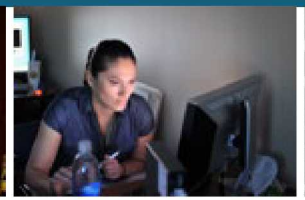
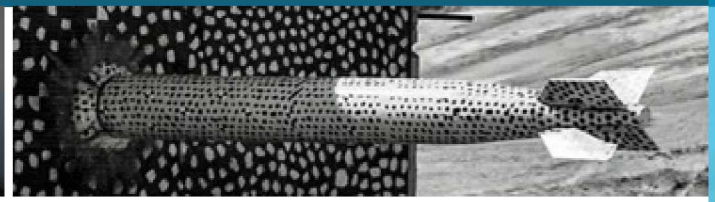
Sandia National Laboratories

SAND2020-0940PE

Experimental Credibility Overview



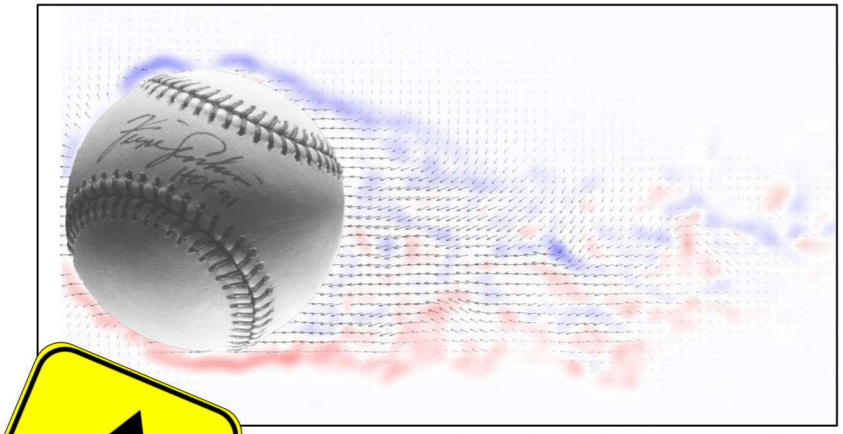


Blake W. Lance and Sarah L. Kieweg

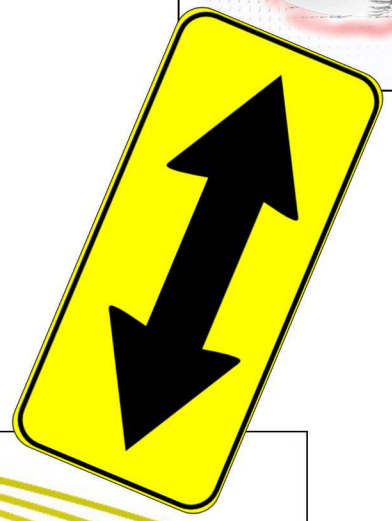
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

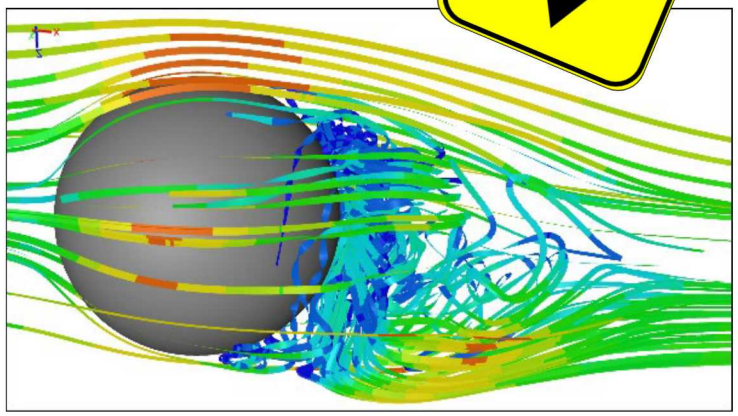
Experiments and simulations have complimentary strengths

- Experiments provide a real-world view of physics
  - Come with cost and schedule
- Simulations provide rapid insights at lower cost
  - Come with potential model form and other errors
- Experimentalists and modelers need to work collaboratively



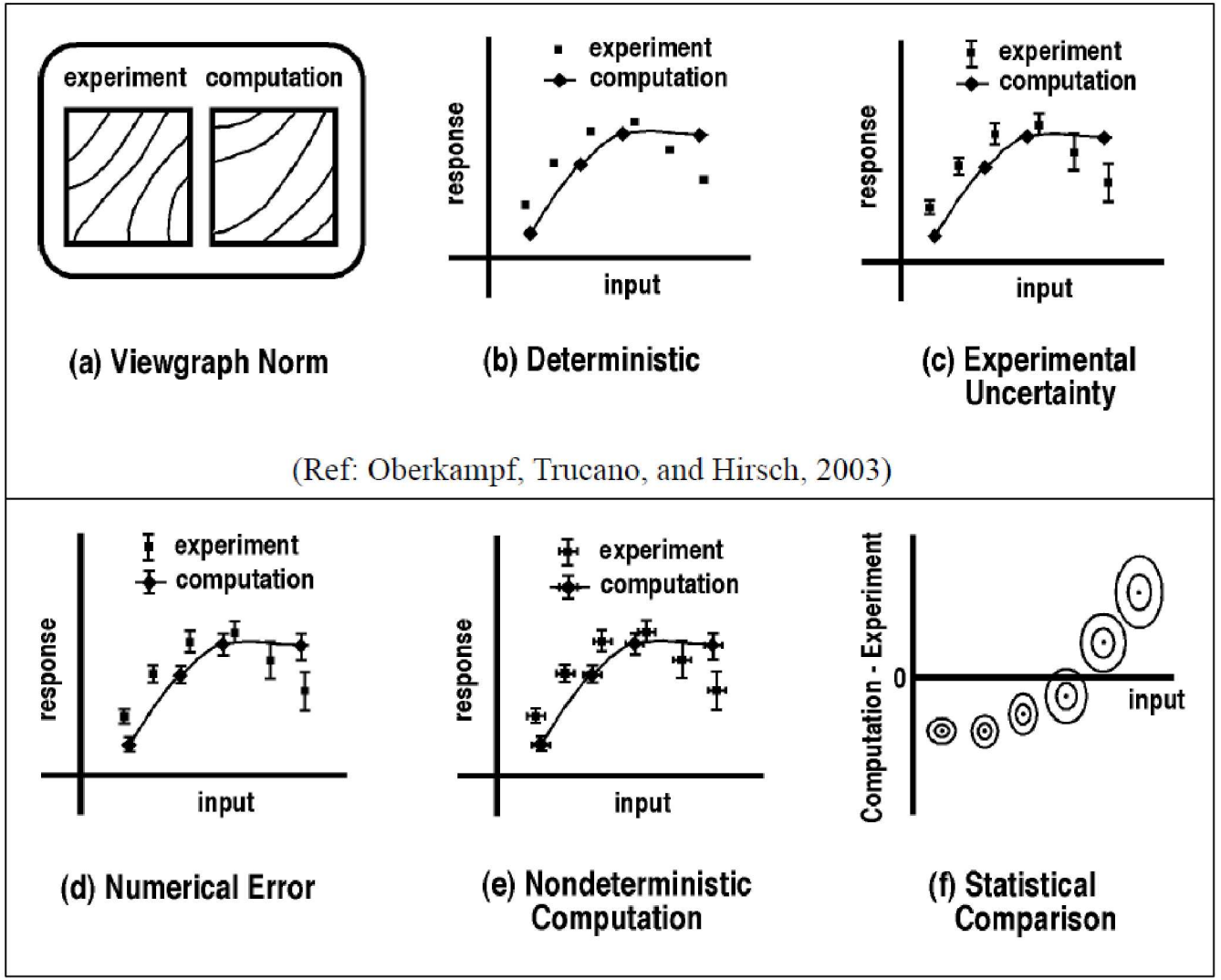
www.baseballaero.com/





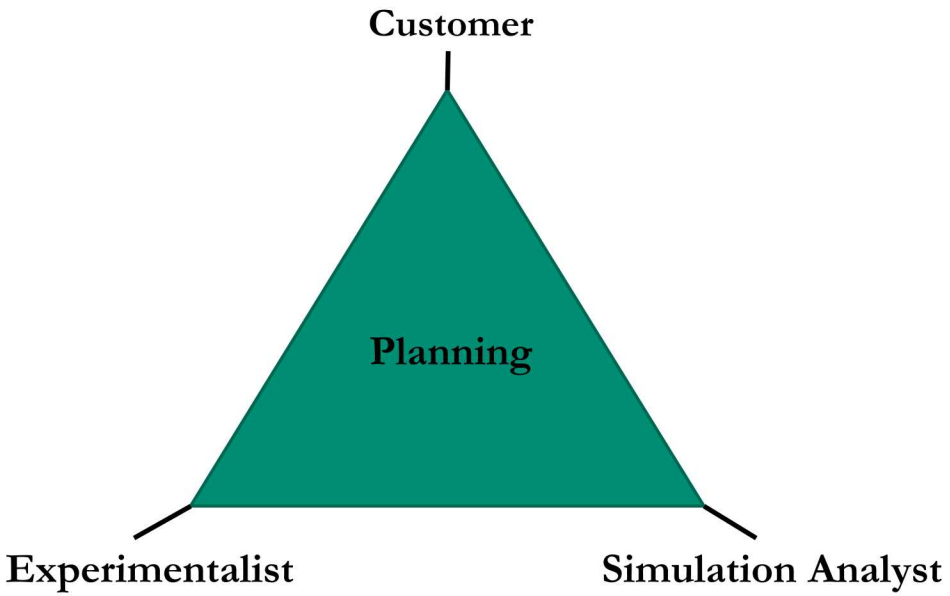
Computer simulations often use experimental data

- Data are used to
  - Calibrate model parameters
  - Validate model predictions
- Simulations inherit the quality/credibility of the experiments



The Experimental Credibility Tool guides a process

- Structured method to assess experiments used for simulations
  - Correctness
  - Completeness
  - Applicability to intended use
- Process that encourages
  - Early planning of experiments
  - Communication between stakeholders: experimentalists, analysts, system stakeholders
  - Documentation of experimental credibility, to aid simulation credibility



Tool: Plan and Assess Experiment

- "Plan and Assess Experiment" Tool
  - "Tool" is a spreadsheet or table that team completes
  - Seven elements
  - Team of experts and users:
    - Computational Analysts
    - Experimentalists
    - Customer
    - V&V partner
  - Team discusses prompts; strengths/weaknesses
  - Team writes assessment commentary
  - Team identifies action items
- Elements
  - Planning
  - Intended Use
  - Sample, Geometric, and/or Material Fidelity
  - Experimental & Environmental Fidelity
  - Experimental Verification
  - Uncertainty Quantification
  - Peer Review and Documentation

Plan and Assess 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**. 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 intended use?
Element	**Prompts to Consider**	**Assessment Commentary**
Planning	• Is purpose of the test known to the experimentalist and end-user (e.g. CompSim analyst)? What is the intended use of test? • How much communication will there be between the experimentalist, customer, and end-use analyst during both the planning and post-test stages? Does this create any strengths or weaknesses to the outcomes? • Describe how CompSim will be involved in the planning of this experiment?	
Sample / Geometric / Material Fidelity	• Is 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 intended use? How? • What documentation and general/specific understanding do you have of the pedigree? • Is 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 intended use? • What could be changed to improve the applicability?	
Experimental Verification	• What methods will be used to verify testing apparatus control/code is performing as desired? • How is the post-process of 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 (e.g. validation, calibration, materials characterization	• Describe how the test conditions will be characterized for the intended use? Will any conditions be missing, not well-characterized, or in doubt? • Describe how the output measurements will be characterized for the intended use? Will enough quantities of interests be measured, and will the right ones be measured? • For validation, will the validation metrics and criteria be 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 the intended use? • 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 will be reviewed by subject matter experts? Which elements will not, and of those, which may need further review and why? • Which of the above elements will be documented? Will the documentation serve the needs of the intended use, and help write the credibility evidence for the CompSim? Or is there anything missing that would have improved the validation process?	

Element 1: Planning

- Test Purpose
  - What is the overall goal for the experiment and simulation for all the stakeholders?
  - Who is the end user?
  - What is the intended use?
- How will analyst be involved in test planning?
- How much communication between
  - Experimentalist
  - Analyst
  - Customer

Assess 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**. 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 intended use?
Element	Prompts to Consider	Assessment Commentary
Planning	• Is purpose of the test known to the experimentalist and end-user (e.g. CompSim analyst)? What is the intended use of test? • How much communication will there be between the experimentalist, customer, and end-use analyst during both the planning and post-test stages? Does this create any strengths or weaknesses to the outcomes? • Will CompSim be involved in the planning of this experiment, and in what way?	
Sample / Geometric / Material Fidelity	• Is 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 intended use? • Do you know the pedigree? • Is 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 intended use? • 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 (e.g. validation, calibration, materials characterization	• Will the test conditions be characterized well enough for the intended use? Will any conditions be missing, not well-characterized, or in doubt? • Will the output measurements be characterized well enough for the intended use? Will enough quantities of interests be measured, and will the right ones be measured? • For validation, were 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 the intended use? • 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 will be reviewed by subject matter experts? Which elements will not, and of those, which may need further review and why? • Which of the above elements will be documented? Will the documentation serve the needs of the intended use, and help write the credibility evidence for the CompSim? Or is there anything missing that would have improved the validation process?	

Element 5: Intended Use

- Possible uses
  - Calibration
  - Validation
  - Materials characterization
- For the intended use
  - Describe the degree test conditions will be known
  - Describe how the measurements provide the required information
  - For validation, were metrics and acceptance criteria specified?
  - Could a range of conditions be helpful?

Assess 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**. 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 intended use?
Element	**Prompts to Consider**	**Assessment Commentary**
Planning	• Is purpose of the test known to the experimentalist and end-user (e.g. CompSim analyst)? What is the intended use of test? • How much communication will there be between the experimentalist, customer, and end-use analyst during both the planning and post-test stages? Does this create any strengths or weaknesses to the outcomes? • Will CompSim be involved in the planning of this experiment, and in what way?	
Sample / Geometric / Material Fidelity	• Is 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 intended use? • Do you know the pedigree? • Is 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 intended use? • 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 (e.g. validation, calibration, materials characterization	• Will the test conditions be characterized well enough for the intended use? Will any conditions be missing, not well-characterized, or in doubt? • Will the output measurements be characterized well enough for the intended use? Will enough quantities of interests be measured, and will the right ones be measured? • For validation, were 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 the intended use? • 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 will be reviewed by subject matter experts? Which elements will not, and of those, which may need further review and why? • Which of the above elements will be documented? Will the documentation serve the needs of the intended use, and help write the credibility evidence for the CompSim? Or is there anything missing that would have improved the validation process?	

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

- Sample, Geometric, Material Fidelity
  - How representative is the test article to the application?
  - Can you describe the test article pedigree and any pre-processing?
- Experimental and Environmental Fidelity
  - How relevant are the test conditions to the application?
  - Are improvements possible/needed to improve applicability?
  - Are all simulations inputs measured?

Assess 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**. 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 intended use?
Element	**Prompts to Consider**	**Assessment Commentary**
Planning	• Is purpose of the test known to the experimentalist and end-user (e.g. CompSim analyst)? What is the intended use of test? • How much communication will there be between the experimentalist, customer, and end-use analyst during both the planning and post-test stages? Does this create any strengths or weaknesses to the outcomes? • Will CompSim be involved in the planning of this experiment, and in what way?	
Sample / Geometric / Material Fidelity	• Is 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 intended use? • Do you know the pedigree? • Is 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 intended use? • 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 (e.g. validation, calibration, materials characterization	• Will the test conditions be characterized well enough for the intended use? Will any conditions be missing, not well-characterized, or in doubt? • Will the output measurements be characterized well enough for the intended use? Will enough quantities of interests be measured, and will the right ones be measured? • For validation, were 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 the intended use? • 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 will be reviewed by subject matter experts? Which elements will not, and of those, which may need further review and why? • Which of the above elements will be documented? Will the documentation serve the needs of the intended use, and help write the credibility evidence for the CompSim? Or is there anything missing that would have improved the validation process?	

Element 4: Experimental Verification

- How do you know you measured what you think you measured?
- Describe any testing of experimental control software
- How are test equipment calibration and quality implemented?
- Describe any testing of data post-processing codes
- Could instrumentation affect test conditions?
- How could repeatability be confirmed?

Assess 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**. 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 intended use?
Element	**Prompts to Consider**	**Assessment Commentary**
Planning	• Is purpose of the test known to the experimentalist and end-user (e.g. CompSim analyst)? What is the intended use of test? • How much communication will there be between the experimentalist, customer, and end-use analyst during both the planning and post-test stages? Does this create any strengths or weaknesses to the outcomes? • Will CompSim be involved in the planning of this experiment, and in what way?	
Sample / Geometric / Material Fidelity	• Is 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 intended use? • Do you know the pedigree? • Is 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 intended use? • 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 (e.g. validation, calibration, materials characterization	• Will the test conditions be characterized well enough for the intended use? Will any conditions be missing, not well-characterized, or in doubt? • Will the output measurements be characterized well enough for the intended use? Will enough quantities of interests be measured, and will the right ones be measured? • For validation, were 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 the intended use? • 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 will be reviewed by subject matter experts? Which elements will not, and of those, which may need further review and why? • Which of the above elements will be documented? Will the documentation serve the needs of the intended use, and help write the credibility evidence for the CompSim? Or is there anything missing that would have improved the validation process?	

Element 6: Experimental Uncertainty Quantification (UQ)

- To what degree will the uncertainty be quantified for:
  - Test conditions
  - Measurements of outputs
- What types of uncertainty measurements would be helpful for the UQ calculations/simulations?
- Can level of repeatability be quantified?

Assess 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**. 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 intended use?
Element	**Prompts to Consider**	**Assessment Commentary**
Planning	• Is purpose of the test known to the experimentalist and end-user (e.g. CompSim analyst)? What is the intended use of test? • How much communication will there be between the experimentalist, customer, and end-use analyst during both the planning and post-test stages? Does this create any strengths or weaknesses to the outcomes? • Will CompSim be involved in the planning of this experiment, and in what way?	
Sample / Geometric / Material Fidelity	• Is 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 intended use? • Do you know the pedigree? • Is 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 intended use? • 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 (e.g. validation, calibration, materials characterization	• Will the test conditions be characterized well enough for the intended use? Will any conditions be missing, not well-characterized, or in doubt? • Will the output measurements be characterized well enough for the intended use? Will enough quantities of interests be measured, and will the right ones be measured? • For validation, were 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 the intended use? • 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 will be reviewed by subject matter experts? Which elements will not, and of those, which may need further review and why? • Which of the above elements will be documented? Will the documentation serve the needs of the intended use, and help write the credibility evidence for the CompSim? Or is there anything missing that would have improved the validation process?	

"Assess Experimental Uncertainty" Tool

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).	

Element 6: Experimental Uncertainty Quantification

Pre-Test

Define Measurand(s) and Connections to QOI(s):

- How are they related?
- Do they require post-processing?

Define Measurement Process and Manage Uncertainties:

- How are the measurement and calibration methods described?
- Were expected uncertainties considered in test design?

Pre-Test *and* Post-Test

Expected and Estimated Uncertainties

- Is there an uncertainty inventory for sensors and expected values?
- Are the estimated uncertainties small enough to meet test requirements?

Uncertainty Propagation and Sensitivity Analysis

- What uncertainty sources are large compared to others?
- If total uncertainties are too large, can the largest sources be reduced?
- Could some uncertainties be better characterized?

Element 7: Peer Review and Documentation

- Assessed for each of the Elements of this tool.
- What additional peer review or documentation will be needed?
  - What level of rigor is expected in the documentation?

Peer Review

- Which of Elements 1-6 will be reviewed by subject matter experts?

Documents

- Will all of the Elements be documented?
- Are tabulated experimental results tied to their description and archived?

Assess 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**. 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 intended use?
Element	**Prompts to Consider**	**Assessment Commentary**
Planning	• Is purpose of the test known to the experimentalist and end-user (e.g. CompSim analyst)? What is the intended use of test? • How much communication will there be between the experimentalist, customer, and end-use analyst during both the planning and post-test stages? Does this create any strengths or weaknesses to the outcomes? • Will CompSim be involved in the planning of this experiment, and in what way?	
Sample / Geometric / Material Fidelity	• Is 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 intended use? • Do you know the pedigree? • Is 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 intended use? • 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 (e.g. validation, calibration, materials characterization	• Will the test conditions be characterized well enough for the intended use? Will any conditions be missing, not well-characterized, or in doubt? • Will the output measurements be characterized well enough for the intended use? Will enough quantities of interests be measured, and will the right ones be measured? • For validation, were 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 the intended use? • 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 will be reviewed by subject matter experts? Which elements will not, and of those, which may need further review and why? • Which of the above elements will be documented? Will the documentation serve the needs of the intended use, and help write the credibility evidence for the CompSim? Or is there anything missing that would have improved the validation process?	

Conclusions

- The Experimental Credibility Tool is a process to improve and communicate experimental results, their uncertainties, and documentation
- It is intended to be used by integrated teams of experimentalists, analysts, and customers
- It can accelerate development timelines by leveraging complimentary strengths of experiments and simulations