ITEA Journal 2008; 29: 294–300
Copyright © 2008 by the International Test and Evaluation Association

Tests and Experiments: Similarities and Differences

Richard A. Kass, Ph.D.
GaN Corporation,
U.S. Army Operational Test Command, Fort Hood, Texas

Test and experimentation are integral to the capability development process. This is the second of a two-part discussion on experimentation. This article considers the similarities and differences between experimentation and testing. While the two endeavors address different questions and exhibit some differences in the planning and execution process, overall similarities outweigh differences especially in event resources suggesting potential gains from sharing resources.

Key words: Demonstration; interdependence; shared resources; terminology; tests experiments; training.

Test and experimentation are two primary information venues in the Department of Defense (DoD) research and development. Testing is associated with system acquisition. Developmental and operational testing assess system progress toward acquisition milestones. Warfighting experiments[1] on the other hand, are associated with concept development. This is the second of two articles on experimentation. The previous article (Kass 2008) illustrated the uses, components, and validity requirements for warfighting experiments. Building on that description, this article discusses the similarities and differences between a test and an experiment.

This experiment versus test presentation[2] is intended to start the discussion. As experimentation and testing continue to evolve, the characteristics contrasted here will certainly change. The main thesis of this comparison is that while notable differences are evident, overall similarities are more significant than differences. Given the similarities, this article suggests that the resources employed in both endeavors can be shared to the mutual benefit of both. In the process of comparing and contrasting experimentation and testing, associated aspects of training are discussed. This will further illustrate the interconnectedness of DoD activities.

Terminology confusion between tests and experiments

Our language promotes confusion between tests and experiments.

We conduct experiments to test hypotheses.
We employ an experiment design to test systems.
Experimental systems undergo testing.

This confusion is exacerbated by common practices. Some test-like activities are renamed as "assessments" or "demonstrations" in order to reserve "testing" to specific agencies with identified acquisition requirements or to avoid consequences of negative results. Likewise, the "experiment" title can be attached to a number of activities that others would call "wargame" or "demonstration."

Terminology confusion suggests a close connection between test and experiment. The following definitions are provided:

Test: to assess the presence, quality, or genuineness of anything (Random House 1982);

Experiment: to explore the effects of manipulating a variable (Kass 2008).

Tests are one way to assess the quality of something. Other means include reliance on logical and mathematical relationships, authority, historical precedent, and natural observations. Assessments derived from testing imply empirical measurements under specified conditions. An example will illustrate the different but complimentary focus of experiment and test.

A math *test* is given to confirm whether students have attained certain levels of math proficiency using familiar letter-grade scale of A through F. A math *experiment* has a different purpose. Math experiments are designed to explore something new, for example, to determine the best way to teach math. The primary purpose of a math *experiment* is not to assess participants' level of math ability; but rather to examine the effect of various teaching methods on participants' math ability. During the *experiment*, each participant's math ability will be assessed by a math *test* to determine higher math ability

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **2008**	2. REPORT TYPE	3. DATES COVERED **00-00-2008 to 00-00-2008**

4. TITLE AND SUBTITLE **Tests and Experiments: Similarities and Differences**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Army Operational Test Command,91012 Station Avenue,Fort Hood,TX,91012 Stat**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**7**	

from lower. The purpose of this *test* is not to pass or fail the participants; but to quantify the effect of the *experiment* treatment. The *experiment* hypothesis might be: If teaching methods (a) are used, then math scores and (b) will increase. The way to determine whether math scores increased is to give the students a math *test* before and after the treatment.

The example signifies that testing is a method for assessing trial outcome. An experiment can be viewed as a sequence of tests. Each experiment trial is a test of one experimental treatment condition. An experiment is a systematic sequence of individual tests to examine a causal relationship, while a test is conducted to quantify an attribute.

Misperceptions of test and experiment distinctions

Given the interconnectivity of experiment and test, it is inevitable that misperceptions arise. One often hears experimenters caution their visitors: "Remember, this is an experiment, not a test." Why this admonition? Acquisition systems that do poorly in tests are in jeopardy of being cancelled. Tests include the idea of pass or fail. Experiments do not. Failure to produce a hypothesized experimental effect is more forgiving: "Let's try this and see what happens."

Experimenters sometimes push the forgiving nature of experiments too far in the phrase—"Tests can fail, experiments never fail." If this statement is interpreted to indicate that experiments rarely impact system acquisition decisions, the statement is understandable. If however, the statement is interpreted to mean, "there are no useless experiments" the statement is wrong. As discussed in the previous article, experiments can fail to provide sufficient information to resolve the experiment hypothesis.

Another misperception is that "Experimenting is messy, but testing is precise." This perception may reflect difficulties in representing the complexity of warfighting in experiments. It is difficult to conduct precise experiments in the operational environment. However, it is equally difficult to conduct precise operational tests in a realistic representative environment for the same reasons. This then cannot be the basis for distinguishing warfighting experiments from operational tests. Both depend on the expertise and experience of the experimenter and tester to balance the requirement for realistic operations against the needs to detect a change and to understand why the change occurred.

A third misperception is that "testing requires detailed data, while experiments use only high-level data." This distinction would not apply to warfighting experiments conducted in constructive or human-in-the-loop simulations because simulation outputs in these experiments are very precise and the experimenter is often inundated with detailed second-by-second interaction data on every entity in the simulation.

*Table 1. Four different perspectives of hypothesis elements**

Hypothesis: If capability A (new sensor), then effect B (increased detections).	
Demonstration	Show how A works to produce B
Training	Practice using A to produce B
Experiment	Determine better way to produce B
Test	Determine if A works to produce B

* Adapted from Figure 42 in Kass, R. A. 2006. *The Logic of Warfighting Experiments.* Published in 2006 by the Command and Control Research Program (CCRP) of the ASD/NII. Used with the permission of the CCRP.

This "data" distinction is derived from the circumstances in which tests and experiments are conducted in the field environment. Test agencies have accumulated sophisticated data-collection instrumentation for use in field tests. When the acquisition community needs to make a decision on a multibillion dollar program, it can justify the development of sophisticated instrumentation to provide maximum information to the acquisition decision. Conversely, experiments designed to examine the potential of a new technology do not have the same incentive today to invest large resources in a detailed answer.[3]

Conceptual difference between tests and experiments

The difference between an experiment and test cannot be based on precision or level of data alone. So is there a difference? One way to formulate an answer is to compare how various disciplines approach a new capability exemplified in the experiment hypothesis paradigm:

If capability A (new sensor), then effect B (increased detections).

Table 1 depicts how the elements of this hypothesis are viewed from the perspective of a demonstration, training, experiment, and test.

A ***demonstration*** is an event orchestrated to show how a process or product works. Demonstrations exhibit how a capability can produce an effect. In the military arena, demonstrations are commonly used as the initial step in training. An instructor demonstrates the correct procedures to follow with A to produce B. In the commercial world, product demonstrations are useful to convince others to buy the product or to illustrate the correct way to use it. While tests and experiments examine the effectiveness of capabilities, demonstrations assume the product works.

Training can be characterized as practice with A in order to accomplish B. This is easy to see when B is defined as a task with conditions and standards (more on this later). If the general task is to detect targets, the task conditions would specify the environment in which detections need to occur. The task standard would indicate the percent of the targets to be detected to meet the training objective.

Most ***experiments*** begin with a "capability gap." In our example, detections need to be increased as indicated on the right-hand side of the hypothesis. An experiment then is a trial-and-error process in search of a good solution for the left side of the hypothesis paradigm to fill the capability gap. Typical experiment questions are expressed in broad terms such as—"Does this approach produce a favorable outcome?" and "Can this problem be solved with X?"

In contrast, ***tests*** can be viewed as examining the goodness of a particular solution with respect to producing its intended effect. Tests are not searches for solutions, but rather a search for the strength of a solution's relationship to its effect. Typical questions for testing are expressed as, "How well does this item work?" and similarly, "How well does this item meet its requirements."

Thus far we have discussed some useful, and some not so useful, ways to think about the differences between experiments and tests. The remainder of this article will address the practical similarities and differences when it comes to planning, executing, and resourcing each event.[4]

Planning process

Planning coordination

Planning processes for experiments and tests employ different terminology but are quite similar functionally. Large tests are collaboratively designed and resourced using test and evaluation working-level integrated product teams (T&E WIPTs). This group meets periodically and is chaired by the capability Program Manager or operational test agency (OTA) depending on whether it is planning a developmental or operational test. Subgroups devoted to M&S, scenario, instrumentation, training, and so forth meet more frequently and less formally. A series of test readiness reviews (TRRs) brings together senior stakeholders to assess progress in the development of the system-under-test (SUT), test planning, and test-resource commitments.

Similar planning processes occur for major experiments. A concept development conference is followed by three planning conferences—initial planning conference (IPC), mid planning conference (MPC), and final planning conference (FPC). These serve the same purpose that T&E WIPTs and TRRs serve in testing. Again, smaller experiment planning IPTs can be formed to focus on M&S, scenario, analysis and data collection, training, and initiative development. An "initiative development" IPT is the experiment corollary to the capability Program Manager. Often capability initiatives for experimentation begin as "good-ideas" that need to be fleshed out so a concrete instantiation can be brought to the experiment. Capability instantiation can include adjustments to simulation, creation of new procedures, early prototypes when available, or implementation of low-fidelity surrogates when prototypes are not available.

Event rigor

The previous article identified four validity requirements for rigorous experiments:

1. Ability to employ the new capability;
2. Ability to detect a change;
3. Ability to isolate the reason for change;
4. Ability to relate results to real operations.

These experimentation requirements are applicable to testing. If the test unit is not able to employ the new system, or if the tester cannot detect a change in performance when the new system is employed, or cannot isolate the reason for any observed performance, or cannot relate the test environment and test results to actual operations; then the test has validity deficiencies. Experimenters and testers consult the same "design of experiment" textbooks to design their events.

While tests and experiments have similar validity requirements, they have different review processes. Test agencies and ranges conducting developmental testing have detailed test protocols and test plans that have increased in rigor through refinement over many years. Deviations from these protocols often require prior approval from both the tester and program manager. Operational testing includes an external review. Operational test plans of major acquisition systems are formally reviewed by the Director, Operational Test and Evaluation (DOT&E) and operational testing cannot begin until DOT&E approves the test plan.

Experiment agencies typically do not have the historical heritage found in the major test ranges and do not have detailed experiment protocols. Experiment plans are often reviewed internally and these reviews tend to focus on scenario realism, adequacy of the experiment initiative instantiation, and availability of experiment resources.

Results utilization

Testing has a major advantage over experimenting in results utilization. The results of developmental test (DT) or operational test (OT) support decisions about capability development programs. Test results assist

*Table 2. Three different portrayals of measures and goals**

	Measure	Goal
Training standard	Minutes to complete attack after target identification	Criterion (provided by commander)
Test criterion	Time to complete task after target identification (MOE/MOP)	Threshold (X minutes)
Experiment measure	Time to complete task after target identification (MOE/MOP)	(not usually available)

* Adapted from Figure 43 in Kass, R. A. 2006. *The Logic of Warfighting Experiments.* Published in 2006 by the Command and Control Research Program (CCRP) of the ASD/NII. Used with the permission of the CCRP.

program managers in assessing whether system performance is on schedule and where to focus system corrections.

In contrast, it not so easy for experimentation programs to show examples where their experiment results have changed the military environment. One reason for this is that most experiments are conducted on future prototypes or concepts outside the programmed acquisition realm. Any good ideas from experiments are initially unfunded and will struggle to find a "funded home."

Interestingly, the impact of many experiment programs may be more indirect than direct. One of the most visible legacies of the Millennium Challenge Field Experiment conducted in 2002 (MC02) was the follow-on creation of the Joint National Training Capability (JNTC) in Joint Forces Command (JFCOM). JNTC was built on the distributed live, virtual, and constructive (LVC) simulation architecture designed for experiment execution. While not the focus of the experiment, most experiment agencies can point to technologies developed to support their experiments that have found use (reuse!) in the operational forces as enhancements to the training environment or operations themselves.

Execution process

Type event

Experiments have an advantage over tests in flexibility—design space—to explore new ideas. Acquisition tests are restricted to testing something concrete—in hand, a component or prototype—even if it is only software algorithms. Experiments, in contrast, have few reality constraints. Experiments can be conducted on future weapons that exist only as concepts. These experiments can be executed entirely in simulation as constructive experiments or as analytic wargames. The focus of these experiments is not "does it work;" but on the potential impact of these ideas on future warfighting operations.

Unit tasks and measures

Tests and experiments are both concerned with realistic scenarios based on defense planning scenarios (DPS). Both look to the Joint and Service description of strategic, operational, and tactical tasks with their associated conditions and standards to provide the basis for unit activity during the event trial. Joint tasks and standards are identified in the Universal Joint Task List (CJCSM 2002)[5] (UJTL). Test and experiment use of the standardized tasks, conditions, and standards originally developed by the training community has been a positive development. The training, testing, and experimentation community can now speak a common language.

The UJTL conditions can provide the basis for the test or experiment trial conditions and UJTL standards can provide a starting point for developing the measures of effectiveness (MOE) and measures of performance (MOP) for tests and experiments. A closer look at the terminology for standards and measures will show that differences in terminology might blur similarities across the three communities. The UJTL task "Provide firepower in support of operations" includes the standard provided in the first row in *Table 2*.

The UJTL notes that training standards have two parts: a *measure* and *criterion*. While numerous quantifiable *measures* are provided in the UJTL, the *criterion* component is not included. The UJTL document asserts that the criterion, the specific time (in this example) in which the task is to be completed, is to be provided by the commander of the unit undergoing training. The commander might select 6 minutes as the task criterion and the unit would continue to re-execute the task until they accomplish it within the allotted time. It is a common misperception that the UJTL includes training standards—it only includes the measure portion of the standard. Training measures without criteria are still quite useful to testers and experimenters.

Starting points for measuring success in the test community are requirements identified in the initial capability document (ICD) and deployment capability document (DCD). While translating acquisition requirements directly into test criteria can be challenging,[6] some are relatively straightforward. Requirements for mean-time-between-failures, message completion rates, and detection ranges would be associated with measures like *time-between-system-failures*, *percent of messages successfully completed*, and *range of detection* with associated *"thresholds."* Notice the shift in terminology between the test and training community. For trainers, the "criterion" represents only the threshold component of the training standard. For the testers, "criterion" includes the measure and the threshold.

Testers and experimenters use identical terms for measures. The primary difference is that experimenters avoid the term "criterion" because they rarely have available thresholds to evaluate success. Consequently, experiments rely on comparative analysis based on alternate treatment conditions—different proposed capabilities or a single capability under different scenario conditions.

Table 2 above highlights the ease in bridging terminology differences. Use of common measures among trainer, tester, and experimenter would increase mutual synergies among the three communities. Operational forces continuously undergo training that could yield realistic, operational thresholds of baseline-force capability based on a heterogeneous mixture of units under a wide variety of operational conditions. This training data, if systematically collected and used by experimenters, would greatly enhance the relevancy of experimentation in answering "so what" questions. Even if an experimental capability performs better under some conditions than others, how much better it is than what is available today?

The test community could also benefit when quantifiable thresholds on current mission performance are available from the training community. Test criteria are based on *system* performance rather than *unit* mission accomplishment. In operational testing, there is increasing emphasis on assessing system capabilities and limitations with respect to overall unit mission accomplishment—especially on the Joint battlefield.[7] While *system performance thresholds* are readily available from requirement documents, there is yet no agreement on how to arrive at a *unit mission success threshold*. A systematic data collection effort of unit mission successes during training exercises might provide the operational baseline for testing system contribution to mission success.

Event resources

If you fell into the middle of a warfighting field experiment, operational test, or training exercise it would be difficult to know which one you had fallen into. In any one, you would observe military operators performing tasks to accomplish a mission. In the extreme case, one might detect operators employing novel procedures or equipment. This could indicate an experiment or test on advanced technology. With this exception, almost nothing else during actual execution from the player perspective would indicate experiment, test, or training. It is only when the purpose of the event is known as discussed above that subtle differences between experiment, test, and training may be evident.

*Table 3. Comparison of resource requirements**

Requirement	Experiment	Test	Training
System Realism			
Simulated (*constructive*)	X	x	X
Simulator (*virtual*)	X	x	X
Prototype (*live*)	X	XX	-
Operational system (*live*)	X	x	X
Trained operators	X	X	x
Instrumentation			
System-level diagnostic collection	-	XX	-
System-range interactions collection	X	X	X
Feedback	X	X	XX
Networks/communications	X	X	X
Exercise Management			
Controllers	X	X	X
Observers	X	X	XX
Trainers	x	x	X
OPFOR unit/equipment	X	XX	XX
Analysts	XX	XX	-

* Adapted from Figure 44 in Kass, R. A. 2006. *The Logic of Warfighting Experiments.* published in 2006 by the Command and Control Research Program (CCRP) of the ASD/NII. Used with the permission of the CCRP.
"x," "X", and "XX" indicate increasing emphasis.

Given this similarity in execution, it is not surprising that the resources to execute an experiment, test, or training exercise are quite similar with only a few notable exceptions. *Table 3* provides a comparison of resource requirements.

System realism. Earlier it was noted that experimentation is the most flexible enterprise. It can be executed with live, virtual, or constructive systems as the primary system of interest. Testing has the most stringent system requirements. Developmental and operational testing are conducted on live prototypes and preproduction systems. While testing does employ some constructive and virtual representations, these are used to save resources in populating a realistic test environment, not to represent the primary SUT. While most operator training is conducted on operational systems, use of air, ground, and sea virtual simulators is continuing to expand to save training resources. Most operational staff training is also accomplished in an LVC environment.

Instrumentation. Range instrumentation and communication networks for event control, execution, distribution, data collection, and player feedback is becoming more similar for test and training as a result of several ongoing initiatives to further common support architectures.[8] The primary exception to eventual complete commonality is that testing requires diagnostic performance data from the SUT, as discussed earlier.

Exercise management. The lead-in to this section suggested that event execution was quite similar among the three communities; this is also true for event management resources. All three communities require similar event controllers, event observers or data collectors, and analysts for interpreting results. Similarly all three require a representative threat force. Operational test approval often requires that threat representativeness be certified by an external agency. While the training community emphasizes training, test and experiment communities employ trainers to ensure operators can use the new capability. The table highlights that experiments and tests place a higher premium on statistical analysis of the event data.

Summary

Experimentation and testing are both important to capability development. They provide empirical data for different questions. However, they have more similarities than differences. They have similar planning processes; similar validity requirements; use similar language for tasks and measures; and for the most part, employ the same resources to design, execute, and report events.

There are some differences to keep in mind. Experiments have greater flexibility to explore a wider variety of warfighting questions and alternatives using virtual and constructive simulations, since experiments do not have to wait for actual prototypes. Conversely, experimentation has far less formal methodological oversight and it is not always easy to link experiment outcomes to implemented operational changes. Tests, on the other hand, have more explicit measures of success (e.g., system procurement requirements) and provide far greater system-diagnostics data collection to know what to fix. Moreover, test results always impact capability development.

Confluence of test and training architecture is assisted by the fact that they have clearly delineated sponsors in the DoD—Under Secretary of Defense (Personnel and Readiness) manages training while the Under Secretary of Defense (Acquisition, Technology, and Logistics) and Director, Operational Test and Evaluation provide test management. There is no corresponding high-level sponsor and policies for experimentation. Consequently, experimentation policy is decentralized making it more difficult to build a coalition with the test and training communities from the top down. However, the sharing of expertise, data, and resources can only benefit all three.

A realization from these comparisons is that predominantly the same resources can be used for experimentation and testing, as well as training. This suggests that efficiencies can be gained if experimentation, testing, and training continue to progress towards shared resources. Increased emphasis is being directed at finding interdependent investment strategies for overlapping infrastructure to support testing and training.[9] A similar interdependency case can be, and should be, made for testing and experimenting. ❑

RICK KASS has 25 years in designing, analyzing, and reporting on operational field tests and military experiments. He held multiple positions as test officer, analyst, and test director for 18 years with the U.S. Army Test and Evaluation Command (USATEC) and was chief of analysis for 7 years with the U.S. Joint Forces Command (USJFCOM) joint experimentation program. Currently Rick works for GaN Corporation supporting the Army's Operational Test Command at Fort Hood, Texas. He has authored over twenty-five journal articles on methods for research, experimentation, and testing and was the primary architect establishing the permanent Warfighting Experimentation Working Group in the Military Operations Research Society (MORS). Rick is a graduate of the National War College and holds a Ph.D. in psychology from Southern Illinois University. E–mail: rick.kass@us.army.mil

Endnotes

1Warfighting experiments are distinguished from experiments used in medical research and early technology research.

2This article expands on ideas previously printed in Kass R.A., 2006. The Logic of Warfighting Experiments, published by the Command and Control Research Program (CCRP) of the ASD/NII, which has graciously granted permission to include the material in this work.

3In the 1970s and 1980s the U.S. Army sustained a Combat Development Experimentation Center (CDEC) with dedicated operational forces, advanced range instrumentation, and scientific methodology for experiments. This center no longer exists. There are costly experiments today. These costs are mostly associated with force operations and M&S development and execution; not the cost of collecting detailed system performance data.

4The following comparison of test and experimentation is more applicable to developmental and operational testing following early prototype development that are assessed against military tasks as opposed to engineering thresholds.

5CJCSM 3500.04C, July 2002. The Services have augmented the Joint list with their respective tactical tasks: Army Universal Tactical List (AUTL), Universal Navy Task List (UNTL), and Air Force Task List (AFTL).

6See Kass, R. A. "Writing measures of performance to get the right data." *The ITEA Journal of Test and Evaluation*, June/July 1995, vol. 16

(2)) for a discussion of pitfalls when translating requirement statements into performance measures for test plans.

[7]In 2007 DOTE chartered a JT&E to develop methodology to conduct and assess system performance within a system-of-system approach to accomplishing military missions in a Joint environment. This JT&E is called Joint Test and Evaluation Methodology (JTEM).

[8]Some examples are the Central Test and Evaluation Investment Program (CTEIP) Common Range Integrated Instrumentation System (CRIIS), Army's One-Tactical Engagement Simulation System (ONE-TESS), the Joint Mission Environment Test Capability (JMETC), and Test and Training ENablingArchitecture (TENA).

[9]Joint memorandum "Test and Training Interdependency Initiative" September 7, 2006 signed by the Under Secretary of Defense (Acquisition, Technology, and Logistics), the Under Secretary of Defense (Personnel and Readiness) and the Director, Operational Test and Evaluation provided a common vision for interdependent test and training solutions to achieve a single, more realistic operational environment. See the Test Resource Management Center *FY2007 Annual Report* (January 2008) for implications of this memorandum.

References

Kass, R. A. 1995. "Writing measures of performance to get the right data." *The ITEA Journal of Test and Evaluation*. June/July 1995, Volume 16-2, pp. 2–6.

Kass, R. A. 2006. *The Logic of Warfighting Experiments*. Washington, D.C.: Command and Control Research Program (CCRP) of the Assistant Secretary of Defense and Information Integration (ASD/NII), Available from the CCRP website at http://www.dodccrp.org.

Kass, R. A. 2008. "Framework for understanding experiments." *The ITEA Journal of Test and Evaluation*. Volume 29-2, pp. 167–174.

Random House. 1984. Random House College Dictionary New York, NY: Random House, Inc.

The ITEA White Sands Chapter introduces...

LIVE-VIRTUAL-CONSTRUCTIVE CONFERENCE

Formerly known as the ITEA M&S Conference

JANUARY 12 - 15, 2009

El Paso, Texas

*Our conference committee, chaired by **Mr. Thomas Berard**, White Sands Missile Range Executive Director, has decided to take the highly successful ITEA Modeling and Simulation Conference that has been held in the month of December in Las Cruces, New Mexico for the past thirteen years and move it to El Paso, Texas in January. This conference will still have the traditional components of the M&S conference with new topics for you to consider. We hope that you continue to support this event by providing outstanding abstracts, exhibiting your products and services, and contributing your sponsorship dollars for the scholarship program.*

CONFERENCE PLANNING COMMITTEE

PROGRAM CHAIR
Gilbert Harding, Director of Systems Engineering Directorate

TECHNICAL CO-CHAIRS
Doug Messer • 575-430-2951 or 1825 • dmesser99@aol.com
Hank Newton • 575-385-7270 • hnewton@aoc.nrao.edu

TOPICS

- Distributive Live-Virtual-Constructive and T&E
- High Performance Computing in Test and Experimentation
- Civil-Military Operations
- Collaborative Simulation & Testing
- Testing in a Networked Environment
- Requirements for M&S in Test & Evaluation
- M&S Tools
- M&S Technologies
- Autonomous and Cognitive Systems
- Netcentric Testing

CALL FOR PAPERS

All sessions will be at The Wyndham Hotel and will be unclassified. Abstracts will be reviewed for presentation during a conference session or as a poster paper. Papers selected for presentation will be published in proceedings on the World Wide Web. Visit **www.itea-wsmr.org** for previous Modeling & Simulation proceedings.

EXHIBITS

Your company or government organization will want to take advantage of the premium and yet limited space that is available for you to display and demonstrate products and services for the test and evaluation community. To obtain an application to exhibit or to see the floor plan, visit the ITEA website.

SPONSORSHIP

Fours levels of sponsorship are available for your company to participate in: Platinum $2500, Gold $1000, Silver $500 and Bronze $250. Your sponsorship dollars will defray the cost of this event and support the ITEA scholarship fund, which assists deserving students in their pursuit of academic disciplines related to the test and evaluation profession. For more information on the benefits of sponsorship, or to obtain a pledge form, please visit www.itea.org.

For the latest information, including an updated agenda, applications to exhibit and sponsor, the floor plan, lodging information, register on line, and much more visit

www.itea.org