



Good Science—Good Practice

Marcus W. Dickson
Wayne State University



Jamie Madigan
Ameren Services

In last issue's column, we discussed the SIOP symposium entitled "Unproctored Internet Testing: What Do the Data Say?" and one of the presentations we highlighted was the one by Lahti and Dekoikkoek on "ROI for Proctored Versus Unproctored Assessment Programs: Estimates From Multiple Utility Models and Identification of Moderators." We received a very kind e-mail from **Ken Lahti** thanking us for mentioning their work but emphasizing some of the findings that we did not. Specifically, Ken said:

You correctly noted our finding that, because most UIT programs omit cognitive ability tests and the omission of cognitive ability usually results in lower overall validity, the resulting decrement in program validity often results in lower program ROI/utility for UIT versus proctored testing (i.e., validity matters!). However, I thought the most interesting finding was actually the moderating effect of program size/scale: The efficiency ROI gains in high-volume UIT programs can actually overcome the ROI lost from assumed validity deficiencies. We found this for the high-volume selection program we modeled (5,000 hires/year) even when we assumed only \$15 per candidate savings for UIT versus onsite testing (likely a very conservative underestimate of actual savings). I think this finding is noteworthy because it shows real tradeoffs between science and practice issues very concretely (i.e., in \$) and could stimulate additional interesting dialogue about how to balance such issues for maximum organizational effectiveness.

We appreciated Ken's note and the additional emphasis on the important findings coming out of their work. Other authors whose work we mention, feel free to follow up as well!

We also received e-mail from **Neil Christiansen**, recommending that we focus on a recent field experiment conducted by Marianne Bertrand and Sendhil Mullainathan (2003), economists at the University of Chicago and MIT. They conducted a field experiment, entitled "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," in order to assess racial discrimination in the labor market. Following the lead of earlier research, they randomly assigned names that

seemed likely to belong to either African-American applicants or White applicants to fictitious résumés in response to job ads placed in two major metropolitan areas. Their study was much larger than prior work in this area (having sent out approximately 5,000 application letters in response to about 1,300 ads), and their design allowed them to assign all names to each of several different résumés (with different levels of skills, experience, education, etc.) multiple times, thus allowing better isolation of the effects of racial perceptions compared to other factors. The results were pretty stark: White-sounding names received a callback for (on average) 1 out of 10 letters submitted, but African-American-sounding names received a callback for (on average) 1 out of 15 letters submitted. In addition, employers who explicitly mention “equal opportunity employer” in their job ads showed no difference in the callback rates from those not listing EEO. In short, the authors summarize their findings by saying that “Based on our estimates, a White name yields as many more callbacks as an additional 8 years of experience.” Clearly, there are significant liability issues for organizations for whom it can be demonstrated that résumés of equivalent quality but different “sounding” names are treated differently. There are also significant issues in terms of lost access to quality human capital if some viable candidates are not considered due to perceptions of their race or ethnicity.

The recent article “Team Mental Models and Team Performance: A Field Study of the Effects of Team Mental Model Similarity and Accuracy” by **Beng-Chong Lim** and **Katherine Klein** (2006) is an excellent example of taking constructs originally examined in a laboratory setting and examining them in actual organizational settings. Lim and Klein focused on the question of whether shared team mental models had an impact on the performance of actual combat teams in the Singapore Armed Forces. The teams that were studied were existing combat teams, trained over a period of 2 years and functioning as a team for that time period, and assessed using standard Singapore Armed Forces assessment exercises (in this case, a 1-day combat circuit exercise, in a jungle environment). Lim and Klein’s analyses showed that when team members have similar mental models about the tasks to be performed, and about the nature of teamwork, the teams perform more effectively. Teams with accurate shared models about the tasks to be performed, and about the nature of teamwork, also performed more effectively. Potential implications of these findings are that investments of time and resources in ensuring that team members are “on the same page” about projects and about how to work as a team are likely to yield results. I (Marcus) remember very well the time we spent arguing with clients back in my consulting firm days over our recommendations for more time to be spent in planning a project, with clients pushing to reduce planning and coordination time and to “just do the work.” Lim and Klein provide data suggesting that preparation and coordination time—time spent developing shared mental models—is likely time well spent.

As a former co-PI on the GLOBE Project, I (Marcus) have heard many times the comment that GLOBE was a significant accomplishment but that it is hard to really know what a manager should do as a result of the data presented in the GLOBE book (House et al., 2004), or other publications relating to GLOBE. In the recently renamed *Academy of Management Perspectives*, Mansour Javidan, **Peter Dorfman**, Mary Sully de Luque, and Bob House begin to translate some of the GLOBE findings into action recommendations for American managers finding themselves in each of several different cultural settings. Their paper, “In the Eye of the Beholder: Cross Cultural Lessons in Leadership From Project GLOBE,” begins with an overview of some of the well-known findings from the project, including GLOBE’s conceptualization of dimensions of culture, the six second-order factors of leadership styles, and the 10 country clusters that were identified in the project. They then use the idea of a hypothetical American manager finding himself or herself in different countries from different culture clusters. Aspects of the culture relevant to the work environment are described, along with a section called “When in Brazil...” (or France, Egypt, or China), which include specific recommendations for managers on things to focus on in order to be effective in that setting. Of course, in an article-length manuscript, it isn’t possible for the authors to go into great depth on any one culture, and they are only able to address a few cultures, but this is an excellent example of using the wealth of data available in GLOBE for addressing practical issues managers face.

Martínez-Tur, Peiró, and Ramos (2005) recently looked at factors influencing customer satisfaction in service sector organizations. They were specifically interested in the balance between social constraints to providing good customer service (e.g., poorly trained employees, conflict between employees), and technical constraints to providing good customer service (e.g., lack of financial resources, lack of space leading to overcrowding). They gathered constraint data from managers of health and fitness facilities in Spain, and general and facet satisfaction data from customers of those facilities. Overall, they found that the two types of constraints each accounted for significant and independent variance and that social constraints had a larger unique contribution than did technical constraints. In other words, throwing money at resolving technical issues won’t resolve all of the customer service issues, and throwing training or selection systems at the people involved to resolve the social issues won’t resolve all of the customer service issues, either. Of course, it is possible that technical constraints like overcrowding of health center facilities, or social constraints like poorly trained staff, play an even larger role than shown in this study because those most dissatisfied with those issues may have moved their memberships to other facilities and thus were unavailable for inclusion in the study’s sample. Nonetheless, it is important to highlight the unique ways in which both social and technical constraints can diminish customer satisfaction.

Finally, Roth, Bobko, and Switzer recently published an article in *Journal of Applied Psychology* that illustrates how practices can sometimes drive research instead of the other way around. The authors model the behavior of the “4/5ths Rule” for determining the presence of adverse impact in a selection system, but they do so using a variety of computer simulations in both hypothetical and realistic situations. For those of you in need of a primer, the 4/5ths rule, whose origin it turns out is more indeterminable than you might guess, is a relatively simple rule of thumb that says that a selection system creates adverse impact if a protected class’s selection ratio is less than 80% (i.e., four fifths) of the selection ratio for the most often selected class. This procedure is unfettered by complex statistical significance tests and thus preferred by courts and government agencies who don’t want to require such specialized knowledge of key decision makers when it comes to evaluating adverse impact claims.

But driven by a need to provide a better (i.e., more scientific) answer to underlying questions of whether or not systematic discrimination exists, Roth, Bobko, and Switzer contrived data sets where real group differences in test scores did and did not exist. They then examined the performance of the 4/5ths rule in terms of signaling the presence of adverse impact and looked at what happened when you added more rigorous statistical tests of group differences. The results showed that the 4/5ths rule resulted in many false positive results (i.e., signaling adverse impact where none really existed), particularly with small sample sizes. Including statistical tests eliminated most of these mistakes.

Many have argued that this issue is an important one because the 4/5ths rule is lacking when compared to more rigorous tests that better accommodate the statistical properties of the data and are more appropriate for the constructs being considered. In other words, the 4/5ths test is not the best test to answer the research question of whether or not adverse impact exists. Some others may argue that this is a moot point because validity is an acceptable defense against adverse impact, and we should only use tests that are valid in the first place. The authors correctly point out, however, that the simple presence of adverse impact, even if it’s the result of a false positive, can trigger a variety of expensive and distracting problems—lawsuits, audits, grievances, and decisions to look for alternative tests with less (possibly perceived) adverse impact. So the performance of the 4/5ths rule is not a trivial issue, either to academics or practitioners.

As always, we hope to hear from you with recommendations for articles that advance theory and science, and which have direct implications for practice as well. We can be reached at marcus.dickson@wayne.edu or at HMadigan@ameren.com.

References

- Bertrand, M., & Mullainathan, S. (2003). *Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination*. MIT Department of Economics Working Paper No. 03-22. Available online at SSRN: <http://ssrn.com/abstract=422902>.
- House R. J., Paul J. Hanges P. J., Javidan M., & Dorfman P. W., Gupta V. (Eds.). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage.
- Javidan, M., Dorfman, P. W., De Luque, M. S., & House, R. J. (2006). In the eye of the beholder: Cross cultural lessons in leadership from Project GLOBE. *Academy of Management Perspectives*, 20, 67–90.
- Lahti, K. Personal communication. July 10, 2006.
- Lim, B-C., & Klein, K. J. (2006). Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy. *Journal of Organizational Behavior*, 27, 403–418.
- Martinez-Tur, V., Peiró, J. M., & Ramos, J. (2005). Linking situational constraints to customer satisfaction in a service environment. *Applied Psychology: An International Review*, 54, 25–36.
- Roth, P., Bobko, P., & Switzer III, F. (2006). Modeling the behavior of the 4/5ths rule for determining adverse impact: Reasons for caution. *Journal of Applied Psychology*, 91, 507–522.

Need to contact SIOP? Here's how!

SIOP Administrative Office
P.O. Box 87
520 Ordway Ave.
Bowling Green, OH 43402-0087
Phone: 419-353-0032 FAX 419-352-2645
e-mail: SIOP@siop.org Web: www.SIOP.org

Email Addresses:

David Nershi-Executive Director

dnershi@siop.org

**Jamie Jackson-Manager of
Membership, CE, and Sales**

jjackson@siop.org

Larry Nader-IT Manager

lnader@siop.org

Graduate Training Program

gtp@siop.org

Book Orders

Books@siop.org

SIOP Foundation
SIOPFoundation@siop.org

**Graphic Design and
Ad Submissions**

Graphics@siop.org

Webmaster

Webmaster@siop.org