

# Legally Defensible Test Development

David Arnold, Ph.D., J.D.

Wonderlic, Inc.

dave.arnold@wonderlic.com

Lisa W. Borden, J.D.

Baker, Donelson, Bearman, Caldwell & Berkowitz, P.C.

lborden@bakerdonelson.com

James C. Sharf, Ph.D.

Employment Risk Advisors, Inc.

jim@jimsharf.com

# Uniform Guidelines on Employee Selection Procedures, Sec.2(B).

- Employment decisions. “These guidelines apply to tests and other selection procedures which are used as a basis for any employment decision. Employment decisions include but are not limited to hiring, promotion, demotion, membership (for example, in a labor organization), referral, retention, and licensing and certification, to the extent that licensing and certification may be covered by Federal equal employment opportunity law (emphasis added)...”

# Uniform Guidelines on Employee Selection Procedures

- Uniform Guidelines’ focus: “job performance,” not “necessary prerequisite” knowledge:
  - “The fundamental principle underlying the guidelines is that employer policies or practices which have an adverse impact on the opportunities of any race, sex, or ethnic group are illegal under Title VII and the Executive Order unless justified by business necessity... If adverse impact exists, it must be justified on grounds of business necessity. Normally this means by validation which demonstrates the relation between the selection procedure and performance on the job” (emphasis added).

# Uniform Guidelines Q&A#7

- “Q: Do the Guidelines apply to the licensing and certification functions of state and local governments?”
- A: The Guidelines apply to such functions to the extent that they are covered by Federal law. The courts are divided on the issue of such coverage. The Government has taken the position that at least some kinds of licensing and certification which deny persons access to employment opportunity may be enjoined in an action brought pursuant to Sec. 707 of the Civil Rights Act of 1964, as amended.”

# Uniform Guidelines on Employee Selection Procedures

- Uniform Guidelines do allow for sampling “necessary prerequisite” knowledge
- “In the case of a selection procedure measuring a knowledge, the knowledge being measured should be operationally defined as that body of learned information which is used in and is a necessary prerequisite for observable aspects of work behavior of the job” (emphasis added).

## “Applicant” Defined by OFCCP (1979)

- “The precise definition of the term ‘applicant’ depends upon the user’s recruitment and selection procedures. The concept of an applicant is that of a person who has indicated an interest in being considered for hiring, promotion or other employment opportunities.” Federal Register (March 2, 1979), 44(43).

# New Proposed Q&A

- In March 2004, the Equal Employment Opportunity Commission (EEOC) and the Department of Labor's Office of Federal Contract Compliance Programs (OFCCP) SEPARATELY proposed the Adoption of Additional Questions and Answers to Clarify and Provide a Common Interpretation of the Uniform Guidelines on Employee Selection Procedures as They Relate to the Internet and Related Technologies.

# New Proposed Q&A

- The intent of the proposed Qs & As was to guide employers in their efforts “to comply with requirements of Federal law prohibiting employment practices which discriminate on grounds of race, color, religion, sex, and national origin” 29 C.F.R. ¶ 1607.1(B).

# New Proposed Q&A

- “OFCCP requires covered federal contractors to obtain, where possible, gender, race and ethnicity data on applicants and employees.” “OFCCP ‘selects’ contractors for compliance audits based on statistical analyses of gender, race and ethnicity data contractors submit to OFCCP.” Federal Register, 69(60), Mar.29, 2004, 16446.

# New Proposed Q&A

- EEOC: “These Additional Questions and Answers are intended to clarify how the Uniform Guidelines on Employee Selection Procedures apply in the context of the Internet and related technologies.” Federal Register, 69(43), Mar.4, 2004, 10152.

# New Proposed Q&A

- OFCCP: “The proposed rule would.... incorporate the new category of ‘Internet Applicant’ ... to distinguish between ‘applicants,’ i.e., submissions of interest that are not submitted through the Internet and related electronic technologies, and ‘Internet Applicants’.”

# New Proposed Q&A

- OFCCP: “The proposed rule creates differing standards for data collection for traditional applicants versus Internet Applicants for the same job. Accordingly, if an employer’s recruitment processes for a particular job involve both electronic data technologies, such as the Internet, and traditional want ads and mailed, paper submissions, the proposed rule would treat these submissions differently for that particular job.”

# New Proposed Q&A

- OFCCP: “Under the proposed rule, the agency will rely on labor force statistics or other relevant data for enforcing E.O. 11246 with respect to recruitment processes that occur prior to collection of gender, race and ethnicity data. This approach is consistent with the longstanding approval of such statistics in hiring discrimination litigation and is especially appropriate because the proposed definition of ‘Internet Applicant’ relates to ‘advertised basic qualifications’.”

# New Proposed Q&A

- EEOC: “While the Internet and related technology has transformed recruitment and job hunting in recent years, our country’s employment nondiscrimination laws, such as Title VII of the Civil Rights Act of 1964 and Executive Order 11246 continue to apply to all aspects of employment including recruitment.”

# New Proposed Q&A

- EEOC: “The Uniform Guidelines on Employee Selection Procedures were issued in 1978 by the EEOC, the Department of Labor, the Department of Justice, and the Office of Personnel Management... serv(ing) two major purposes:... 1) recordkeepng, and ... 2)detail(ed) methods for validating tests and selection procedures that are found to have a disparate impact.”

# New Proposed Q&A

- EEOC: “Disparate impact is when an employer uses a practice or standard, like a hiring or promotion requirement or an employment test, that has a statistically significant disproportionate negative effect on a protected group, even though the standard or test is not intentionally discriminatory. Such a practice or standard is unlawful under Title VII if it is not job-related and consistent with business necessity” (emphasis added).

# New Proposed Q&A

- EEOC: “Interpreting the definition of ‘applicant’ in the context of the Internet and related electronic processing technology is the focus of this document. With this interpretation, the UGEST agencies are providing guidance about when employers should identify the race, gender, and ethnicity of their applicant pool when they use the Internet and related technologies.”

# New Proposed Q&A

- OFCCP: “See, e.g. Griggs v. Duke Power (relying on Census data about the general population to find that a high school degree requirement had a disparate impact on African Americans); Dothard v. Rawlinson (‘The application process itself might not adequately reflect the actual potential applicant pool, since otherwise qualified people might be discouraged from applying because of a self-recognized inability to meet the very standards challenged as being discriminatory.’)...”

# New Proposed Q&A

- In proposing to define who is a bona fide “e-Applicant,” however, the enforcement agencies blurred the previous distinction between recruitment practices which were not burdened under the Uniform Guidelines and selection practices which were.

# New Proposed Q&A

- “Thus, OFCCP will compare the proportion of women and minorities in the contractor’s relevant applicant pool with labor statistics or other data on the percentage of women and minorities in the relevant labor force. If there is a significant difference between these figures, OFCCP will investigate further as to whether the contractor’s recruitment and hiring practices conform with E.O. 11246 standards.”

# New Proposed Q&A

- Among other issues are: 1) whether an employer is obligated when merely announcing that they are hiring (i.e. “broadcast” recruitment) or obligated only under the more restrictive, “position specific” announcement; and 2) what the obligation of a third party web site (e.g., Monster) is with regard to the collection, retention and sharing of RSN data.

# New Proposed Q&A

- The thrust of the proposed Q&As is to clarify that recruiting announcements used to develop their pool of qualified applicants (licensed / certified) are to be newly subjected to disparate impact analysis and that such analysis “can be based on Census or workforce data.”

# New Proposed Q&A

- In other words, the enforcement agencies are licensing themselves to generalize an inference of adverse impact based on their presumption of what an employer's internet applicant flow should resemble in terms of RSN composition rather than who actually applied. Census occupational classification codes will likely be the possibly over-broad benchmarks against which internet candidate or applicant pools will be compared.

# New Proposed Q&A

- Unresolved also is whether pencil & paper applicant selection procedures will be similarly analyzed. As one labor economist opined, “The proposed Q & As are my retirement annuity.”

# US Department of Commerce: “Falling Through the Net”

- People with a college degree are eight times more likely to have a PC at home and 16 times more likely to have Internet access at home than those with an elementary school education.
- A high-income household in an urban area is 20 times more likely to have Internet access than a rural, low-income household.
- A child in a low-income white family is three times more likely to have Internet access than a child in a comparable black family, and four times more likely than if he were Hispanic. The Economist (24 June, 2000). Government and the Internet. p.24.

# US Department of Commerce: “Falling Through the Net”

- A wealthy household of Asian descent is 34 times more likely to have Internet access than a poor black household.
- A child in a two-parent white household is twice as likely to have Internet access as a child in a single-parent household. If the child is black, he is four times more likely to have Internet access than his single-parent counterpart.
- Disabled people are nearly three times less likely to have home access to the Internet than people without disabilities.

# Implications for the Future

- If adopted as proposed, recruitment practices including meeting minimum qualifications (“MQs”) such as educational diplomas and passing licensing and certification exams will be newly challenged under the Uniform Guidelines.

# Implications for the Future

- As a practical matter, minimum qualifications (e.g., diploma requirements, certification exams) resulting in adverse impact (i.e., not representatively attained by each RSN subgroup in the labor market) must be “job related for the position in question and consistent with business necessity” as defined in the Civil Rights Act of 1991.

# Implications for the Future

- Wholly aside from burdening occupational licensing and certification, imagine the psychometric defense (“psychomagician”?) of the content validity of a high school diploma for an hourly position or BS in liberal arts prerequisite for a salaried position.

- Gulino v. The Board of Education of the City School District of the City of New York & the New York State Education Dept
- New York State Board of Regents Certification Requirements
- The New York State Board of Regents has statutory authority to define educational policy for state-supported schools from kindergarten through graduate school.

- The Board of Regents appoints the Commissioner of the State Education Department (SED) who is accountable to the legislature for: overseeing curricula, developing and administering state-wide Board of Regents' examinations, and certifying the competence of prospective teachers. A certified prospective teacher is awarded a provisional teaching certificate (State certificate) which good for five years.

- In 1987, the Ambach Commission's Blueprint for Learning recommended that the Board of Regents examine teacher candidates statewide to certify each person's grounding in the liberal arts and sciences. In 1990, the Board of Regents amended examination regulations for provisional State certification to require a candidate's demonstrating minimum competency in the liberal arts and sciences.

# NES: Liberal Arts and Sciences Test (1993 – present)

- First administered in 1993, the Liberal Arts and Sciences Test (LAST) was designed to be a valid sample of those competencies: “The LAST consists of multiple-choice test questions and a written assignment. Candidates taking the LAST are asked to demonstrate conceptual and analytical skills, critical-thinking and communications skills, and multicultural awareness. The test covers scientific and mathematical processes, historical and social scientific awareness, artistic expression and the humanities, communication skills, and written analysis and expression.”

# ETS: National Teacher Exam Core Battery Tests (1984 – 1993)

- Between 1984 and 1993, candidates were required to demonstrate minimum competency on the National Teacher Exam (NTE) Core Battery Tests.

# NTE Core Battery

- Professional knowledge: factors affecting the learning environment.
- General Knowledge:
  - Social Sciences: both knowledge and interpreting and analyzing data,
  - Science: principles, theories, facts, and the scientific method,
    - » continued

# NTE Core Battery

- Literature and Fine Arts: recognition of the basic elements of literature and art,
- Mathematics: skills essential to explaining quantitative data, and
- Communication skills

# Teacher Certification Examinations

- In addition to LAST, the Board of Regents' requirements for lifetime teaching licensure include:
  - An MS degree, completion of teacher preparation coursework, passing a written Content Specialty Test (CST),
  - passing a written Assessment of Teaching Skills – Written (ATS-W), and
  - passing an Assessment of Teaching Skills – Performance (ATS-P) which is a video sample of classroom teaching.

# Teacher Certification Examinations

These comprehensive written and performance exams are known as the New York State Teacher Certification Examinations (NYSTCE) the purpose of which “is to help identify candidates for teacher certification who have demonstrated the level of knowledge and skills that are important for performing the job of a teacher in New York State public schools.”

## CERTIFICATION = EMPLOYMENT DECISION ?

- In carrying out their legislated mandate to certify the minimum liberal arts and sciences competency of each prospective teacher candidate, the Board of Regents makes no employment decision and is therefore not legally considered to be an employer.

# According to the National Research Council:

- “Licensure is a state function. It is aimed, above all, at protecting the public. As defined by the federal government, licensure is ‘the process by which an agency of government grants permission to persons to engage in a given profession or occupation by certifying that those licensed have attained the minimal degree of competency necessary to ensure that the public health, safety, and welfare will be reasonably well protected’ (U.S. Department of Health, Education, and Welfare, 1971)...”

# According to the National Research Council:

- “Licensure is distinct from hiring. Although state licenses grant permission for teachers to teach in public schools, local agencies – school districts and, in many cases, schools – actually hire the teachers and so determine who will teach and what they will teach. The districts and schools that hire teachers may decide to use criteria in addition to the holding of a license in deciding which teachers to hire.”

# New York City Board of Education

## Licensure Requirements

- The Board of Education of the City School District of the City of New York (Board of Education) sets educational policy for public schools in 31 school districts across 5 boroughs. The Board of Education is accountable for making employment decisions that are in compliance with the State Board of Regents' regulations. The Board of Education has the authority to hire by granting a provisional teaching license and to fire those who fail to fulfill additional licensure requirements within five years.

# New York City Board of Education Licensure Requirements

- Beginning in 1991, the School Board required a prospective teacher candidate to possess a State certificate in addition to meeting the City-imposed licensure requirements.
- The New York City Board of Education is an employer in that it has the statutory authority to examine a teacher candidate with a State certificate using its own employment examinations, to grant a lifetime teaching license to those who pass, and to make the hire / no hire employment decision for those who are licensed.

## The New York State Board of Regents is not an employer

- Based on my contemporaneous experience as an employment testing consultant to the Transportation Security Administration (TSA), it is my opinion that the State of New York Board of Regents' legislated authority to require a State certificate for prospective teachers is indistinguishable from Congresses' legislated authority to TSA requiring U.S. citizenship for prospective airport security screeners.

# Not an employer . . .

- Both the State certificate and U.S. citizenship requirement are legislatively-imposed employment prerequisites. Obtaining the New York State certificate requires demonstrating minimal competence on a test. Obtaining U.S. citizenship also requires demonstrating minimal competence on a test.
- Administering legislatively-imposed employment prerequisites no more makes the Board of Regents an employer for aspiring teachers in New York State than it makes the Immigration and Naturalization Service an employer for aspiring Transportation Security Screeners nationwide. Simply stated, the New York State Board of Regents is not an employer.

# Board of Regents' Burden: Meeting “Professional Standards”

- The contract between the Board of Regents and National Evaluation Systems, Inc. requires that the LAST be developed “in conformance with appropriate professional standards” without referencing Federal obligations under Title VII of the Civil Rights Act of 1964.
- I am of the opinion that Board of Regents is not subject to the Uniform Guidelines on Employee Selection Procedures (Uniform Guidelines) which define an employer's burden when employment decisions adversely affect classes covered by Title VII of the Civil Rights Act of 1964.

- My opinion on the inapplicability of the Uniform Guidelines to the Board of Regent's use of the LAST is based on the following: 1) the Board of Regents is not an “employer or employment agency” since certification does not fall within “practices” covered by the Uniform Guidelines; and 2) the Uniform Guidelines do not make the case that state licensing and certification tests fall under their scope.

## SED's Methodology Supported by National Research Council

- My conclusion of the appropriateness of the Board of Regents' job analysis is supported by the National Research Council's Testing Teacher Candidates: The Role of Licensure Tests in Improving Teacher Quality.

# National Research Council

- “The Standards for Educational and Psychological Testing (American Educational Research Association et al., 1999)...say that ‘validity refers to the degree to which evidence and theory support the interpretation of test scores entailed by proposed uses of tests’ (p.9) and that the primary purpose of licensure testing is ‘to ensure that those licensed possess knowledge and skills in sufficient degree to perform important occupational activities safely and effectively’ (p.156).

# National Research Council

- “The standards explain that the type of evidence needed to establish a test’s validity is a matter of professional judgment: ‘Professional judgment guides decisions regarding the specific forms of evidence that can best support the intended interpretation and use’ of test scores (p.11).” “The 1999 standards note that at the present time validity research on licensure tests focuses ‘mainly on content-related evidence, often in the form of judgments that the test adequately represents the content domain of the occupation’ (p.57).”

# National Research Council

- “Typically, validity evidence for employment and credentialing tests includes a clear definition of the occupation or specialty, a clear and defensible delineation of the nature and requirements of the job, and expert judgments on the fit between test content and the job’s requirements. Procedurally, test sponsors conduct job analyses to define occupations and develop test specifications (blueprints) for licensure tests.”

# National Research Council

- “These are studies of the knowledge, skills, abilities, and dispositions needed to perform job duties and tasks. Studies of content relevance are then conducted to determine whether the knowledge and skills examined by the tests are relevant to the job and are represented in the test specifications. These data are generally obtained by having subject matter experts rate items on how well they reflect the test specifications, testing objectives, and responsibilities of the job. ...”

# National Research Council

- “Typically, sensitivity reviews are conducted to determine if irrelevant characteristics of test questions or test forms are likely to provide unfair advantages or disadvantages to particular groups of test takers. ... These sensitivity reviews also rely on expert judgment and are designed to remove potentially offensive materials from test forms” (p71-72)

# Uniform Guidelines do allow for sampling “necessary prerequisite” knowledge

- Under the technical standards for content validity section, the Uniform Guidelines do permit sampling knowledge such as that sampled on the NTE and the LAST which “is a necessary prerequisite to performance of critical or important work behavior(s):
- “In the case of a selection procedure measuring a knowledge, the knowledge being measured should be operationally defined as that body of learned information which is used in and is a necessary prerequisite for observable aspects of work behavior of the job” (emphasis added).

# National Research Council (2001):

- Teacher credentialing examinations test knowledge and skills identified by panels of educators as critical for entry into the profession
- “As is the case with the licensing tests used in other professions, teacher licensure tests focus on the knowledge and skills identified by panels of educators as critical for entry into the profession. They cover the material considered to be minimally necessary for beginning teaching.”

# National Research Council

- “Teacher licensure tests are not designed to distinguish moderately qualified teachers from highly qualified teachers. They are not constructed to predict the degree of teaching success a beginning teacher will demonstrate. The tests focus on the knowledge and skills necessary for competent beginning teaching, not on advanced levels or on the full set of knowledge and skills that licensing boards might like candidates to have. Thus, initial licensing tests are designed to discriminate at the score level that separates minimally competent from incompetent beginning practitioners. They do not possess the kinds of measurement qualities needed to make distinctions at the higher performance levels.”

National Research Council has articulated professional consensus regarding validation of teacher credentialing tests:

- “(A) comprehensive plan for gathering logical and empirical evidence for validation should specify the types of evidence that will be gathered (e.g., content-related evidence, data on the test’s relationships to other relevant measures of candidate knowledge and skill, and data on the extent to which the test distinguishes between minimally competent and incompetent candidates), priorities for the additional evidence needed, designs for data collection, the process for disseminating results, and a time line;

# National Research Council

- “the validation plan should include a focus on the fairness of the assessment for candidates and on disparate impacts for major candidate population groups; the plan should specify examination of the initial and eventual passing rates; major stakeholders should have input into the validation plan, and assessment experts should review the plan against current professional standards;

# National Research Council

- “the plan should require periodic review of accumulated validity evidence by external reviewers and appropriate follow-up; evidence should be provided regarding implementation of the validity plan including results of studies undertaken to collect validity evidence and gather fairness data; the validity evidence should be reported as studies are conducted; & assessment experts should review the results of the validity studies against current professional standards.”

# BEHAVIORAL SCIENCE IN COURT

- American Bar Association, Code of Professional Responsibility
- The duty of a lawyer, both to his client and to the legal system, is to represent his client zealously within the bounds of the law.  
(Ethical Canon 7-1)

# BEHAVIORAL SCIENCE IN COURT

- “The bounds of the law in a given case are often difficult to ascertain. The language of legislative enactments and judicial opinions may be uncertain as applied to varying factual situations. The limits and specific meaning of apparently relevant law may be made doubtful by changing or developing constitutional interpretations, inadequately expressed statutes or judicial opinions, and changing public and judicial attitudes. Certainty of law ranges from well-settled rules through areas of conflicting authority to areas without precedent.” (Ethical Canon 7-2)

# BEHAVIORAL SCIENCE IN COURT

- “Where bounds of law are uncertain, the action of a lawyer may depend on whether he is serving as an advocate or an advisor. A lawyer may serve simultaneously as both advocate and advisor, but the two roles are essentially different. In asserting a position on behalf of his client, an advocate for the most part deals with past conduct and must take the facts as he finds them... While serving as an advocate, a lawyer should resolve in favor of his client doubts as to the bounds of the law.” (Ethical Canon 7-3)

# BEHAVIORAL SCIENCE IN COURT

- “The advocate may urge any permissible construction of the law favorable to his client, without regard to his professional opinion as to the likelihood that the construction will ultimately prevail. His conduct is within the bounds of the law, and therefore permissible, if the position taken is supported by the law or supportable by a good faith argument for an extensive modification, or reversal of the law. However, a lawyer is not justified in asserting a position in litigation that is frivolous.” (Ethical Canon 7-4)

# BEHAVIORAL SCIENCE IN COURT

- “The lawyer... is not an umpire, but an advocate. He is under no duty to refrain from making every proper argument in support of any legal point because he is not convinced of its inherent soundness... His personal belief in the soundness of his cause or of the authorities supporting it is irrelevant.” (American Bar Opinion 280, 1949)

# Role of Psychologist as Behavioral Scientist

## (Ethical Standards of Psychologists)

- Objective in full disclosure of data
- Uncompromising in seeking broad principles governing human behavior
- Deductive in thought processes seeking to gain knowledge and explanations derived from broad generalizations
- Motivated by the need to explain and predict
- Impatient with Yes/No, Right/Wrong answers
- Has difficulty in making unequivocal statements on findings given professional constraints of judgment, caution and modesty

# Role of Lawyer as Advocate

- Subjective in choice of most favorable precedents to further client's interests
- Willfully biased in advocating most favorable “theory of law” to further interests of client
- Inductive in his thinking in dealing with court precedents
- Motivated by desire to win (gain favorable decision for client)
- Demands Yes/No, Right/Wrong answers
- Impatient with tentative nature of behavioral research and takes advantage of differing opinions to advantage of client.