

Tuesday Conference Program

► **Breakfast with Exhibitors, Partners & Poster Sessions, 7:30 AM – 9:00 AM**

Location: Sun Ballroom C - D and Sun Lobby

► **2006 ATP Conference Opening Keynote Address, 9:00 AM – 10:15 AM**

Location: Sun Ballroom A - B

Opening Keynote Address: Dave Wilson, President and CEO, Graduate Management Admission Council

"Because It Matters!"

Join us as Dr. David Wilson opens the 2006 ATP Conference! Don't miss this informative keynote address as Dr. Wilson discusses the importance of measurement as a management tool in business and performance management. The role of professional associations such as the Association of Test Publishers in promoting sound measurement and testing practices will be emphasized.

Dr. Wilson received his BCom degree from Queen's University, MBA from the University of California–Berkeley, and Ph.D. from the University of Illinois. He has published extensively, including four text books and numerous articles and cases. He is a frequent speaker to business, professional, and academic groups and has received numerous awards.



► **Breakout Sessions, 10:30 AM - 12:00 PM**

Alternative Technologies for Performance Testing

10:30 AM - 12:00 PM

Location: Tallahassee 2

Type: Cross Divisional

Level: All

This session discusses and demonstrates alternative technologies for the development, delivery and scoring of performance tests. These technologies include simulations, live applications, and virtual laboratory environments. The presenters will discuss the unique features and benefits of their technology for assessing human performance capabilities.

Presenters: Dean Hardy, Hatsize Learning Corporation
Doug Rosenthal, President, Job Performance Systems
Marcus MacNeill, Surgient

Tuesday Conference Program

Developing Innovative Assessments Across a Continuum - Discrete Items, Situated Tasks, and Simulated Environments

10:30 AM - 12:00 PM

Type: Cross Divisional

Location: Miami 2

Level: Intermediate

A host of measurement issues arise when test developers undertake innovative assessment. We address these in terms of assessment structure, veridicality, and cross-dimensional components. Assessment structure defines the way in which the assessment is presented to the examinee, the amount and type of information available to the examinee, and the decisions and actions required of the examinee. It extends from items, through tasks, to environments. The second level of the taxonomy, veridicality, describes the level at which the innovative assessment provides a realistic context for measurement through the levels of complexity, interactivity, and fidelity. The third taxonomy level is comprised of cross-dimensional components that contribute to innovation: media inclusion, response action, and scoring methods. These categorizations of the latest CBT innovations will be illustrated with examples from operational testing and research. Practical issues that test developers or psychometricians may need to address in developing assessment innovations will be integrated throughout.

Presenters: Cynthia Parshall, Measurement Consultant
Christine Harmes, Testing Consultant
Steven Wise, Professor, James Madison University

Rapid Growth Assessment Programs Utilizing Multi-Distribution Channels

10:30 AM - 12:00 PM

Type: Cross Divisional

Location: Daytona 1 & 2

Level: All

This is a cross-divisional session addressing the broad spectrum of private site assessment as seen through the eyes of four dynamic yet very different organizations. The issue that has confronted these organizations is remarkable growth and rapid program evolution. The key to their success has been their ability to manage high-volume workforce assessment. The speakers will discuss highlights of their strategy including: • Private site testing • Scalability to thousands of testing center locations • Speed to market • Cost • Convenience

Presenters: Geraldine Kenney-Wallace, Director Group e-Strategy, Learning & Information Technology, City & Guilds, UK
Suzana Lopes, Director, Business Development, Europe, Middle East and Africa, Promissor
Seth Zimmer, Director-Capability Assessment, BellSouth Corporation
Dale Cyr, Chief Executive Officer, ARDMS
Ian Necus, Director, UK Operations, Promissor

Providing "Business Intelligence" to Test Sponsors, Test Users, and Candidates

10:30 AM - 12:00 PM

Type: Cert/Lic

Location: Miami 1

Level: Advanced

* Bring your laptop to this session with wireless Internet if available

Test sponsors, test users and candidates all have increasing needs for information about their testing programs. This session investigates the use of state-of-the-art business intelligence tools to provide reports and data, manage eligibility and other information for the test sponsors and their affiliates. The session will illustrate how this information can be used to help manage a testing program more effectively. Two testing programs, GMAC[®] and Hewlett-Packard, will provide demonstrations and examples of these systems.

Presenters: Ashok Sarathy, Associate Director, GMAT Program, GMAC
Peg Jobst, Vice President, GMAT Program, GMAC
Tyler Tu, Vendor Manager, Hewlett Packard
Donna Figueroa, CPT - Manager, Exam Design and Management, Hewlett Packard
Stephen Greenfield, Product Manager, Pearson VUE

Tuesday Conference Program

Technology-based Performance Assessment: Value, Innovation, and Measurement

10:30 AM - 12:00 PM

Type: Cert/Lic

Location: Miami 3

Level: Advanced

This session will focus on four different aspects of technology-based performance assessments. Examples of different performance-based item types will be demonstrated, the use and integration of various media will be discussed, and the results of a case study comparing multiple-choice and performance-based items will be presented.

Presenters: Elizabeth Sabin, American Veterinary Medical Association
Yvonne Prefontaine, Solaris Certification Program Manager, Sun Microsystems
Wallace Judd, Executive Director, Performance Testing Council

Protecting Your Intellectual Property in Tests

10:30 AM - 12:00 PM

Type: Clinical

Location: Sanibel 1

Level: All

Intellectual property (IP) is among the most critical assets of test publishers. It mostly takes the form of copyrighted test protocols including test items, manuals, scoring criteria; trade secrets in the scoring algorithms; and trademark rights in the test name. This session will focus on general risk reduction strategies, especially as to IP in test publishing and copyright issues. Multi-Health Systems Inc. will discuss the general topics to cover in an IP audit, goals of IP due diligence, and a wide range of e-commerce related regulatory and legal overlaps, such as heightened privacy concerns and increasing patient access rights in the advent of e-commerce; and Hogrefe Publishing will discuss the problems in protecting test IP that arise because the unique properties of tests often do not reside primarily in the concrete form of the published items, but at a more abstract level. Can copyright laws protect tests against unauthorized "cloning"? A prototypical legal action is presented.

Presenters: Juergen Hogrefe, Publisher/CEO, Hogrefe & Huber Publishers
Rita Chadda, Legal Counsel & Director of Legal Affairs, Multi-Health Systems, Inc.

Improving the Accessibility of Large-scale Reading Assessments for Students with Disabilities

10:30 AM - 12:00 PM

Type: Education

Location: Tallahassee 3

Level: All

Large-scale assessments of student achievement are key components of educational accountability and reform, and no area of achievement receives more attention than reading. The No Child Left behind Act of 2001 (NCLB) requires states to assess student proficiency in "mathematics and reading or language arts". However, for students who have disabilities that affect reading, conventional assessments of achievement can be problematic. The papers presented in this symposium report on two years of research carried out by the National Accessible Reading Assessment Projects (NARAP). The first two papers provide an overview of the two projects that compose NARAP; Designing Accessible Reading Assessments (DARA), and Partners for Accessible Reading Assessments (PARA). The remaining two papers present the outcomes of the first year of research. The authors of the papers will discuss the implications of the research for improving both the design and administration of reading assessments for students with disabilities.

Presenters: Linda Cook, Principal Research Scientist, Educational Testing Service
Martha Thurlow, Director, National Center on Educational Outcomes
Cara Cahalan-Laitusis, Research Scientist, Educational Testing Service
Jamal Abedi, University of California, Los Angeles

Tuesday Conference Program

Reporting Test Scores in Meaningful Ways to Candidates: We Must Do Better!

10:30 AM - 12:00 PM

Type: Education

Location: Captiva 1 - 2

Level: All

Test theory and testing practices have improved considerably over the years with the introduction of computer-based tests, item response theory models, automated test assembly, the emergence of new item types to assess higher-level cognitive skills, and testing agencies are simply better able to handle the basics as well. At the same time, methods for reporting test scores and diagnostic information to candidates remain understudied and undervalued as a problem for credentialing agencies. This is an unfortunate state-of-affairs because there is substantial evidence suggesting that candidates often are confused by the meaning of their test scores and disappointed by the limited amount of diagnostic information they receive from hours of often expensive testing. Examples are easy to find. The goals of this talk will be to expand on the talk given at the 2005 ATP meeting by: (1) providing more evidence about the nature of the failings of many score reporting methods, (2) developing several promising ideas and providing examples for enhancing the clarity and meaning of candidate score reports, and (3) offering suggestions for evaluating possible score report designs.

Presenter: Ronald K. Hambleton, University of Massachusetts

What Kinds of Assessments Improve Learning?

10:30 AM - 12:00 PM

Type: Education

Location: Sanibel 3

Level: Intermediate

The idea that assessments should, in addition to measuring student achievement, also, in some way, contribute to it has broad appeal. For this reason, a number of vendors offer "formative assessment" systems. Although many of these systems allude to the research basis on formative assessment, few actually embody the characteristics that the research literature suggests would be necessary to increase student achievement. This presentation will produce evidence that assessment can improve achievement only if it directly impacts day-to-day classrooms processes to make teaching more responsive to the needs of students in real time. Further, Dr. William will argue that this primarily calls for teacher professional development, rather than the provision of sophisticated data analysis tools.

Presenter: Dylan William, Educational Testing Service

Legal Issues in Personnel Selection: A Summary of Recent Court Decisions

10:30 AM - 12:00 PM

Type: I/O

Location: Tallahassee 1

Level: Intermediate

This presentation will focus on three court opinions that have implications for the use of tests in the employment process. The U.S. Supreme Court opinion in *Smith v. City of Jackson*, 03-1160 (2005) focused on interpreting whether the Age Discrimination in Employment Act ("ADEA") recognizes the doctrine of disparate impact as a basis of recovery in age discrimination suits. The court's holding that disparate impact is a legally recognized claim under the ADEA and the standard for justifying measures that exhibit disparate impact on the basis of age will be examined. Two other cases, *Karraker et al. v. Rent-A-Center*, No. 04-2881 (2005) and *Leonel v. American Airlines*, No. 03-15890 (2005), involving the Americans with Disabilities Act ("ADA") will also be discussed. These opinions from the U.S. Court of Appeals for the 7th and 9th Circuits respectively, address when selection measures can be administered under the ADA. Specifically, the ADA requires that medical examinations and other assessments, which are likely to reveal the existence of a disability, be administered subsequent to the employer tendering a conditional offer of employment. The relevant legislative history of the ADA, as well as the opinions of these two appellate courts will be discussed. Due to the fact that many publications in the HR and legal field have consistently misconstrued the impact of the *Karraker* opinion, significant attention will be given to this decision.

Presenters: David W. Arnold, General Counsel, Wonderlic, Inc.
James C. Sharf, Sharf and Associates, Employment Risk Advisors, Inc.

Tuesday Conference Program

Situational Judgment Tests: Heterogeneity, Validity, Subgroup Differences, and Development Methods

10:30 AM - 12:00 PM

Type: I/O

Location: Sanibel 2

Level: Intermediate

Situational judgment tests have a long history of use in personnel selection. They are best viewed as measurement methods that can and typically do measure multiple constructs. Thus, the measures are typically heterogeneous at the item level and do not lend themselves to typical item analysis procedures (factor analysis, item analysis). Depending on the type of response instructions, situational judgment tests can be viewed as knowledge tests or as assessments of behavioral tendencies. Both criterion-related and construct validity as well as mean racial differences are moderated by response instructions. Relative to situational judgment tests with behavioral tendency instructions, situational judgment tests with knowledge instructions have somewhat higher criterion-related validity, substantially higher correlations with cognitive ability, substantially lower correlations with personality and somewhat higher mean racial differences. The development of situational judgment tests typically rely on substantial input from subject matter experts. Critical incidents written by subject matter experts are typically used to suggest topics for situational judgment item stems. Subject matter experts also provide potential situational judgment item responses and provide ratings on the effectiveness of the response options. Test development time can be reduced by having verbally fluent subject matter experts write situational judgment items targeted to specific KSAs. This presentation has four goals. First, it reviews the heterogeneous nature of situational judgment items. Second, it provides an overview of the validity of situational judgment tests. Third, it reviews mean racial differences in situational judgment tests. Fourth, it reviews common methods for developing situational judgment tests.

Presenters: Michael A. McDaniel, Virginia Commonwealth University
Deborah Whetzel, Work Skills First, Inc.

► **Lunch with Exhibitors and Partners, 12:00 PM – 1:30 PM**

Location: Sun Ballroom C – D and Sun Lobby

► **Breakout Sessions, 1:30 PM - 2:30 PM**

Improving the Validity of Tests and Testing Systems

1:30 PM - 2:30 PM

Type: Cross Divisional

Location: Sanibel 2

Level: All

This session provides theoretical and practical recommendations for investigating and improving the validity of tests and testing systems. This session highlights three complimentary validation approaches: evidence centered design, validity centered design, and validity proposition analysis. The speakers will discuss validity linkages with the following two summary statements: "The [validation] argument must link concepts, evidence, social and personal consequences, and values. (Cronbach, 1988) and "Validity is an inductive summary of both the adequacy of existing evidence for and the appropriateness of potential consequences of test interpretation and use (Messick, 1988)."

Presenters: John Behrens, Sr. Manager, Assessment Development & Innovation, Cisco Systems, Inc
Paul Jones, Psychometrician, Thomson Prometric
James Olsen, VP & Chief Scientist, Alpine Testing Solutions

Tuesday Conference Program

Issues in Unproctored Online Testing

1:30 PM - 2:30 PM

Type: Cross Divisional

Location: Sanibel 1

Level: All

Internet-based test delivery is becoming commonplace in a wide range of applications, such as Human Resource Management (HRM), education, and training and certification. As online testing increases, traditional examination models are evolving and boundaries of practice are being pressed by practical demands. One such practice is the unproctored administration of online tests. At present, this practice is relatively uncharted territory. Little research has been conducted examining the effects and limits of unproctored online testing, and guidelines for test users, where they exist, offer little prescriptive information for practice. This presentation will address key issues, considerations, and questions raised by unproctored administration of online tests, along with possible strategies for minimizing the risk associated with this practice under limited conditions. Critical concerns and threats to measurement, susceptibility vs. suitability of various types of assessments, test use and decision models, and controls and measures that have been taken or proposed to mitigate threats to measurement that are raised by unproctored online testing will be reviewed.

Presenters: John Weiner, Vice President of Products and Services, Psychological Services, Inc.
Doug Reynolds, Vice President of Assessment Technology, Development Dimensions International (DDI)

Leveraging Internet Technology in the Test Development Process

1:30 PM - 2:30 PM

Type: Cross Divisional

Location: Miami 2

Level: All

This session will examine how the Internet is revolutionizing the test development process. The presentation will focus on how, in place of traditional test development methods, virtual meetings and Internet based on-site meetings provide high quality tests with reductions in time to market, errors, and overall cost. Session speakers will conduct actual on-line sample test development meetings.

Presenters: Deborah Schnipke, Virtual Psychometrics
David Scrams, Group Program Manager, Microsoft Corporation
Kirk Becker, Promissor

Testing in the European Union

1:30 PM - 2:30 PM

Type: Cross Divisional

Location: Sanibel 3

Level: All

European Union regulations and standards are quickly evolving to ensure the professional use of assessments. This session will describe some of the trends in EU standards that will affect testing programs over the next 5 to 10 years. Policy, psychometric, and regulatory issues will be highlighted. Additionally some new approaches to assessment will be described as case studies on how testing can be both innovative, professionally sound, and responsive to governmental policies and standards.

Presenter: Jill Lewis, Driving Standards Agency UK

Tuesday Conference Program

Web-based Surveys: A Useful New Tool for Testing Organizations

1:30 PM - 2:30 PM

Type: Cross Divisional

Location: Tallahassee 2

Level: All

* Bring your laptop to this session with wireless Internet if available

As the world becomes better connected and as our practices increasingly move across more borders, the Internet has proven to be an increasingly more important tool. One of the many useful features of the Internet is the ability to conduct Web-Based Surveys (WBSs). Within the context of licensure and certification testing, WBSs can be a useful tool for deploying a job analysis/role delineation survey. Other uses include obtaining marketing information from a specific constituent group. For example, some organizations conduct WBSs to understand baseline salary data across work settings and geographic regions. Another WBS use is to provide constituents an opportunity to provide feedback on agency-provided services.

Presenters: Reed Castle, Director of Research & Development, SMT
Casey Marks, Associate Executive Director, National Council of State Boards of Nursing, Inc.
Larry Fabrey, Vice President, Measurement Research, AMP
Tancy Stanbery, Certification Officer, Board of Certification for Emergency Nursing

Competency Directories: The Journey Continues: The Data Have Arrived

1:30 PM - 2:30 PM

Type: Cert/Lic

Location: Miami 1

Level: All

An information-sharing trend focused on job competencies is emerging among a growing number of associations. Competency directories provide information about core competencies common to professionals across associations, specific jobs/professions, and professional development needs of individuals in different jobs. In this session, you will learn how competency directories are developed. A completed directory will be shown. How competency directories are used for informed decision making, their value, and accessibility to organizations will also be discussed.

Presenters: Linda Montgomery, Director Research Services, Thomson Prometric
Susan Jurow, National Association of College and University Business Officers

Item Response Theory Models for Unfolding: Where We Have Been and Where We Are Going?

1:30 PM - 2:30 PM

Type: Clinical

Location: Tallahassee 1

Level: Intermediate

Interest in item response theory (IRT) models for unfolding has increased steadily during the last fifteen years. Unfolding IRT models are appropriate for measuring attitudes using responses to traditional (Likert or Thurstone) attitude questionnaires. They can also be used to measure individual preferences for alternative stimuli and individual status in certain developmental processes that occur in stages. Like other IRT models, unfolding IRT models provide the potential for sample invariant interpretations of item parameters, item invariant interpretations of person parameters, and indices of measurement precision at the individual level. This presentation will review the distinctions between unfolding and cumulative models in the attitude measurement literature. It is argued that using cumulative models to analyze responses to Likert or Thurstone attitude questionnaires can sometimes lead to a specific form of invalidity in which individuals with the most extreme attitudes receive scores that are indicative of more moderate opinions. Use of an unfolding model circumvents this problem.

Presenter: James S. Roberts, University of Maryland

Tuesday Conference Program

Online SAT Essay Scoring: Process, Challenges, and Benefits

1:30 PM - 2:30 PM

Location: Tallahassee 3

Type: Education

Level: Intermediate

The purpose of this session is to provide information on the processes, challenges, and benefits involved in online scoring of the new essay on the SAT Reasoning Test™. Online essay scoring includes remotely training and monitoring thousands of readers, who must reliably complete the scoring of scanned images of handwritten essays within a limited time frame. In March 2005, the new SAT was first administered. Since then, there have been an additional six administrations, with about a half million student essays being scored after each. Essays are scored using a holistic approach in which the piece of writing is considered as a total work. Each essay is independently scored by two qualified readers who take into account such aspects as complexity of thought, substantiality of development, and facility with language. A reader does not judge a work based on its separate traits, but rather on the total impression it creates. The process of selecting the writing prompts, rangefinding pretest samples for anchor papers, and developing annotated online training resources will be explained and illustrated with sample materials. The online qualifying of potential readers, the prompt-specific training of the readers who qualify, the monitoring of reader agreement, and the procedures for validating reader accuracy will also be explained. Both the challenges and benefits of this many-faceted endeavor will be summarized. Suggestions to programs faced with mass scoring of open-ended items or extended writing to prompts will also be discussed.

Presenters: Kathleen Williams, Vice President, Test Development, The College Board
Daisy Vickers, Pearson Educational Measurement

Applicant Reactions to Testing and Selection Procedures

1:30 PM - 2:30 PM

Location: Miami 3

Type: I/O

Level: All

Interest in applicant reactions to tests and other personnel selection procedures has grown over the last decade with the development of models to explain how applicants perceive the hiring process, and how applicant perceptions in turn affect important attitudes and behaviors. The dominant model in this area has been that proposed by Gilliland (1993), which uses organizational justice theory to describe how selection procedures, the HR staff, and the selection context all influence important outcomes such as job acceptance decisions, organizational attractiveness, motivation, test validity, and litigation. Specifically, Gilliland suggests that organizations can take specific steps to mitigate applicants' negative perceptions regardless of whether an applicant is hired or passes a test. In this presentation, models of applicant reactions and the empirical research to date, focusing on whether and when applicants' perceptions of the selection process really matter will be discussed. In addition, how test-taker predispositions such as test-taking self-efficacy and test-taking motivation may affect a range of outcomes including test validity will be examined. Key questions that remain unanswered in the empirical work, current issues regarding high-tech testing and screening, and recommendations for future research will also be reviewed. Finally, specific ways that organizations and test publishers can improve applicant perceptions will be described.

Presenter: Donald Truxillo, Professor, Portland State University

Platinum Sponsor Session (ACT): Filling the Pipeline: Preparing Tomorrow's Professionals

1:30 PM - 2:30 PM

Location: Captiva 2

Type: Sponsor

Level: All

This overview of the ACT Research and Policy briefs brings participants up to date on current issues and recommendations regarding preparation in America for professionals in the coming years. Highlights:

- Too few students are ready for college and work in all three academic areas—English, mathematics, and science. The percentage of ACT-tested high school graduates who met or exceeded all three College Readiness Benchmarks is alarming—a mere 22 percent of the 1.2 million students tested in 2004.
- College Retention Data collected by ACT show that up to one-fourth of all students at four-year colleges do not return for their second year of school.
- Employee Turnover Reduction –Owensboro Medical Health System has seen its turnover rate reduced by 32 percent since implementing an employee assessment and training program.
- Certification: What are the pros and cons of voluntary vs. mandatory certification and why does it matter?

Presenter: Terry Ausman, Director, Business Development, ACT

Tuesday Conference Program

Platinum Sponsor Session (AMP): Which is More Important . . . Technology or How it is Applied?

1:30 PM - 2:30 PM

Type: Sponsor

Location: Captiva 1

Level: All

With an appropriate focus on technology, organizations must also be cognizant of the way in which technology is applied. This session will focus on providing examples of ways in which technology solutions developed by Applied Measurement Professionals is being used to the advantage of clients, candidates, and other customers. Efficiency for client volunteer groups has been realized through web-based meetings and other remote applications. Convenience and service to candidates have been realized through the development of web application and scheduling processes. And assurance of security has been enhanced through the development of technologically sound monitoring systems. However, with the advent of these systems, it is important to remember to maintain a balance of advanced applications of technology with people who care enough to provide excellent customer service.

Presenters: Bill Hogan, Vice President, Marketing, AMP
Patrick Gallagher, Director of Marketing, AMP

Platinum Sponsor Session (Pearson VUE): Pearson VUE's Six Service Model: A Disciplined Approach and Methodology for Computer-based Testing Programs

1:30 PM - 2:30 PM

Type: Sponsor

Location: Daytona 2

Level: All

Whether national or global, computer-based testing programs are extremely complex and afford numerous opportunities for failure. If realized, these failures can lead to incessant customer service issues or even major program incidents that can affect the very integrity of your testing program. Pearson VUE uses a six service model, a highly disciplined approach and framework to fully plan, implement and operate all aspects of a testing program. This Sponsor Session will discuss the lifecycle of a computer based testing program and demonstrate how Pearson VUE uses the six service model to achieve technical and operational excellence.

Presenters: Kris Kohls, Pearson VUE
Mark Poole, Director of Test Security, Pearson VUE
Jodi Peddycoart, Account Executive, Pearson VUE
Deborah Smith, Product Manager, Pearson VUE

Platinum Sponsor Session (Promissor): Florida, The Model of Assessment

1:30 PM - 2:30 PM

Type: Sponsor

Location: Daytona 1

Level: Intermediate

According to a 2005 Census Bureau report, the State of Florida led all states with 14 rapidly growing counties. This expansion means exponential growth in the areas of teacher certification, real estate, insurance, and other certification and licensing programs. No where has this growth been felt more then at Promissor®. We deliver over 225,000 examinations annually for the State of Florida into more than 25 Promissor-owned, private, and partner centers and in 150 Promissor-owned centers nationwide. Promissor receives 575,000 candidate and licensee calls annually for 11 Florida boards and processes 25,000 reservation and payment transactions monthly. The phenomenal growth in Florida is evidenced by the call volumes in Promissor's call center. In October 2004, Florida averaged 27,468 calls per month and by August of 2005, calls increased to 47,748 for the same programs. Promissor offers key services to these agencies ensuring efficient, flexible, secure, and cost-effective credentialing solutions - no matter how large the volume. In this vendor session, guest speakers from the Florida Departments of Financial Services and Business and Professional Regulation will provide examples of successful strategies for rapidly expanding programs with a focus on speed, security, and accuracy in large scale assessment. They will discuss the synergistic power of working with Promissor who shares the same goals and principles and is committed to the success of their programs.

Presenters: Audrey Huggins, Florida Department of Insurance
Gus Aschoo, Florida Department of Business and Professional Regulation

Tuesday Conference Program

► **Breakout Sessions, 2:45 PM – 3:45 PM**

ATP: A Lobbying and Legal Force for the Testing Industry

2:45 PM – 3:45 PM

Location: Sanibel 1

Type: Cross Divisional

Level: All

Whether it is at the state or federal level, if legislation is being discussed that is relevant to the testing industry, ATP can - and does - influence the legislation so that it is more fair and accommodating to the test publishers. ATP also is a leader among professional testing organizations in protecting test publishers' intellectual property rights and promoting best testing practices. The speakers, the President of ATP and legal counsel to ATP, will talk about specific instances where ATP has played a role, and the ways in which ATP can play a role in the future. Come see how ATP is working for you.

Presenters: John Oswald, Senior Vice President & General Manager, Educational Testing Service
Alan Thiemann, Partner, Thiemann Aitken & Vohra, LLC
Larry Snowwhite, Vice President, Government Relations, Houghton Mifflin Company

Credential Fraud & Privacy Policies – Can They Mix?

2:45 PM – 3:45 PM

Location: Miami 2

Type: Cross Divisional

Level: All

Credential fraud is a significant problem today and threatens to compromise the integrity and credibility of the credential industry and erode market demand for many credentials. Credential Sponsors already recognize the importance of security measures such as secure item banks and proctored exam delivery as means to protect the integrity of their exams and credentials. Credential Sponsors possess the control, authority and responsibility and must take proactive measures to address the issue that a lack of credential verification exposes a critical security gap in the credentialing process. How do global privacy policies govern the release of candidate information? What are some of the best practices used to provide credential verification services? This session will provide an overview of the credential fraud issue as well as a review of the Fair Credit Reporting Act (FCRA) and the Family Educational Rights and Privacy Act (FERPA).

Presenters: Clarke Porter, Co-Founder, CEO, Integral7, Inc.
Mark Poole, Director of Test Security, Pearson VUE

Large Scale Online Assessment: Where Have We Been, What is Needed, Where Are We Going?

2:45 PM – 3:45 PM

Location: Tallahassee 2

Type: Cross Divisional

Level: Intermediate

*** Bring your laptop to this session with wireless Internet if available**

More and more states are moving their traditional paper and pencil (P&P) assessment programs to various approaches (fixed form, adaptive, sequential, etc.) of online computer-based testing (CBT). Today's many state RFPs explicitly call for providers to deliver stable computerized systems as an option. In the years ahead, there will be more and more movement toward CBT as the dominant approach to school testing. The presence of technology alone in schools will cause this to occur. There is little doubt that CBTs will offer advantages. But what of the transition between CBT and today's in-place paper and pencil (P&P) assessments. How does Education move to online testing while assuring fairness and valid appraisal in our nation's schools? Can we create online computerized tests that support instruction not only evaluate it? Is it possible, defensible and necessary for states and schools to operate "dual" programs based on the untested assumption that there may be differences in performance based on the mode of testing? Do P&P test item formats and structures need to change? Is there a real demonstrable and sure need to be attentive to learners whose experience with online accountability testing line in the absence of prior experiences? Can online computerized "early warning assessments" and formative testing be fruitful in informing and guiding instruction and student learning? Also, as states move to embrace online testing are they ready for adaptive testing? These are some of the crucial and pending questions that need to be addressed. It these topics that this session shall address in an effort to inform and guide future and emerging large scale assessment practices.

Presenter: John Poggio, University of Kansas

Tuesday Conference Program

Systematic Tools to Respond to and Remediate Security Incidents

2:45 PM – 3:45 PM

Type: Cross Divisional

Location: Miami 3

Level: All

This session looks at different tools used to detect and enforce test security policies. It will begin by reviewing IP protection strategies and Testlet design with examples illustrating each. A Security Incident Response matrix that catalogs the types of security breaches and their associated sanctions will also be reviewed. Organizations can make decisions and take action against perpetrators in a consistent, legally defensible manner when such tools are employed to manage test theft, fraud, and piracy. The presentation will also explore additional strategies and techniques that can be used to: - compare student papers against a database of previously authored papers and reference sources to detect plagiarism - monitor Internet sites and discussion groups to detect academic honesty policy violations - prevent test content theft by using white boards, virtual calculators, and word processing utilities at computer-based test sites - conduct psychometric and content analyses to identify instances of cheating and test content theft - determine the extent to which specific examinees may have copies of paraphrased items.

Presenters: Patrick Jones, Dean of Assessment, Excelsior College
Adrienne Acosta, TDP Program & Account Manager, Cisco Systems, Inc.
F. Jay Breyer, Executive Director, Psychometrics, Thomson Prometric
Victoria Quinn-Stephens, Manager, Security Enforcements, Cisco Systems, Inc.

Testing Across Languages and Cultures: Methods for Ensuring and Evaluating Measurement Equivalence

2:45 PM – 3:45 PM

Type: Cross Divisional

Location: Miami 1

Level: Intermediate

The global marketplace, international comparisons of educational achievement and cross-cultural research require measuring common attributes and perceptions across cultural and linguistic boundaries. This presentation will cover quality control procedures in test translation and statistical methods for evaluating the comparability of different language versions of tests and surveys. Many of these statistical methods are also appropriate for evaluating cultural bias in tests administered within a single language as well as for evaluating measurement equivalence across standard and accommodated test administrations.

Presenter: Steve Sireci, UMASS, School of Education

Converting from Paper/Pencil to CBT: A Tale of Two Boards

2:45 PM – 3:45 PM

Type: Cert/Lic

Location: Tallahassee 3

Level: Beginner

This session will discuss various considerations and issues for migrating from pencil/paper-based exams to computer-based testing. Two case studies will be discussed: 1) The National Registry of Emergency Medical Technicians (NREMT) will discuss why a transition was necessary as well as the process and tools necessary to successfully transition to CBT while maintaining the current program. 2) The American Board of Pediatrics will discuss the impact of migration on exam reliability (as the exam will be shorter) and pass/fail classifications (does a shorter exam result in the same folks passing/failing?). Additionally, cut score method changes and equating will also be discussed.

Presenters: Linda Althouse, Director of Psychometrics, American Board of Pediatrics
Lance Blackstone, Product Manager-Item Banking, Pearson VUE
Phil Dickison, Associate Director, National Registry of Emergency Medical Technicians

Tuesday Conference Program

The Evolution of the D.S.M. and How it Effects Test Publishing

2:45 PM – 3:45 PM

Location: Sanibel 2

Type: Clinical

Level: Beginner

The purpose of this session is to introduce test publishers to the process of developing DSM-V, while providing them a road map of what their test publishing company can do now to contribute to this massive effort. Although formal work on the preparation of DSM-V is not scheduled to begin until workgroups are formed in 2006 or 2007 (with eventual publication slated for 2011), several projects are currently underway that aim to enrich the empirical database that will be available to the DSM-V workgroups for their future deliberations. The first phase was the development of "A Research Agenda for DSM-V," published in May 2002, that covered cross-cutting issues in diagnosis (e.g., role of neuroscience and genetics in diagnostic criteria). The second phase involves ten diagnosis-specific research conferences convening over the next three years. In addition, a DSM-V web page (www.dsm5.org) has been established to allow comments about problems with DSM-IV and suggestions for DSM-V to be routed to the future workgroups. The presentation concludes with a consideration of future diagnostic issues and controversies in the area of Dissociative Disorders. This workshop is especially relevant for those test publishers that offer measures of normal personality and clinical psychopathology. This workshop also serves as a networking opportunity for ATP members with the key architect of DSM-V.

Presenters: Jack Jones, President, IPAT
Michael First, Professor of Clinical Psychiatry, Columbia University

Doing CAT Right and Doing CAT Wrong

2:45 PM – 3:45 PM

Location: Tallahassee 1

Type: Education

Level: Intermediate

The Graduate Management Admission Test® was the first large scale assessment to employ computer adaptive testing (CAT). Over the past 8 years CAT has proven to be quite successful in reducing candidate test time, and more importantly, providing a high level of test security. There are numerous issues and numerous pitfalls that must be addressed for CAT to be successful. This session will discuss criteria for a successful CAT and then discuss decisions about algorithms, item pool formation, and item pool rotation. Examples of decisions and their affect on item utilization and security will be presented.

Presenters: Kara Owens, Manager, Assessment Research, Graduate Management Admission Council
Lawrence Rudner, Executive Director for Research and Development, Graduate Management Admission Council

Perspectives on Corporate Testing from an I/O Practitioner

2:45 PM – 3:45 PM

Location: Sanibel 3

Type: I/O

Level: All

This presentation will focus on the state of testing, its uses, its misuses, etc. in the corporate arena. The presenter, Dr. Alan Frost, will discuss real-life experiences in some major U.S. corporate settings. Topics will include examples where testing has been implemented in very effective ways with impressive results, examples where tests have been misused or misunderstood, and some practical methods learned to overcome resistance.

Presenter: Alan Frost, Vice President of Organizational Effectiveness and Staffing, Hughes Supply, Inc.

Tuesday Conference Program

Platinum Sponsor (Cito): Improving Education From the Start

2:45 PM – 3:45 PM

Location: Daytona 1

Type: Sponsor

Level: All

Pyramid is the first elaborated educational program produced for 3-6 year old children. It is a cohesive entity addressing all developmental areas. Teachers observe their pupils, but it is difficult for teachers to objectify their observations. Therefore, Cito has developed monitoring systems for different age groups. The Pre-school monitoring system, consisting of observation lists and a series of computer-based tests within the Pyramid program, helps the pre-school teacher to monitor not only a pupil's individual development, but also the proficiency level of an entire group or class. The data can be transferred to the Pupil monitoring system for the 4-12 year old children, Cito's complete system for longitudinal assessment of a pupil's progress throughout primary education. Cito's Platinum Sponsor Session will show you the effects of implementing a method in pre-school based on pilots in the USA and the importance and value of using high-quality monitoring systems throughout the educational process.

Presenter: Cheryl Smith, Director of Educational Development, Cito, USA

Platinum Sponsor Session (Harcourt Assessment): Creating More Value for Teachers through Innovative Item Types

2:45 PM – 3:45 PM

Location: Captiva 2

Type: Sponsor

Level: All

Test publishers have a significant opportunity to expand the information that tests provide to help teachers meet the individual needs of students. This presentation will focus on innovative item types that help teachers deliver differentiated instruction aligned to where an individual student's learning breaks down within specific content standards. The theory behind this innovation, empirical evidence validating it, and models of instructional action plans driven by data from these items will be presented.

Presenter: Margie Jorgensen, Senior Vice President, Harcourt Assessment, Inc.

Platinum Sponsor Session (Thomson Prometric): Test Development: One Size Doesn't Fit All

2:45 PM – 3:45 PM

Location: Captiva 1

Type: Sponsor

Level: All

The need to choose a test delivery model that meets the requirements of the intended audience and other stakeholders has long been understood. All too often, however, there is a failure to recognize that test development procedures and services must be tailored to meet the specific needs of individual markets. Taking a "one size fits all" approach results in costly inefficiencies and threatens the acceptance and validity of the testing program. This session will highlight the different requirements of various markets and discuss the value of leveraging those commonalities that exist.

Presenters: Stephen Williams, Vice President, Test Development Services, Thomson Prometric
Nikki Eatchel, Executive Director, Test Development, Thomson Prometric

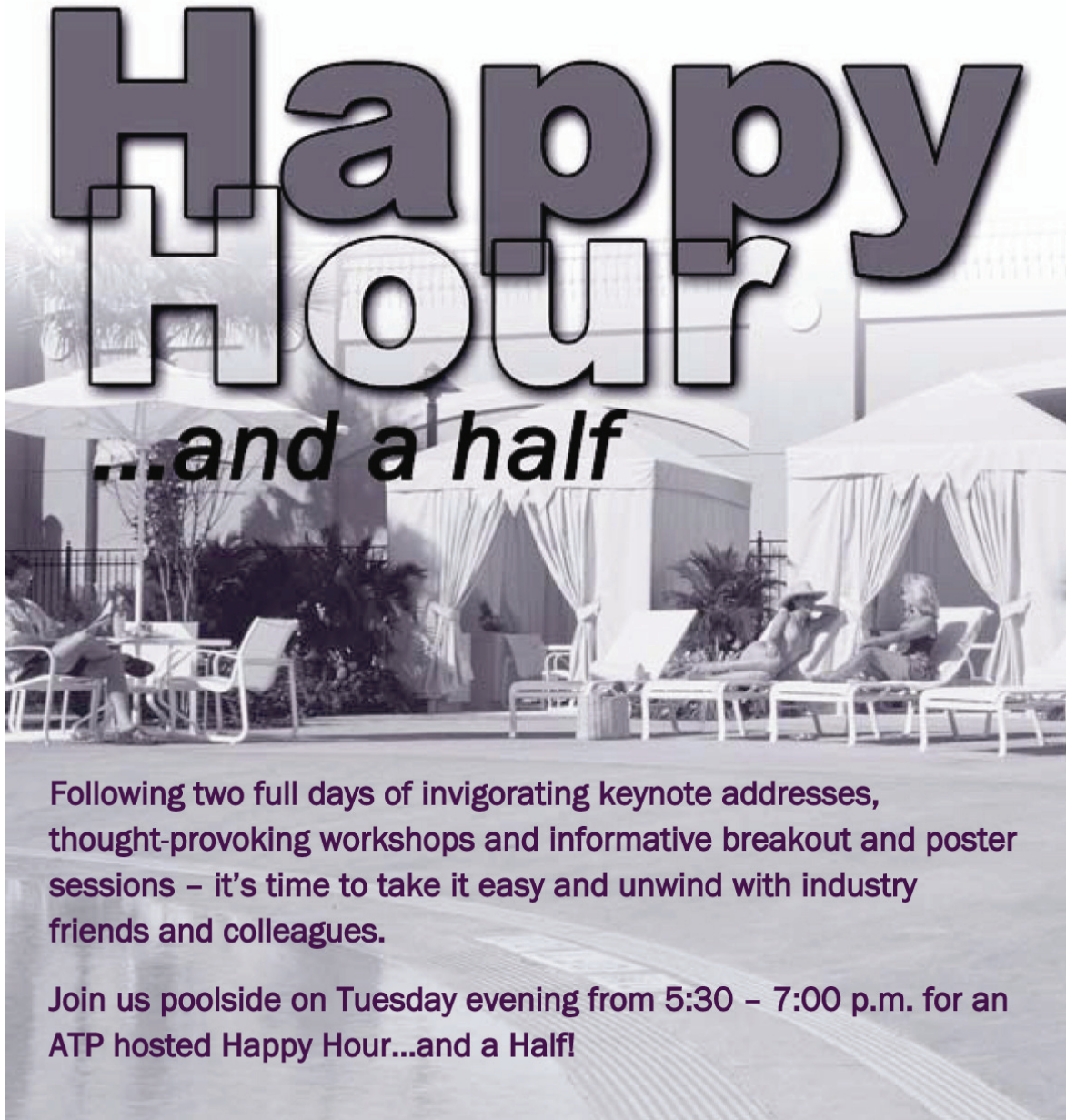
► Exhibits, Partners & Poster Sessions Open, 3:45 PM – 5:30 PM

Location: Sun Ballroom C – D and Sun Lobby

Tuesday Conference Program

► ***Happy Hour...and a Half! 5:30 PM – 7:00 PM***

Location: South Beach Pool



Happy Hour

...and a half

Following two full days of invigorating keynote addresses, thought-provoking workshops and informative breakout and poster sessions – it's time to take it easy and unwind with industry friends and colleagues.

Join us poolside on Tuesday evening from 5:30 – 7:00 p.m. for an ATP hosted Happy Hour...and a Half!