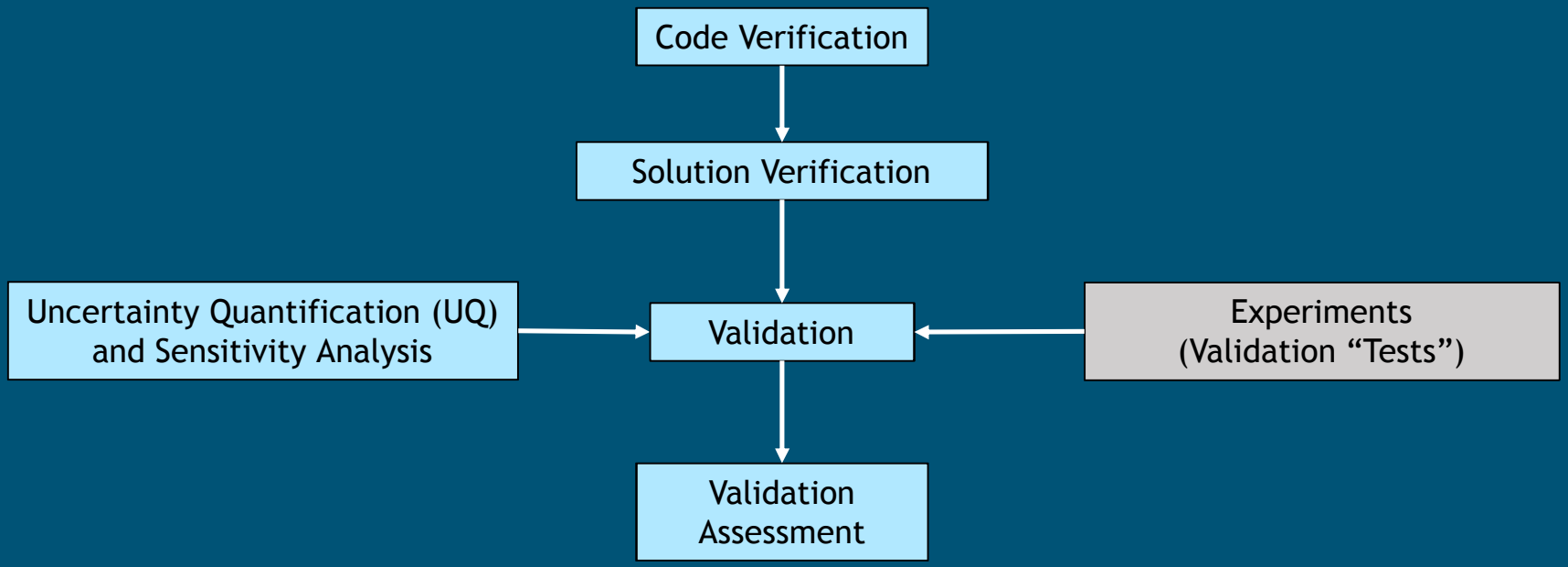
Experimental Credibility and its Role in Model Validation and Decision Making

PRESENTED BY

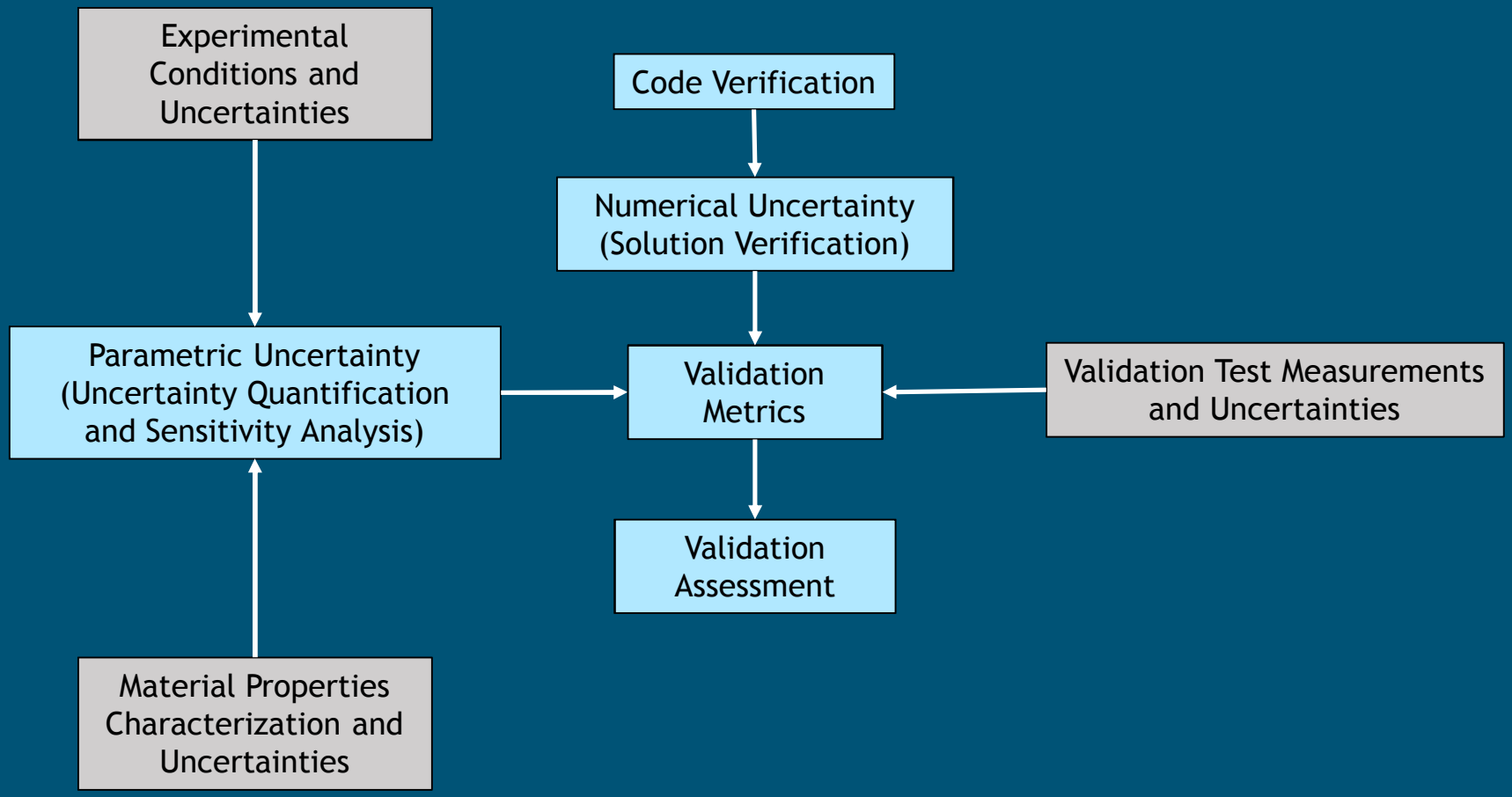
Sarah L. Kieweg PhD and Walt R. Witkowski PhD

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND2018-######

Model Validation: V&V and UQ



Role of Experiments in Model Validation and Decision Making



What is CompSim Credibility and Experimental Credibility?

CompSim Credibility – May Include:

- Trust a user or decision-maker has in the computational simulation (CompSim) outcomes.
- Assessment of correctness or completeness of MVUQ process.
  → PCMM is an assessment tool *
- Assessment of uncertainty of outcomes and predictions.
- Evidence convincingly communicated and documented.

Experimental Credibility – May Include:

- Assessment of correctness and completeness of several experimental elements.
- Assessment of experiment's use with its intended purpose (e.g. model validation).
- Assessment of uncertainty of measurements.
- Evidence convincingly communicated and documented.

* Oberkampf, W.L., M. Pilch, and T.G. Trucano, *Predictive Capability Maturity Model for Computational Modeling and Simulations*. SAND2007-5948 (UUR). 2007: Sandia National Laboratories.

Goals of Experimental Credibility Framework

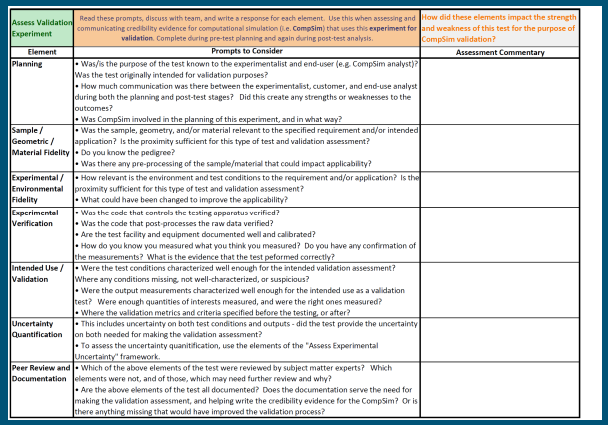
- Structured method to assess experiments used for validation:
  - Correctness
  - Completeness
  - Applicability to intended application

- Tools that encourage:
  - Early planning of validation experiments
  - Communication and documentation of experimental credibility, to aid in assessment of CompSim credibility

Proposed Tools

"Assess Validation Experiment" Tool

- Seven elements
- Team of experts and users:
  - Computational Analyst ("user" of validation test data)
  - Experimentalist
  - V&V partner
- Team discusses prompts
- Team writes assessment commentary
- "Tool" is a spreadsheet or table that team fills in

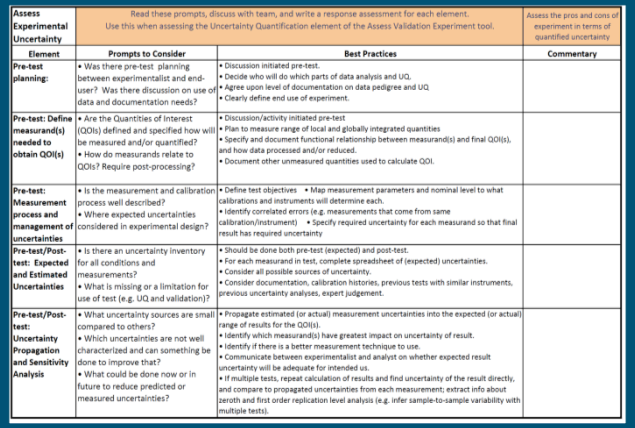


Assess Validation Experiment	Read these prompts, discuss with team, and write a response for each element. Use this when assessing and communicating credibility evidence for computational simulation (i.e. CompSim) that uses this experiment for validation. Complete during pre-test planning and again during post-test analysis.	How did these elements impact the strength and weakness of this test for the purpose of CompSim validation?
Element	Prompts to Consider	Assessment Commentary
Planning	• Was/is the purpose of the test known to the experimentalist and end-user (e.g. CompSim analyst)? Was the test originally intended for validation purposes? • How much communication was there between the experimentalist, customer, and end-use analyst during both the planning and post-test stages? Did this create any strengths or weaknesses to the outcomes? • Was CompSim involved in the planning of this experiment, and in what way?	
Sample / Geometric / Material Fidelity	• Was the sample, geometry, and/or material relevant to the specified requirement and/or intended application? Is the proximity sufficient for this type of test and validation assessment? • Do you know the pedigree? • Was there any pre-processing of the sample/material that could impact applicability?	
Experimental / Environmental Fidelity	• How relevant is the environment and test conditions to the requirement and/or application? Is the proximity sufficient for this type of test and validation assessment? • What could have been changed to improve the applicability?	
Experimental Verification	• Was the code that controls the testing apparatus verified? • Was the code that post-processes the raw data verified? • Are the test facility and equipment documented well and calibrated? • How do you know you measured what you think you measured? Do you have any confirmation of the measurements? What is the evidence that the test peformed correctly?	
Intended Use / Validation	• Were the test conditions characterized well enough for the intended validation assessment? Where any conditions missing, not well-characterized, or suspicious? • Were the output measurements characterized well enough for the intended use as a validation test? Were enough quantities of interests measured, and were the right ones measured? • Where the validation metrics and criteria specified before the testing, or after?	
Uncertainty Quantification	• This includes uncertainty on both test conditions and outputs - did the test provide the uncertainty on both needed for making the validation assessment? • To assess the uncertainty quantification, use the elements of the "Assess Experimental Uncertainty" framework.	
Peer Review and Documentation	• Which of the above elements of the test were reviewed by subject matter experts? Which elements were not, and of those, which may need further review and why? • Are the above elements of the test all documented? Does the documentation serve the need for making the validation assessment, and helping write the credibility evidence for the CompSim? Or is there anything missing that would have improved the validation process?	

Images of Tools: Details shown on subsequent slides

"Assess Experimental Uncertainty" Tool

- Addresses uncertainty quantification (UQ) element of first tool
- Five Elements



Assess Experimental Uncertainty	Read these prompts, discuss with team, and write a response assessment for each element. Use this when assessing the Uncertainty Quantification element of the Assess Validation Experiment tool.		Assess the pros and cons of experiment in terms of quantified uncertainty
Element	Prompts to Consider	Best Practices	Commentary
Pre-test planning:	• Was there pre-test planning between experimentalist and end-user? Was there discussion on use of data and documentation needs?	• Discussion initiated pre-test. • Decide who will do which parts of data analysis and UQ. • Agree upon level of documentation on data pedigree and UQ • Clearly define end use of experiment.	
Pre-test: Define measurand(s) needed to obtain QOI(s)	• Are the Quantities of Interest (QOIs) defined and specified how will be measured and/or quantified? • How do measurands relate to QOIs? Require post-processing?	• Discussion/activity initiated pre-test • Plan to measure range of local and globally integrated quantities • Specify and document functional relationship between measurand(s) and final QOI(s), and how data processed and/or reduced. • Document other unmeasured quantities used to calculate QOI.	
Pre-test: Measurement process and management of uncertainties	• Is the measurement and calibration process well described? • Where expected uncertainties considered in experimental design?	• Define test objectives • Map measurement parameters and nominal level to what calibrations and instruments will determine each. • Identify correlated errors (e.g. measurements that come from same calibration/instrument) • Specify required uncertainty for each measurand so that final result has required uncertainty	
Pre-test/Post-test: Expected and Estimated Uncertainties	• Is there an uncertainty inventory for all conditions and measurements? • What is missing or a limitation for use of test (e.g. UQ and validation)?	• Should be done both pre-test (expected) and post-test. • For each measurand in test, complete spreadsheet of (expected) uncertainties. • Consider all possible sources of uncertainty. • Consider documentation, calibration histories, previous tests with similar instruments, previous uncertainty analyses, expert judgement.	
Pre-test/Post-test: Uncertainty Propagation and Sensitivity Analysis	• What uncertainty sources are small compared to others? • Which uncertainties are not well characterized and can something be done to improve that? • What could be done now or in future to reduce predicted or measured uncertainties?	• Propagate estimated (or actual) measurement uncertainties into the expected (or actual) range of results for the QOI(s). • Identify which measurand(s) have greatest impact on uncertainty of result. • Identify if there is a better measurement technique to use. • Communicate between experimentalist and analyst on whether expected result uncertainty will be adequate for intended us. • If multiple tests, repeat calculation of results and find uncertainty of the result directly, and compare to propagated uncertainties from each measurement; extract info about zeroth and first order replication level analysis (e.g. infer sample-to-sample variability with multiple tests).	

First Tool: Assess Validation Experiment

Seven Elements in Tool:

1. Planning
2. Sample, Geometric, and/or Material Fidelity
3. Experimental & Environmental Fidelity
4. Experimental Verification
5. Intended Use
6. Uncertainty Quantification ➔ Expands to second tool: Assess Experimental Uncertainty
7. Peer Review and Documentation

"Assess Validation Experiment" Tool – *details on subsequent slides*

Assess Validation Experiment	Read these prompts, discuss with team, and write a response for each element. Use this when assessing and communicating credibility evidence for computational simulation (i.e. **CompSim**) that uses this **experiment for validation**. Complete during pre-test planning and again during post-test analysis.	How did these elements impact the strength and weakness of this test for the purpose of CompSim validation?
Element	**Prompts to Consider**	**Assessment Commentary**
Planning	• Was/is the purpose of the test known to the experimentalist and end-user (e.g. CompSim analyst)? Was the test originally intended for validation purposes? • How much communication was there between the experimentalist, customer, and end-use analyst during both the planning and post-test stages? Did this create any strengths or weaknesses to the outcomes? • Was CompSim involved in the planning of this experiment, and in what way?	
Sample / Geometric / Material Fidelity	• Was the sample, geometry, and/or material relevant to the specified requirement and/or intended application? Is the proximity sufficient for this type of test and validation assessment? • Do you know the pedigree? • Was there any pre-processing of the sample/material that could impact applicability?	
Experimental / Environmental Fidelity	• How relevant is the environment and test conditions to the requirement and/or application? Is the proximity sufficient for this type of test and validation assessment? • What could have been changed to improve the applicability?	
Experimental Verification	• Was the code that controls the testing apparatus verified? • Was the code that post-processes the raw data verified? • Are the test facility and equipment documented well and calibrated? • How do you know you measured what you think you measured? Do you have any confirmation of the measurements? What is the evidence that the test peformed correctly?	
Intended Use / Validation	• Were the test conditions characterized well enough for the intended validation assessment? Where any conditions missing, not well-characterized, or suspicious? • Were the output measurements characterized well enough for the intended use as a validation test? Were enough quantities of interests measured, and were the right ones measured? • Where the validation metrics and criteria specified before the testing, or after?	
Uncertainty Quantification	• This includes uncertainty on both test conditions and outputs - did the test provide the uncertainty on both needed for making the validation assessment? • To assess the uncertainty quanitification, use the elements of the "Assess Experimental Uncertainty" framework.	
Peer Review and Documentation	• Which of the above elements of the test were reviewed by subject matter experts? Which elements were not, and of those, which may need further review and why? • Are the above elements of the test all documented? Does the documentation serve the need for making the validation assessment, and helping write the credibility evidence for the CompSim? Or is there anything missing that would have improved the validation process?	

First Tool - Element 1: Planning

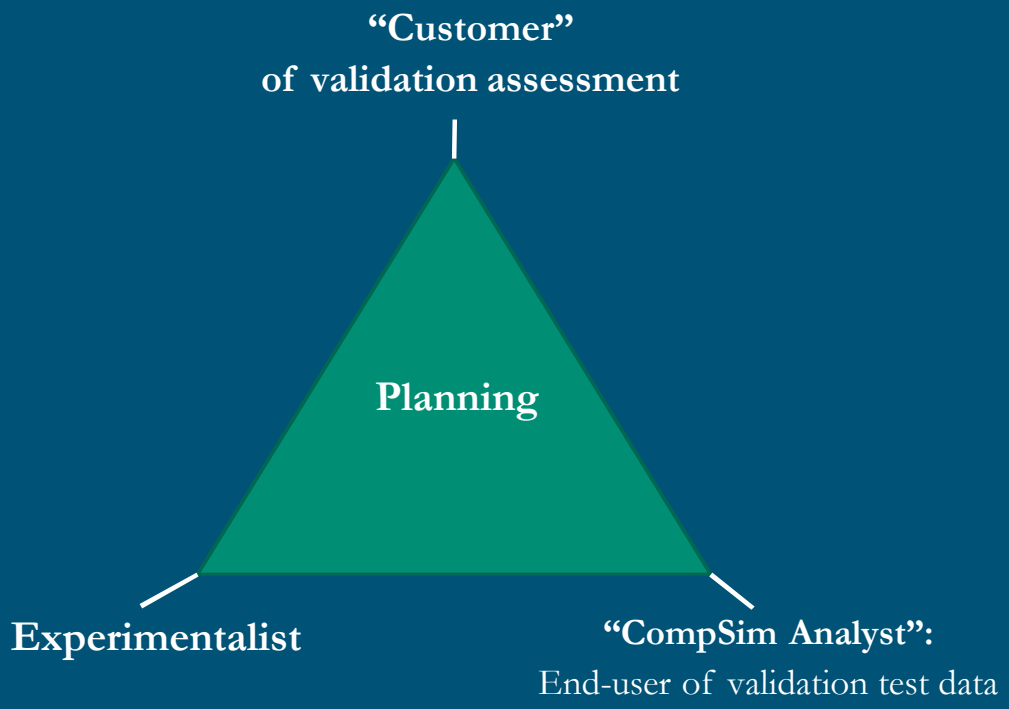
Communication

- How much communication between:
  - Experimentalist
  - Analyst
  - Customer

during planning and post-test stages?

Test Purpose

- Was test originally intended for validation?
- Who was the original end user?
- Did the experimentalist know the original purpose of the test?



Partnering with CompSim Analyst

- Was computational simulation (CompSim) involved in the design of the experiment?

First Tool –
Element 2: Sample, Geometric, and/or Material Fidelity
Element 3: Experimental & Environmental Fidelity

Pedigree of Sample

- Source and specs of sample known and documented?
- Pre-processing of sample known?

Fidelity

- Relevance to:
  - Requirements
  - Intended application
- Proximity sufficient for this validation assessment?
- What would improve applicability?

First Tool - Element 4: Experimental Verification

How do you know you measured what you think you measured?

Test Control Code
- Code that controls testing apparatus.
- Was it verified?

→

Test Equipment
- Documentation
- Calibration

→

Post-processing Code
- Post-processes raw data.
- Was it verified?

First Tool - Element 5: Intended Use

Intended Use
- Where test conditions characterized well enough for intended validation assessment?
- Where output measurements?
- Where validation metrics and criteria specified before testing, or after?

First Tool - Element 6: Experimental Uncertainty Quantification

Is the uncertainty quantified for:

- Test Conditions?
- Measurements of outputs?

This element is examined with the second tool: "Assess Experimental Uncertainty".

Why need well-defined experimental uncertainty quantification (UQ)?:

- Test condition uncertainty contributes to parametric uncertainty of simulation.
- Output measurement uncertainty needed for probabilistic-based validation metrics.

First Tool - Element 7: Peer Review and Documentation

- Assessed for each of the Elements of this tool.

Peer Review

- Which of Elements 1-5 were reviewed by subject matter experts?

Documents

- Are all of the Elements documented?
- What documentation is missing that would have improved the validation process?

Second Tool: Assess Experimental Uncertainty

Five Elements in Tool:

1. Planning of data analysis and UQ
2. Definition of measurand(s) needed to obtain QOI(s)
3. Definition of measurement process and management of uncertainties
4. Expected and estimated measurement uncertainties
5. Uncertainty propagation and sensitivity analysis

This tool feeds back into the element for Uncertainty Quantification in the first tool.

Contains best practices summarized from:

- ASME Performance Test Codes Supervisory Committee, ASME PTC 19.1-2013: *Test Uncertainty* 2013, New York, NY: ASME.
- Coleman, H.W. and W.G. Steele. *Experimentation, Validation, and Uncertainty Analysis for Engineers*. 3rd ed. 2009, New York: Wiley.

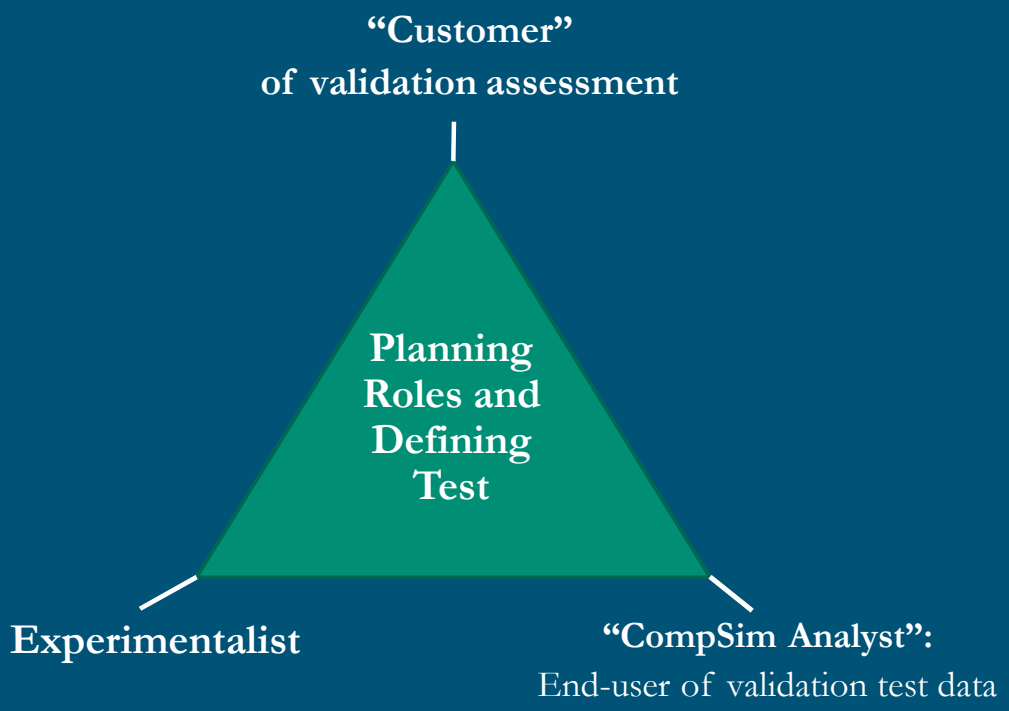
"Assess Experimental Uncertainty" Tool – *details on subsequent slides*

Assess Experimental Uncertainty	Read these prompts, discuss with team, and write a response assessment for each element. Use this when assessing the Uncertainty Quantification element of the Assess Validation Experiment tool.		Assess the pros and cons of experiment in terms of quantified uncertainty
Element	**Prompts to Consider**	**Best Practices**	**Commentary**
Pre-test planning:	• Was there pre-test planning between experimentalist and end-user? Was there discussion on use of data and documentation needs?	• Discussion initiated pre-test. • Decide who will do which parts of data analysis and UQ. • Agree upon level of documentation on data pedigree and UQ • Clearly define end use of experiment.	
Pre-test: Define measurand(s) needed to obtain QOI(s)	• Are the Quantities of Interest (QOIs) defined and specified how will be measured and/or quantified? • How do measurands relate to QOIs? Require post-processing?	• Discussion/activity initiated pre-test • Plan to measure range of local and globally integrated quantities • Specify and document functional relationship between measurand(s) and final QOI(s), and how data processed and/or reduced. • Document other unmeasured quantities used to calculate QOI.	
Pre-test: Measurement process and management of uncertainties	• Is the measurement and calibration process well described? • Where expected uncertainties considered in experimental design?	• Define test objectives • Map measurement parameters and nominal level to what calibrations and instruments will determine each. • Identify correlated errors (e.g. measurements that come from same calibration/instrument) • Specify required uncertainty for each measurand so that final result has required uncertainty	
Pre-test/Post-test: Expected and Estimated Uncertainties	• Is there an uncertainty inventory for all conditions and measurements? • What is missing or a limitation for use of test (e.g. UQ and validation)?	• Should be done both pre-test (expected) and post-test. • For each measurand in test, complete spreadsheet of (expected) uncertainties. • Consider all possible sources of uncertainty. • Consider documentation, calibration histories, previous tests with similar instruments, previous uncertainty analyses, expert judgement.	
Pre-test/Post-test: Uncertainty Propagation and Sensitivity Analysis	• What uncertainty sources are small compared to others? • Which uncertainties are not well characterized and can something be done to improve that? • What could be done now or in future to reduce predicted or measured uncertainties?	• Propagate estimated (or actual) measurement uncertainties into the expected (or actual) range of results for the QOI(s). • Identify which measurand(s) have greatest impact on uncertainty of result. • Identify if there is a better measurement technique to use. • Communicate between experimentalist and analyst on whether expected result uncertainty will be adequate for intended us. • If multiple tests, repeat calculation of results and find uncertainty of the result directly, and compare to propagated uncertainties from each measurement; extract info about zeroth and first order replication level analysis (e.g. infer sample-to-sample variability with multiple tests).	

Second Tool - Element 1: Pre-Test Planning

Best Practices:

- Agree on roles – who will do data analysis and UQ?
- Clearly define end use of experiment.
- Agree on documentation



Second Tool - Elements 2- 5

Pre-Test

2 - Define Measurands and Connections to QOI(s):

- How are they related?
- Do they require post-processing?
- Range of local and global quantities?

3 - Define Measurement Process and Manage Uncertainties:

- Are the measurement and calibration methods well described?
- Where expected uncertainties for each measurand considered in test design?

Pre-Test *and* Post-Test

4 – Expected and Estimated Uncertainties

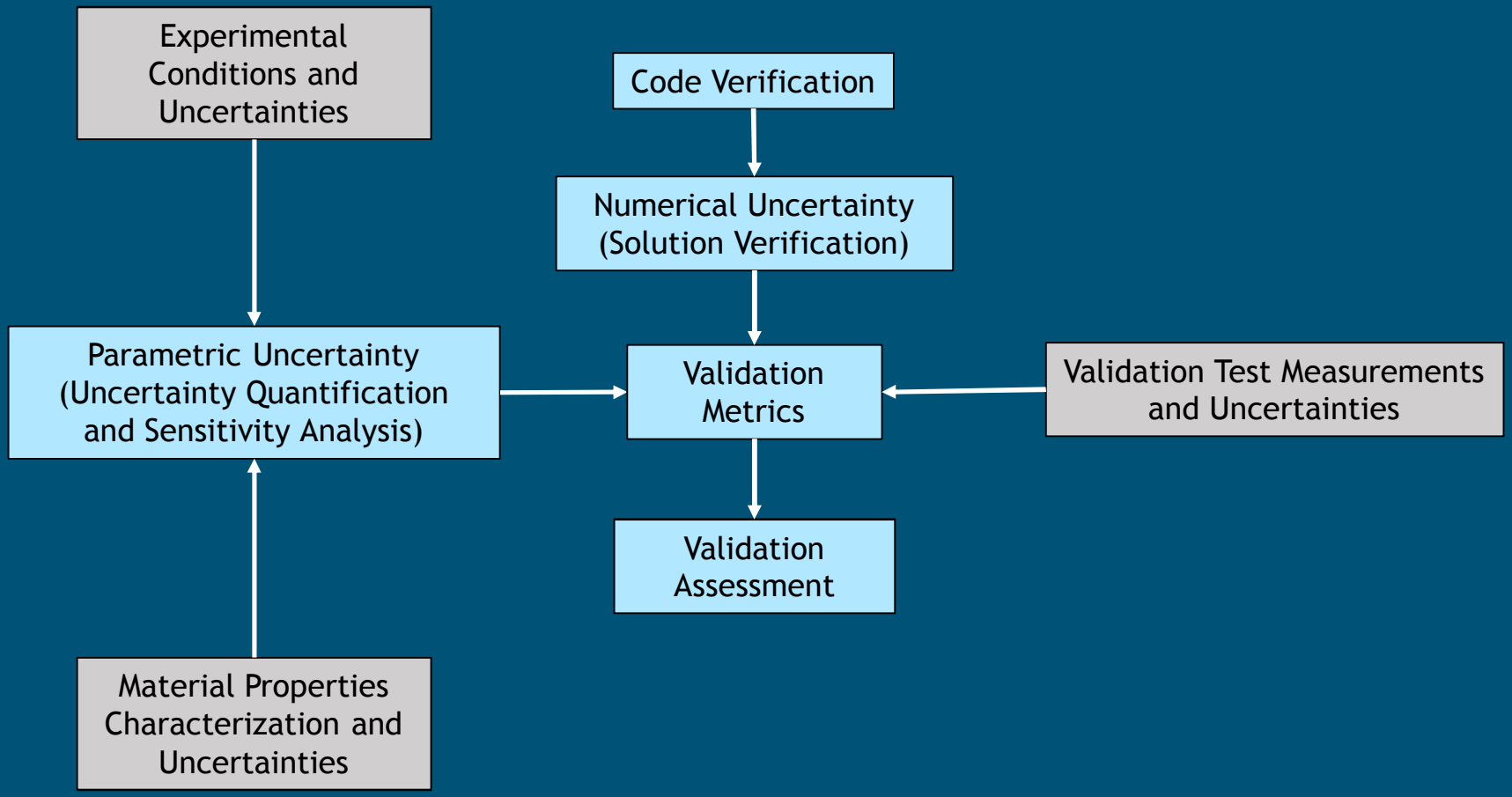
- Is there an uncertainty inventory for all conditions and measurements?
- What is missing for computational UQ and validation needs?

5 – Uncertainty Propagation and Sensitivity Analysis

- What uncertainty sources are large compared to others?
- Which uncertainties are not well characterized?
- What could be done now or in future to reduce or better define uncertainties?

Conclusions

- Anticipated outcomes of tool usage:
  - Communication of applicability and limitations of experiments.
  - Provide evidence of how experiments support overall simulation credibility.



Acknowledgements

- Martin Pilch
- Angel Urbina
- Advance Simulation and Computing (ASC) Program of the Department of Energy's National Nuclear Security Administration
- Weapons Systems Engineering Assessment Technology (WSEAT – now Delivery Environments) Program at Sandia National Laboratories