The **Uniform Guidelines and Personnel Selection: Identify and Fix the Right Problem**

RICHARD F. TONOWSKI  
_U.S. Equal Employment Opportunity Commission and University of Maryland University College_

McDaniel, Kepes, and Banks (2011) argue that the _Uniform Guidelines on Employee Selection Procedures_ (Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, and Department of Justice, 1978) “have substantial influence” on personnel selection but are “inaccurate and inconsistent with professional practice” as summarized in the _Standards for Educational and Psychological Testing_ (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999) and _Principles for the Validation and Use of Personnel Selection Procedures_ (SIOP, 2003). Without a doubt, the phrasing in the 1978 _Uniform Guidelines_ is dated (and inscrutable in places) and reflects the state of the art more than 3 decades ago. But surely things such as the current way to discuss what was formerly called content validity cannot have given rise to the use of personnel practices unsupported by scientific evidence, the gerrymandering of personnel selection practices, and a general disregard of the ethics of such practices that the authors cite. McDaniel et al. seem to have in mind a problem far weightier. Indeed, there is a problem here. Although the authors have done us a great service by engendering discussion on what has been detrimental to the science and practice of personnel selection, revising or rescinding the _Uniform Guidelines_ is an insubstantial proxy for action on the real issue. I explore the status of the _Uniform Guidelines_ with respect to regulation, influence, legal enforcement, responsibility for updating, and agendas of stakeholders. I elaborate on comments of the focal article authors to clarify the nature of the problem and propose ways to fix it.

As background, I note that the _Uniform Guidelines_ are the last in a series of testing development and use guidelines produced by the EEOC and other federal agencies between 1966 and 1978. “Uniform” indicates that the 1978 _Uniform Guidelines_ resolved a situation where there were both EEOC guidelines and federal “executive agency” guidelines regarding testing. My experience has been primarily with EEOC practices.

What the Problem Is and Is Not

_What regulations?_ The authors repeatedly reference the _Uniform Guidelines_ and its predecessors as federal regulations. Apart from some record-keeping provisions, the
Uniform Guidelines are not regulations; they do not have the force of law. They are a declaration of the agencies’ enforcement policy, never proposed or subjected to review as regulation. With their call for the start of formal rule-making procedure, the authors would seem to involve the federal government in regulating selection testing where it has no role currently.

What influence? Perhaps, then, the Uniform Guidelines have a stranglehold on the profession because the courts are enthralled by them. True, the U.S. Supreme Court (Griggs v. Duke Power, 1971) accorded “great deference” to a predecessor of the Uniform Guidelines. But the historical reasons for that have waned and so has the influence of such pronouncements. The Uniform Guidelines are entitled to deference but not obedience, declared a federal appellate court when the Uniform Guidelines were still fresh (Guardians Association of New York City Police Department Inc v. Civil Service, 1980). More recently (Stagi v. AMTRAK, 2010), a federal appellate court explained that the Uniform Guidelines, as an agency pronouncement not coming from regulatory authority, get no more than Skidmore v. Swift & Co. (1944) deference, and that deference is only the power to persuade when the agency knows what it is talking about. A U.S. Supreme Court Justice parenthetically commented that agency influence based on persuasion does not involve deference (Kasten v. Saint-Gobain Performance Plastics Corp., 2011; Scalia, dissenting).

What enforcement? The Uniform Guidelines are definitely enforcement policy, and so perhaps the signatory agencies have been twisting the profession’s arms through selective enforcement, despite lack of regulatory authority and judicial support. As of this writing, the EEOC does not have even one substantive suit involving a formal test; its last suit involving a formal test (developed by nonpsychologists) was successfully concluded in 2006. (Suits involving use of applicant background information such as credit history, highly subjective selection practices, or subpoena enforcement to get validity information about a test are separate issues.) Of close to 100,000 private-sector discrimination charges received in 2010, perhaps 10 will involve a formal test. This does not imply that enforcement action will ensue from any of them. Selection testing is a tiny piece of the agency’s workload, and it commands interest and resources accordingly. The Department of Justice brings high-profile testing suits against state and municipal agencies, but very few. The Office of Personnel Management provides test use guidance to federal agencies on selection rather than engaging in enforcement; its online guidance follows, and directly references, the guidance on the SIOP Web site. The mission of the Office of Federal Contract Compliance Programs (OFCCP) arguably is the promotion of affirmative action for specified groups covered by law or executive order as a condition of doing business with the federal government. Discriminatory selection practices are an obvious barrier to that goal, but the goal goes beyond just doing valid selection and may not be dependent on the Uniform Guidelines.

Whose responsibility? If the agencies have not actively used the Uniform Guidelines to hamstring the profession, then perhaps theirs is a sin of omission. The authors discuss the “unfulfilled promises of the Uniform Guidelines,” apparently meaning that the federal agencies have the responsibility to keep up with the latest developments in the field and publish guidance accordingly. Agencies such as EEOC are not shy about providing guidance about things where they have authority, responsibility, and expertise, as casual perusal of EEOC’s Web site will confirm. Why should we care about what agencies having no regulatory authority or essential interest regarding the technical aspects of test validation did not do? Analogy with federal agencies whose purpose is closely bound up to science and the regulation of its application fails with the Uniform Guidelines signatory agencies. It seems that it is not enough that the Uniform Guidelines, as the focal article authors
themselves note, gave free rein to test users to interpret the Uniform Guidelines in the light of current professional standards, standard textbooks, and journals in the field of personnel selection.

Whose agenda? The McDaniel et al. list of unaddressed issues includes situational specificity, validity based on content similarity, construct validity, “types” of validity, validity generalization, transportability, differential validity and differential prediction, “false assumptions regarding adverse impact,” and “the diversity–validity dilemma.” The problem with the Uniform Guidelines is not so much that they prohibit practices that the focal article authors favor but that they are silent. So, as with adverse impact and the “diversity–validity dilemma,” the Uniform Guidelines “implicitly” deny the authors’ interests. There is a logical problem of inferring a position from silence, compounded by what the Uniform Guidelines themselves said about their interpretation.

One example that the focal article authors bring up is differential validity. We are told that differential validity does not exist and that the Uniform Guidelines “have not been revised to be consistent with current knowledge.” Yet one source of current knowledge, using a meta-analysis of historical cognitive ability test results, “calls into question claims by previous researchers that differential validity does not exist” (Berry, Clark, & McClure, 2011).

The authors also state that the Uniform Guidelines embrace the situational specificity hypothesis and have an emphasis on local validation studies. This appears to be an inference from silence. The history of situational specificity with respect to cognitive tests is recounted by the authors. Not mentioned is situational specificity “embraced” more recently in the research literature with personality tests (Tett & Christiansen, 2007). And although it is certainly true that an enforcement agency such as EEOC will demand validity evidence when investigating a test implicated in discrimination, in my personal experience the inference that this demands local validation evidence is simply untrue.

My point is not to debate the authors on specific assertions. Debate tends to produce advocacy for just one side of an issue. We evaluate the available evidence, reach our sincere conclusions, and ignore what does not fit the conclusions. This would be fine if other professionals did not come to different conclusions after following the same process. Getting wording in an authoritative document so that it actually means something but still reflects consensus and not just the views of a few can be a challenge. And that would be true dealing just with industrial and organizational (I–O) psychologists.

Consider the resolution effort with stakeholder differences regarding the Uniform Guidelines. It is not just a discussion among I–O psychologists. Besides the views of the various stakeholders regarding selection and EEO law, there is the political agenda of the federal executive branch to consider. As noted above, there may also be a lack of uniform views among the signatory agencies. The resulting document will likely be as much a political statement as a technical statement. It will also, of necessity, be a legal statement. The pen that finalizes its wording will be wielded by a lawyer not a psychologist.

The heart of the matter. Between 1966 and 1978 the federal agencies needed to get involved in technical selection issues because case law was rapidly evolving and professional guidance for test users was not up to what the agencies saw as their enforcement mission. Those days are past. The onus for providing professional guidance is back where it belongs, on the profession.

Although federal agencies do not regulate selection testing, EEO statute and case law demand that tests be demonstrated to work for legitimate purposes when the tests are implicated in alleged discrimination. Despite the introduction of the Principles and multiple revisions to the Principles and Standards since 1978, we still lack the certainty of generally accepted professional
practice. This cannot come from government fiat; it has to be the way that the profession actually practices. There may be several reasons why this stability has not been achieved. Ours is a dynamic field, and yesterday’s conventional wisdom may not hold today. Honest disagreement on the implications of research can lead to a lack of consensus on generally accepted practice and can water down the wording of authoritative professional documents, further leading to divergent practices. Test validation is not a paint-by-numbers exercise, and sound professional judgment is essential; the problem may be that our science is not yet mature and so there is too much judgment and too few standards. But the real problem is, I think, one which the profession has created and maintained for itself.

McDaniel et al. point to federal employees, human resources consultants, and labor lawyers, saying that they “can manufacture uncertainty about scientific findings,” and one can pay “a ‘scientist’ to testify to almost anything.” Somehow they missed the worst offenders of all: I–O psychologists. Who, if not they, are the enablers of endless litigation? Plaintiffs can find an appropriately credentialed expert to testify that whatever the test developer did could have been done some other way and so is fatally flawed. Employers can counter with their own experts who will testify that however shoddy the selection procedure, it is good enough. These experts hardly limit themselves to the Uniform Guidelines in advocating for their respective clients; fixing the Uniform Guidelines does not fix the problem. Abuse of expert testimony is hardly confined to our profession. But its net effect, I fear, is that we have gone a long way to convincing important stakeholders (such as judges, enforcement officials, and civil rights advocates) that we really do not know what we are talking about. That, in turn, means that potential test users must act defensively. Having any selection procedure open to challenge at any time by applicants, and having test users avoid responsibility with contrived argument or stonewalling on production of validity evidence, surely does not help either science or practice.

Fix the Problem

There may be a fix available to those with faith that good practice eventually drives out bad practice and the fortitude to press on until good practice is eventually reached.

Just do it. As the focal article authors rightly state, “SIOP’s inaction is counter to its mission.” SIOP and other professional organizations need to move ahead with establishing generally accepted professional practice. They do not need the federal government to do this. This implies more than updating the Standards and Principles every few years. Take on the issues that the authors say are suffering from silence. Come up with standard practice that represents the profession and satisfies the legitimate expectations of other stakeholders. The Uniform Guidelines will recede into history.

Science is science. Law is law. We operate at the intersection of the two, but they are distinct. For instance, validity may generalize, but enforcement is likely to be situational specific. Besides general issues of validity, an enforcement investigation will look to what psychologists term as construct contamination and construct deficiency in the specific way the test is used. Botched test administration may not be seen as reflecting on the test itself, but it will be of concern for enforcement if the error adversely impacted a protected class. Being clear on what type of issue is under discussion keeps people from talking past each other.

Engage the agencies. Moving ahead on what is within our expertise does not mean ignoring the federal agencies with respect to their legal enforcement role. McDaniel et al. find that the Uniform Guidelines do not address the legality of unspecified proposed solutions to the “diversity–validity dilemma.” At EEOC, the Office of Legal Counsel is charged with providing legal interpretation both to
internal staff and to employers. If there is a real issue as to how a practice relates to the agency’s enforcement, someone should ask for a private letter or a more formal determination. But we need discrete, substantive issues.

Find common ground. EEOC and OFCCP are busy with cases involving selection mechanisms that are not formal tests. Arguably both economic efficiency and EEO would be better served by more effective selection instruments. This may be particularly the case with small employers who, as the authors indicate, do not have resources to develop formal tests. Presenting stakeholders with opportunity to take on common concerns can lead to results beyond what anyone could accomplish by themselves.

Engage the courts. The Federal Judicial Center provides the federal judiciary with a wealth of information on scientific and technical matters that judges might encounter at trial. The Handbook of Scientific Evidence, scheduled to be released in its third edition in 2011, is a valuable resource. But neither it nor other materials from the center addresses selection. This is something that SIOP should pursue.

Celebrate good practice. Things that work need to be disseminated. This includes practical material presented at conferences that may not be suitable for research journals. Things that do not work (such as unsuccessful expert testimony in court) should be analyzed for lessons to be learned. SIOP has had good legal commentators for years; expand their coverage. Consider doing what enforcement agencies cannot: Provide third-party review for major test development projects. Above all, make it normative that I–O psychologists advocate primarily for sound professional practice, not just for whoever is paying them.

Persevere. Government agencies will back away from addressing the interface of technical issues and law because the issues are unfamiliar or they think they are being set up for someone’s agenda. Judges will translate misunderstood matters of fact into ill-founded matters of law. Those who do not understand what I–O psychology is about will malign the profession. We keep going with our efforts to improve professional practice and get it accepted. In the end, we prevail—at least until new issues make us do it all over again.

References


Guardians Association of New York City Police Department Inc. v. Civil Service (1980, 2nd Cir.). 630 F. 2d 79.


