

The U.S. Equal Employment Opportunity Commission

Meeting of May 16, 2007 - Employment Testing and Screening**Statement of Kathleen K. Lundquist, Ph.D.,
Co-Founder and President,
APT, Inc.**

Good morning. Thank you for the opportunity to share some of my thoughts with you about the role employment testing can play in creating a fair and effective workforce. As the last of your speakers today, I will keep my remarks relatively brief.

In a number of recent class actions, concern has been raised about the potential for “excessive subjectivity” in such selection procedures as performance evaluations and unstructured interviews. In contrast, employment tests, when well-developed, can provide a consistent and fair basis on which to make employment decisions, making them less subject to intentional or unintentional bias.

Professionally-developed selection procedures serve a legitimate business purpose: they allow employers to base hiring and promotional decisions on solid, job-related information. The evidence that a selection procedure measures behavior consistently (i.e., its reliability) and is an accurate measure of job performance (i.e., its validity) is the basis on which a selection procedure is shown to be job-related. Job-related procedures ensure that employees possess the necessary skills to perform the job; such procedures can be used by employers to predict which candidates will be able to successfully perform the job. In short, good selection procedures are fair to candidates (i.e., standardized and objective in their administration and scoring) and useful to organizations (i.e., result in gains in overall productivity).

The key steps in developing a good selection procedure or test include:

1. Determining the need and appropriateness of testing;
2. Determining the necessary job requirements;
3. Developing tests to measure the job’s requirements;
4. Collecting data to evaluate test quality;
5. Collecting data to evaluate job-relatedness;
6. Establishing administrative procedures and controls, and
7. Implementation and ongoing monitoring of the test.

One of the most important considerations underlying any selection procedure used for employment decision-making is the assumption that the procedure is asking the right questions, i.e., is measuring the characteristics necessary for successful job performance. This is determined by conducting a job analysis. Job analysis involves systematically analyzing what tasks individuals are expected to perform on a job, and what knowledge, skills, abilities and other characteristics (KSAOs) are required to perform the job tasks. For example, oral communication skills and the ability to organize work would be two of the KSAOs required to perform the tasks in the job of an attorney, while the knowledge of advanced mathematics might not be required for this job. The job analysis thus assists in identifying the content to be measured in the selection procedure and the criteria against which to evaluate test performance. Further evidence of job-relatedness is obtained by formally conducting a validation study to demonstrate the connection between the test and the job, typically either through linking the content of the test with the content of the job (content validity) or by showing a statistically significant relationship between test performance and job performance (criterion-related validity).

One of the major challenges in employment testing has been ensuring that the selection procedure itself fairly represents what an individual will be able to do on the job. Recent applications of technology and research in the testing field have permitted employment tests to measure skills in ways that are more similar to how the individual actually would perform on the job. Research has shown that “high fidelity” selection procedures, such as work samples, video simulations and assessment center exercises, enhance candidate acceptance and often reduce adverse impact.

The use of traditional paper and pencil tests in personnel selection has long been debated based on the trade-off between validity and adverse impact. On the one hand, cognitive ability tests have been shown to be highly predictive of job performance. On the other hand, tests of cognitive ability have also shown higher adverse impact than other types of assessment. Traditionally, White candidates have performed better on cognitive ability tests, demonstrating a 1.0 standard deviation difference (i.e., effect size¹) between African American/Blacks and Whites, a 0.7 standard deviation difference between Hispanics and Whites, and virtually no difference between men and women.

In an attempt to search for less adverse alternatives in our recent work on the Ford Apprentice Testing program, we reviewed the research literature on alternative testing measures that demonstrate good validity with less adverse impact². Test administration format or test medium is an important factor in this equation. Research has shown video-based testing to have comparable validity to paper-based tests, with lower adverse impact. Adverse impact was found to be reduced by enhancing applicants’ job relatedness perceptions, positively impacting test-taker motivation, and reducing the reading comprehension demands of the test. Table 1 shows a comparison of the research on different test administration formats.

**Table 1
Comparison of Test Administration Formats**

	Paper-and-Pencil	Computer-Based	Video-Based
Assessment of KSAOs	(–) Not as wide a range of KSAOs can be assessed vs. computer and video-based tests	(+) Wider range of KSAOs can be assessed vs. paper-and-pencil tests	(+) Wider range of KSAOs can be assessed vs. paper-and-pencil tests (+) Tests can look more like the job
Validity		(+) Same validity as paper-and-pencil tests	(+) Same validity as paper-and-pencil tests
Adverse Impact		(+) Same adverse impact as paper-and-pencil tests	(++) Reduces adverse impact
Development Costs	(+) Less development costs	(–) Increased development costs	(–) Increased development costs
Administration	(+) Cost effective and practical for large group administration	(–) Facilities with many secure computers with internet access required (–) Smaller test sessions needed (+) Test Administrator responsibilities reduced	(+) Practical for large group administration (–) Increased costs for video presentation equipment to ensure adequate viewing for all candidates
Test Security	Security of test materials needed during printing, shipping, and storage	(–) Greater exposure of test content over repeated administrations threatens test security	Security of test materials needed during shipping and storage
			(–) Delayed scoring and

Scoring	(–) Delayed scoring and reporting of results	(+) Real-time scoring and results	reporting of results (answers captured in paper-and-pencil form)
Alternate Forms	(+) Can develop alternate forms of some tests and less costly to develop	(+) Can develop alternate forms of some tests	(–) Difficult and costly to develop alternate forms

In addition, research has shown that, for some jobs, measuring more than just cognitive ability can result in better prediction of overall job success and result in lower adverse impact. Considering both cognitive abilities and non-cognitive personal characteristics (e.g., conscientiousness or customer focus) may give a more complete picture of the qualifications of the candidate. Often, it is not just what you know, it’s how you show it.

Work sample or situational judgment tests have also been shown to be promising ways to maintain validity and decrease adverse impact. These assessments are designed to mirror or simulate the actual tasks performed on the job, for instance through a Manager In-Basket exercise or a video simulation of a production line. Such tests measure the ability to identify and understand job-related issues or problems and to select the proper course of action to resolve the problem. Their good validity stems from having the candidate actually perform a part of the job and their reduced adverse impact appears to result from candidate acceptance and motivation.

In our work on the Ford Apprentice Testing program, we have combined a cognitive test, a non-cognitive assessment and a video-based simulation to measure candidates’ qualifications for the apprenticeship. These tests will undergo validation later this summer.

I would like to focus my final comments to the Commission on two critical areas: the importance of operational validity and the need to provide recognition and encouragement to good testing programs.

Beyond the content and initial development of the test, it is important to ensure that the selection procedures continue to be job-related in actual use over time (operational validity). This involves monitoring and revising the selection procedures when jobs change, ensuring that those who use and score the tests are fully trained, and regularly monitoring test data to detect and address observed patterns of unfairness. I would encourage the Commission to develop standards and review procedures to audit testing programs for their operational validity, as well as their initial job-relatedness.

Finally, there are many good testing programs which could serve as models for employers wishing to implement effective selection procedures. I encourage the Commission to provide recognition to such programs.

Thank you.

References

Barrick, M.R. & Mount, M.K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44(1), 1-26.

Batram, D. & Brown, A. (2004). Online testing: Mode of administration and the stability of OPQ 32i scores. *International Journal of Selection and Assessment*, 12(3), 278-284.

Bobko, P., Roth, P.L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology*, 52, 561-589.

Buchanan, T., & Smith, J.L. (1999). Using the Internet for psychological research: Personality testing on the World Wide Web. *British Journal of Psychology*, 90(1), 125-144.

- Campion, M.A., Pursell, E.D., & Brown, B.K. (1988). Structured interviewing: Raising the psychometric properties of the employment interview. *Personnel Psychology*, 41, 25-42.
- Caretta, T.R., & Ree, M.J. (1997). Negligible sex differences in the relation of cognitive and psychomotor abilities. *Personality and Individual Differences*, 22(2), 165-172.
- Chan, D., & Schmitt, N. (2002). Situational judgment and job performance. *Human Performance*, 15(3), 233-254.
- Chan, D., & Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology*, 82, 143-159.
- Clevenger, J., Pereira, G.M., Wiechmann, D., Schmitt, N., & Harvey, V.S. (2001). Incremental validity of situational judgment tests. *Journal of Applied Psychology*, 86, 410-417.
- Conway, J.M., Jako, R.A., & Goodman, D.F. (1995). A meta-analysis of interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology*, 80, 565-579.
- Curtis, J.R., Gracin, L., & Scott, J.C. (1994, April). Non-traditional measures for selecting a diverse workforce: A review of four validation studies. In J.C. Scott (Chair), *Selecting and managing a diverse workforce: The role of assessment*. Symposium conducted at the Ninth Annual Conference of the Society for Industrial and Organizational Psychology, Inc., Nashville, TN.
- Dalessio, A.T. (1994). Predicting insurance agent turnover using a video-based situational judgment test. *Journal of Business and Psychology*, 9(1), 23-32.
- Dean, M.A. (2004). An assessment of biodata predictive ability across multiple performance criteria. *Applied H.R.M. Research*, 9(1), 1-12.
- Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440.
- Dunnette, M.D. (1976). Aptitudes, abilities, and skills. In M.D. Dunnette (Ed.), *Handbook of Industrial and Organizational Psychology* (pp. 473-520). Chicago: Rand McNally Publishing Company.
- Dyer, P.J., Desmarais, L.B., & Masi, D.L. (1994, April). Multimedia approaches to testing: An aid to workforce diversity? In J.C. Scott (Chair), *Selecting and managing a diverse workforce: The role of assessment*. Symposium conducted at the Ninth Annual Conference of the Society for Industrial and Organizational Psychology, Inc., Nashville, TN.
- Furnham, A. & Medhurst, S. (1995). Personality correlates of academic seminar behavior: A study of four instruments. *Personality and Individual Differences*, 19, 197-208.
- Gallagher, A., Bridgeman, B., & Cahalan, C. (2002). The effect of computer-based tests on racial-ethnic and gender groups. *Journal of Educational Measurement*, 39(2), 133-147.
- Greud, V.A. & Green, B.F. (1986). Equivalence of conventional and computer presentation of speeded tests. *Applied Psychological Measurement*, 10, 23-34.
- Groth-Marnat, G. & Schumaker, J. (1989). Computer-based psychological testing: Issues and guidelines. *American Journal of Orthopsychiatry*, 59, 257-263.
- Harris, M.M. (1989). Reconsidering the employment interview: A review of recent literature and suggestions for future research. *Personnel Psychology*, 42, 691-727.
- Hedges, L.V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of

high-scoring individuals. *Science*, 269, 41-45.

Higuera, L.A. & Riera, M.C. (2004). Validation of a trainability test for young apprentices. *European Psychologist*, 9(1), 56-63.

Hough, L.M. (1998). Personality at work: Issues and evidence. In M.D. Hakel (Ed.), *Beyond Multiple Choice: Evaluating Alternatives to Traditional Testing for Selection* (pp. 131-159). Mahwah, NJ: Lawrence Erlbaum Associates.

Hough, L.M., Oswald, F.L., & Ployhart, R.E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment*, 9(1/2), 152-194.

Huffcutt, A.I., & Arthur, Jr., W. (1994). Hunter & Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, 79(2), 184-190.

Huffcutt, A.I., & Roth, P.L. (1998). Racial group differences in employment interview evaluations. *Journal of Applied Psychology*, 83, 179-189.

Huffcutt, A.I., Roth, P.L., & McDaniel, M.A. (1996). A meta-analytic investigation of cognitive ability in employment interview evaluations: Moderating characteristics and implications for incremental validity. *Journal of Applied Psychology*, 81, 459-473.

Hunter, J.E. & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96(1), 72-98.

Hunter, J.E., Schmidt, F.L., & Hunter, R. (1979). Differential validity of employment tests by race: A comprehensive review and analysis. *Psychological Bulletin*, 86(4), 721-735.

Hurtz, G.M. & Donovan, J.J. (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology*, 85, 869-879.

Hyde, J.S., Fennema, E., & Lamon, S.J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107(3), 139-155.

Judge, T.A., & Barrick, M.R. (2001). *Personality and Work*. Tutorial presented at the Sixteenth Annual Conference of the Society of Industrial and Organizational Psychology, San Diego, CA.

Kehoe, J.F. (2002). General mental ability and selection in private sector organizations: A commentary. *Human Performance*, 15(1/2), 97-106.

Lefkowitz, J., Gebbia, M.I., Blasam, T., & Dunn, L. (1999). Dimensions of biodata and their relationships to item validity. *Journal of Occupational and Organizational Psychology*, 72, 331-350.

Lievens, F., & Coetsier, P. (2002). Situational tests in student selection: An examination of predictive validity, adverse impact, and construct validity. *International Journal of Selection and Assessment*, 10(4), 245-257.

Lyons, T.J., Bayless, J.A., Park, R.K. (2001). Relationship of cognitive, biographical, and personality measures with the training and job performance on detention enforcement offices in a federal government agency. *Applied H.R.M. Research*, 6(1), 67-70.

McDaniel, M.A., Morgeson, F.P., Finnegan, E.B., Campion, M.A., & Braverman, E.P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, 86, 730-740.

McDaniel, M.A., Whetzel, D.L., Schmidt, F.L., & Maurer, S.D. (1994). The validity of employment

- interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599-616.
- Mead, A.D. & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin*, 114(3), 449-458.
- Motowidlo, S.J., Dunnette, M.D., & Carter, G.W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology*, 75, 640-647.
- Motowidlo, S.J., & Tippins, N. (1993). Further studies of the low-fidelity simulation in the form of a situational interview. *Journal of Occupational and Organizational Psychology*, 66, 337-344.
- Mount, M.K., Witt, L.A., & Barrick, M.R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology*, 53, 299-323.
- Neuman, G & Baydoun, R. (1998). Computerization of paper-and-pencil tests: When are they equivalent? *Applied Psychological Measurement*, 22, 71-83.
- Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66(6), 574-583.
- Olson-Buchanan, J.B., Drasgow, F., Moberg, P.J., Mead, A.D., Keenan, P.A., & Donovan, M.A. (1998). Interactive video assessment of conflict resolution skills. *Personnel Psychology*, 51, 1-24.
- Pearlman, K., Schmidt, F.L., & Hunter, J.E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology*, 65, 373-406.
- Ployhart, R.E., Weekley, J.A., Holtz, B.C., & Kemp, C. (2003). Web-based and paper-and-pencil testing of applicants in a proctored setting: Are personality, biodata, and situational judgment tests comparable? *Personnel Psychology*, 56, 733-752.
- Reilly, R.R. & Chao, G.T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology*, 35, 1-62.
- Reilly, R.R. & Israelski, E.W. (1988). Development and validation of minicourses in the telecommunication industry. *Journal of Applied Psychology*, 73, 721-726.
- Richman-Hirsch, W.L., Olson-Buchanan, J.B., Drasgow, F. (2000). Examining the impact of administration medium on examinee perceptions and attitudes. *Journal of Applied Psychology*, 85, 880-887.
- Robertson, I.T. & Downs, S. (1989). Work-sample tests of trainability: A meta-analysis. *Journal of Applied Psychology*, 74, 402-410.
- Robertson, I.T. & Downs, S. (1979). Learning and the prediction of performance: Development of trainability testing in the United Kingdom. *Journal of Applied Psychology*, 64, 42-50.
- Robertson, I.T. & Kandola, R.S. (1982). Work sample tests: Validity, adverse impact and applicant reaction. *Journal of Occupational Psychology*, 55, 171-183.
- Robertson, I.T. & Mindel, R.M. (1980). A study of trainability testing. *Journal of Occupational Psychology*, 53, 131-138.
- Roth, P.L., Bevier, C.A., Bobko, P., Switzer, F.S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, 54, 297-330.

- Rothstein, H.R., Schmidt, F.L., Erwin, F.W., Owens, W.A., & Sparks, P. (1990). Biographical data in employment selection: Can validities be made generalizable? *Journal of Applied Psychology, 75*, 175-184.
- Salgado, J.F. & Moscoso, S. (2003). Internet-based personality testing: Equivalence of measures and assesses' perceptions and reactions. *International Journal of Selection and Assessment, 11*(2/3), 194-204.
- Salgado, J.F., Viswesvaran, C., & Ones, D.S. (2001). Predictors used for personnel selection: An overview of constructs, methods, and techniques. In N. Anderson, D.S. Ones, H.K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of Industrial, Work, and Organizational Psychology* (pp.165-199). London: Sage Publications.
- Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262-274.
- Schmidt, F.L., Gast-Rosenberg, I., & Hunter, J.E. (1980). Validity generalization for computer programmers. *Journal of Applied Psychology, 65*, 643-661.
- Schmitt, N., Clause, C.S., & Pulakos, E.D. (1996). Subgroup differences associated with different measures of some job-related constructs. *International Review of Industrial and Organizational Psychology, 11*, 115-139.
- Schmitt, N., Gooding, R.Z., Noe, R.A., & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigations of study characteristics. *Personnel Psychology, 37*, 407-422.
- Schmitt, N., Rogers, W., Chan, D., Sheppard, L., Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. *Journal of Applied Psychology, 82*, 719-730.
- Siegel, A.I. (1983). The miniature job training and evaluation approach: Additional findings. *Personnel Psychology, 36*, 41-56.
- Smiderle, D., Perry, B.A., & Cronshaw, S.F. (1994). Evaluation of a video-based assessment in transit operator selection. *Journal of Business and Psychology, 9*(1), 3-22.
- Smith, M.C. & Downs, S. (1975). Trainability assessments for apprentice selection in shipbuilding. *Journal of Occupational Psychology, 48*, 39-43.
- Spray, J.A., Ackerman, T.A., Reckase, M.D. & Carlson, J.E. (1989). Effect of the medium of item presentation on examinee performance and item characteristics. *Journal of Educational Measurement, 26*, 261-271.
- Tenopyr, M.L. (1996). Gender issues in employment testing. In R.S. Barrett (Ed.), *Fair Employment Strategies in Human Resource Management* (pp. 193-197). Westport, CT: Quorum Books.
- Weekely, J.A., & Jones, C. (1999). Further studies of situational tests. *Personnel Psychology, 52*, 679-700.
- Weekley, J.A. & Jones, C. (1997). Video-based situational testing. *Personnel Psychology, 50*, 25-49.
- Wernimont, P. & Campbell, J. (1968). Signs, samples, and criteria. *Journal of Applied Psychology, 52*, 372-376.
- Wiesner, W.H., & Cronshaw, S.F. (1988). A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational*

Psychology, 61, 275-290.

Wise, S.L. & Plake, B.S. (1989). Research on the effects of administering tests via computers. *Educational Measurement: Issues and Practice*, 5-10.

Footnotes

¹ Effect size indicates the number of standard deviation units that separates non-protected and protected group means.

² A complete list of the literature reviewed is contained in the reference list at the end of this paper.

This page was last modified on May 16, 2007.



[Return to Home Page](#)