stepping up 2012

CRESST: Advancing assessment, evaluation and technology worldwide

National Center for Research on Evaluation, Standards, & Student Testing

UCLA | Graduate School of Education & Information Studies
Founded in 1966, the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) is a leading research organization contributing to improved learning for American children and adults. As part of UCLA’s Graduate School of Education & Information Studies, CRESST conducts rigorous research studies, develops innovative assessment and learning systems, forges new methodologies, as well as evaluates education and training programs. Throughout all of its work, the center is committed to leveraging the power of new technologies in order to advance assessment and evaluation practice.

Over the past few years, CRESST has grown to meet expanded needs by substantially strengthening our work beyond a historical K–12 audience. Today, our projects include government, military, medical, and pre-K through adult-level organizations. We have successfully scaled up our models-based assessment research to create innovative assessment, learning, and evaluation systems for almost every age and program type. Designated as the national center for research in its mission area, CRESST’s most recent nationally competitive achievement is the five-year, $10 million award for the National Center for Advanced Technology in Schools (CATS). Through CATS, CRESST explores new game designs for assessment and learning purposes. CRESST is proud to have been designated as a Center of Excellence by the US Congress as well as the recipient of numerous military commendations for its simulation-based assessments.
designs in their efforts to develop, assess, and provide feedback on desired capacities.

In the *College Readiness Assessment Initiative*, CRESST researchers are developing and validating designs for integrated literacy assessments, another example of applied CRESST models. Funded by the Bill and Melinda Gates Foundation, the assessments will measure students’ progress in the Common Core Standards for English Language Arts and student understanding of important subject matter concepts and principles in history, science, and literature. The performance measures will help teachers develop, assess, and provide feedback on secondary students’ progress toward college readiness. CRESST is also investigating their applicability for teacher evaluation and accountability purposes.

In the *Advanced Technologies for Automated Performance Assessment* (ATAPA) project, CRESST is implementing a design for simulation-based assessments of situational awareness, decision-making and communication in the context of naval damage control. Funded by the Office of Naval Research and based on a 3D, computer-based simulation depicting the interior of a naval ship, players assume the role of a damage control investigator who is tasked with identifying, addressing, and reporting on a variety of potential, imminent, and existing fires and fire...
hazards. Derived from the CRESST models, center researchers are using Dynamic Bayesian networks to monitor and evaluate in real time all actions and decisions related to situation awareness, communications, and decision-making plus providing both formative and summative feedback on performance.

2. EVALUATION METHODOLOGY FOR RIGOR AND RELEVANCE

Applying our expertise in measurement, research design, and advanced statistical modeling—CRESST is pushing forward the state of the art in evaluation methodology to enhance goals for both rigor and relevance. Current work includes new measures of learning processes; fidelity of implementation and program quality; new approaches for detecting complex performance outcomes; and the synthesis of existing research to inform the development, evaluation, and improvement of Web-based courses, training simulations, and educational games.

CRESST researchers, for example, are capitalizing on the data tracking power of interactive media to devise new, real-time, unobtrusive measures of program process and performance. The challenge in technology-based simulation and games lies not in capturing the raw data but in interpreting what a player’s actions and decisions mean in the broader context of learning. We have overcome that challenge by using keystroke, response latencies, and eye tracking data, in combination with data mining and cluster analysis techniques, to infer fidelity of simulation and video game use and to identify alternative pathways, error patterns, and misconceptions that can inform program improvement in both education and training settings.

In more traditional programs, CRESST researchers have advanced the use of teacher logs and analyses of teacher assignments and student work to create reliable indicators of fidelity implementation and quality of teacher practice. The methodologies have been applied in evaluations of pre-service training, professional development, curriculum innovation, and school reform. Currently, CRESST researchers are using the methodology to examine the effects of IMPACT, a teacher residency program for pre-service teachers in math and science and in studies of teachers’ integration of literacy and content teaching.

CRESST researchers also have been active in advancing new statistical modeling methods to provide more sensitive analysis of research and evaluation data. CRESST explorations of game applications, for instance, have introduced survival analysis as a novel technique to create alternate indicators of learning progress and game success. With more traditional data, CRESST researchers have extended latent variable regression hierarchical modeling to four levels, thereby monitoring school performance over time (using multi-site, multiple cohorts longitudinal data) and deriving estimates of initial status, growth rates, and performance gaps between identified subgroups. We are applying similar methodology to examine the impact of teacher preparation programs on subsequent student growth and/or the effects of multi-site training programs on longitudinal outcomes.
3. MULTIMETHOD EVALUATIONS OF PROGRAM QUALITY

With a focus on theory-based and rigorous design, innovative measurement, and longitudinal analysis, CRESST is expanding its tradition of exceptional K–12 evaluation to new fields. Current projects combine qualitative and quantitative methodologies with multiple measures to evaluate the implementation and effects of military training, medical education, and K-12 school programs. Both high technology and more traditional contexts are subjects of our evaluation inquiry.

In the Evaluation of Simulations for Assessing Medical Education and Training, for instance, CRESST researchers are combining their expertise in learning and cognition with that in measurement and evaluation to generate a framework for analyzing the quality and effectiveness of medical training simulations used throughout the military services. We are also creating replicable designs, common indicators, and exemplary measures for evaluating and improving the impact of these programs. Similarly, CRESST is applying this knowledge to the Sonosimulator Efficacy Study, investigating the value of specific simulations for medical training purposes.

In California, CRESST is conducting a statewide multimethod evaluation of afterschool programs, an area of longstanding interest. The four-year study includes approximately 4,000 elementary and middle school sites and 190 high school sites. We are using archival, survey, and case study data to examine program implementation and effects. Among the prime evaluation questions are program effects on school attendance, homework completion, graduation rates, student achievement, and the relationship between afterschool dosage and outcomes.

CRESST multimethod evaluations encompass programs of varying size and duration. Current projects range from small studies of the arts, to technology-based school and learning innovations, to longitudinal studies of teacher and training effectiveness. Customized to program theories of action and anticipated outcomes, CRESST evaluations employ program sensitive measures that provide trustworthy results in easy to understand formats.

4. ASSESSMENT USE FOR BETTER TEACHING, LEARNING, AND DECISION MAKING

CRESST research has long investigated the design, validation and use of assessment to inform policy and improve practice. In the vanguard of studies examining the impact of testing on schools and teachers, CRESST research and development also has sought to improve the validity and use of data to improve school planning and decision-making. CRESST’s Quality School Portfolio (QSP), for example, provided an early, educator-friendly electronic tool to enable schools to merge, analyze, disaggregate,
report, and use longitudinal student data for school planning, program evaluation, and improvement. QSP classroom features also enabled teachers to monitor and respond to individual student progress. With today’s technology advances, CRESST is creating comprehensive system architectures for integrating assessment, instruction, and learning.

Innovative CRESST research also emphasizes teachers’ use of formative assessment, a strategy that has shown strong effects on student learning. CRESST efforts are examining what it will take to move from current reality to higher quality formative assessment practice. CRESST’s recently completed POWERSOURCE© study, for example, developed a new approach to the design of formative assessment tools in middle school mathematics and created professional development and instructional resources to support teachers’ use. Results from a randomized controlled study showed significant learning effects.

Looking from the design of assessment tools to optimal processes for assessment use, CRESST researchers are using qualitative and quantitative methods to identify and understand critical features of formative practice and their relationship to student learning. In the Efficacy Study of a Diagnostic Formative Assessment for Middle School Science, for example, CRESST researchers are examining relationships between teachers’ content knowledge, formative assessment practice, and student learning, as well as studying the impact of curriculum-embedded formative assessments. Through the Assessment and Accountability Comprehensive Center, CRESST is partnered with WestEd to assist 16 regional centers and all 50 states in developing state and district capacity for data use and to improve teachers’ formative assessment practices.

5. VALIDITY AND FAIRNESS FOR SPECIAL POPULATIONS

A long-term CRESST goal has been evaluation, research, and development that improves assessment validity while helping to narrow achievement gaps between diverse student groups including students with disabilities and English language learners (ELLs). CRESST researchers, for example, pioneered research on the validity of assessment accommodations, providing research exemplars for the field. We also identified important variables for increasing the accessibility of assessment and accountability systems to special population learners. CRESST research investigating the efficacy of language-support accommodation, for example, has produced a widely used protocol for reviewing and revising test items to reduce construct-irrelevant language load. Recommendations derived from this robust research program have been incorporated into federal and state policies and practices, as well as by commercial agencies and professional associations.
Currently, CRESST researchers are developing new assessment options to support English language development for ELLs. In our Interventions for Struggling Adolescent and Adult Readers and Writers project, CRESST researchers are working with ETS colleagues to develop a series of formative assessments, guidelines, and a Web-based reporting tool to help teachers diagnose and respond to the language learning needs of their ELL students. CRESST researchers also are active in the development of a new, federally funded measure of English language development.

CRESST evaluation work also seeks to close achievement gaps. Through studies of charter and magnet schools serving under-represented populations, CRESST researchers are identifying and analyzing the conditions and factors that lead to student success. Evaluations of magnet and charter schools across the country including Green Dot Public school’s Locke Transformation Project, exemplify our long-term commitment to closing achievement gaps.

6. TECHNOLOGICAL INNOVATION IN ASSESSMENT AND EVALUATION PRACTICE

From development of innovative computer learning and assessment simulations to sensor-based measurement methods, CRESST technology and data research provide useful tools, assessments, and reporting systems for diverse purposes and audiences. Researchers in our Center for Advanced Technology in Schools (CATS), for example, are merging psychology, instruction, and assessment research with game learning systems to improve middle school students’ mathematics achievement. As a key partner in GAMECHANGER: Using Technology to Improve Young Children’s STEM Learning, for example, CRESST is developing ground-breaking, Web-based computer games to both stimulate and assess young children’s science learning, while conducting inquiry into principles of effective game development and exploring new crowd sourcing techniques to gather feedback.

In further explorations of the power of emerging technology to create new options for assessing and promoting learning, CRESST researchers are working to create comprehensive, next generation assessment systems. Grounded in ontologies of learning progres-
and refinement of expected learning progressions. One component, for example, the CRESST Assessment Application (CAA) automates on-going monitoring and learning support in military training simulations. CAA provides a Bayes-net driven software system to monitor performance, provide real time feedback, and immediately adapt training sequences.

The Ontological Relations Builder (ORB), developed with funding provided by the Bill and Melinda Gates Foundation and Office of Naval Research, is designed to create, explore, store, compute, visualize, and communicate information. Using ontologies, the ORB stores information about items or entities as well as relations between each. Items or entities may range from something as simple as labels identifying objects or people, to knowledge, skills, attitudes, concepts, standards, and assessments. Relations may be taxonomic or may describe dependencies or system-calculated similarities. Using data-mining algorithms, ontology creation can be done manually or automatically. The ORB design facilitates flexible and consistent modeling of a domain; allowing, for example, the alignment of standards, instruction, and assessment, while supporting the generation of Bayesian nets used in assessment systems like the CAA.

In another prototype development effort funded by the Gates Foundation, CRESST is creating a personal Web teaching assistant (webTA) to help teachers transition to the new Common Core State Standards including cognitive demands and problem solving skills. WebTA will help teachers create lessons and assessments aligned with Common Core progressions, plan and schedule instruction and assessment sequences, monitor student performance, and know what to teach next.

CRESST has also pioneered the use of sensors in assessment systems, specifically to measure psychomotor skills for improved human performance. For instance, to more accurately evaluate military marksmanship skills, CRESST designed a sensor system that measures breathing, muzzle, and trigger movements. When combined with EEG data, the innovative assessment system provides enhanced feedback to instructors and trainees. CRESST is now investigating the application of these techniques to assess medical procedure skills.
CRESST employs more than 110 people at its UCLA location including approximately 20 researchers, plus 45 graduate and undergraduate students working on a broad variety of research projects. All CRESST projects are housed in modern offices on the UCLA campus, with access to over 16,000 square feet of space. The center’s extensive computer systems are linked to a 100/1000 megabit Ethernet network, connected to shared files, databases, and Web servers. Research and data are centrally collected and stored on a secure network in compliance with all human subjects protection requirements.

CRESST staff has access to virtual private networking (VPN) services and campus-wide wireless connectivity. Server-side filters protect the network from SPAM and other malicious activities. CRESST also houses an array of multimedia equipment including DV and HDV cameras, HD video projectors, and recording devices.

Through UCLA Academic Technology Services, CRESST has access to a visualization portal, technology sandbox, comprehensive statistical computing systems, central software, and information technology to support technology access and usability for students, faculty, and staff with disabilities in compliance with federal and state laws. UCLA’s visualization portal includes a 40-seat theater with up-to-date virtual reality technologies used for scientific visualizations and advanced digital technologies supporting both instruction and research.

CRESST has a state-of-the-art video conference system with dual-mode Internet connectivity and multiple location support. The system has been used to enhance graduate level courses and distance learning between UC campuses.

Multiple on-campus UCLA conference centers are capable of supporting more than 2000 conference attendees and as many as three simultaneous large or small conferences. UCLA’s libraries are among the top-ranked research libraries in the United States with over 7.6 million books and 90,000 magazines and academic journals.

Since 1994, the CRESST Web site has been broadly disseminating CRESST research information to the public. The Web site contains more than 800 full-text research reports dating back to 1966, plus hundreds of overhead presentations, newsletters, and assessment resources. Approximately 15,000 products are downloaded each month from the CRESST Web site and more than 180,000 each year. CRESST also actively participates in the Education Resources Information Center (ERIC), the world’s largest digital library of education literature, further expanding access and use of center research and evaluation reports.

CRESST leverages its budget across multiple projects and partners to increase its impact, ensuring that all funding agencies receive maximum value for their investment.
CRESST researchers are among the most widely recognized authorities in their respective fields, serving as members and leaders of national professional associations including the National Research Council’s Board on Testing and Assessment, the National Council on Measurement in Education, the American Psychological Association, the National Academy of Education, the American Educational Research Association (AERA), and the World Education Research Association (WERA). CRESST partners have invented and produced a substantial number of innovative tools including iStatistics and the Quality School Portfolio, both used to evaluate educational quality and improvement.

CRESST research associates have been honored for distinction in research and practice including: the AERA Award for Distinguished Research; the American Psychological Association Distinguished Scientific Award for Applications in Psychology; the American Psychological Society James McKeen Cattell Award; the Educational Testing Service Award for Distinguished Service to Measurement; multiple AERA Presidential Citation Awards; and the AERA Research-to-Practice Award. More than a dozen current or former CRESST researchers or National Advisory Board members have been selected to AERA Fellow status based on their “scientific or scholarly contributions” to educational research on a national or international basis.

CRESST Director Eva Baker is the first full-term president of the World Education Research Association and a former president of the American Educational Research Association (2006-2007). She is a former co-chair of the Joint Committee for the Revision of Standards for Educational and Psychological Testing and was a congressionally appointed member of the National Council on Education Standards and Testing and chair of the Board on Testing and Assessment, National Research Council.

CRESST Director Joan Herman has held a variety of leadership positions in the American Educational Research Association and chaired the National Research Council committee that produced the recently published book, Assessing 21st Century Skills. Dr. Herman is current editor of Educational Assessment, serves on the Joint Committee for the Revision of Standards for Educational and Psychological Testing, and is co-author of numerous practitioner focused books including The Turnaround Toolkit: Managing Rapid, Sustainable School Improvement with Lynn Winters.

CRESST researchers and partners are editors, co-editors or authors of a broad number of technology and education books including: Assessment of Problem Solving Using Simulations; Technology Applications in Education: A Learning View; and Web-based Learning: Theory, Research, And Practice. CRESST Assistant Director for Professional Development Margaret Heritage is author of Formative Assessment: Making It Happen in the Classroom.

CRESST researchers publish dozens of articles every year, primarily in distinguished peer-reviewed journals but also in more practice- or policy-focused publications such as Educational Leadership, Kappan Magazine, and the American School Board Journal, School Administrator. Serving as editors, co-editors, and reviewers of national journals, CRESST partners are regularly quoted for their expertise in America’s top newspapers including the New York Times, USA Today, The Washington Post, Los Angeles Times, Chicago Tribune, Education Week, and more than 100 other newspapers or magazines.

For additional information about specific CRESST projects or to find out if CRESST can help your organization, please contact Ron Dietel at dietel@cse.ucla.edu or 310-794-9168, or visit www.CRESST.org.
The National Center for Research on Evaluation, Standards, and Student Testing (CRESST) is a partnership of many universities across the nation. The CRESST mission focuses on the assessment of educational quality, addressing persistent problems in the design and use of assessment systems to serve multiple purposes. Together, the institutions comprising CRESST are committed to:

1. Developing assessments, tools, procedures, and systems to increase achievement in schools and adult training application;

2. Exerting intellectual leadership in the design, analyses, appropriate interpretation, and use of assessment systems for the research, practice, and policy communities—and the public;

3. Conducting high-quality research to create new knowledge and advance theories that will have long-term impact on conceptions and interpretations of educational quality;

4. Creating and evaluating assessment approaches and accommodations that promote equity in assessment and learning; and

5. Engaging teachers, researchers, policy makers, and the public in reflection and action to improve assessment and its link to educational quality through direct connections with CRESST products and staff.