

Editor:

Morris, Haroun, and Naimark (*J Am Acad Psychiatry Law* 32:231–45, 2004) surveyed forensic psychologists and psychiatrists and failed to find consensus in judgments of competency to stand trial based on a presented vignette. Essentially equal numbers of subjects found the same defendant to be competent as found the defendant to be incompetent. The authors characterized this finding as “not a mere fluke of the sample” and offered some probability calculations to suggest that the likelihood of this result is as low as 1 in 10 quadrillion (a result that requires certain assumptions about an expected higher base rate of examiner agreement). Apparently regarding these results as an embarrassment to forensic expertise, the authors concluded that “the defendant’s fate depends only on who performs the evaluation.”

Before zealous attorneys seize upon this research report as a cross-examination tool, it should be noted that a far simpler and arguably more compelling hypothesis was available. The authors never demonstrated that the brief vignettes employed as research stimuli were psychometrically adequate to the task of eliciting and accurately measuring the subjects’ forensic decision-making. A psychometrically inadequate vignette could effectively reduce the subjects to mere guessing, and random guessing would also produce the observed pattern of equally divided judgments.

Thomas E. Schacht, PsyD
 Professor, Department of Psychiatry and
 Behavioral Sciences
 James H. Quillen College of Medicine
 East Tennessee State University
 Johnson City, TN

Reply

Editor:

We hope that you will consider publishing this letter as our response to Dr. Schacht’s¹ letter and the commentary by Dr. Leong² published on our article, “Assessing Competency Competently: Toward a Rational Standard for Competency-to-Stand-Trial As-

sessments” (*J Am Acad Psychiatry Law* 32:231–45, 2004). We thank both Drs. Schacht and Leong for their thoughtful responses, but we respectfully disagree with their analyses and conclusions.

Whatever the methodological limitations of our study (which will be addressed herein), the responses of the 273 forensic psychiatrists and psychologists who participated in our study clearly indicated confusion about the meaning of the different tests of competency to stand trial. Some respondents were unaware that the federal statutory test was construed by the Supreme Court to be the *Dusky* test. Some differentiated among the three tests; some did not. Some decided competency on the basis of mental disorder alone or on treatment considerations that are irrelevant to the determination of competency. In addition, some courts distinguished the rational manner test from the rational understanding test; some did not. The existing confusion must be addressed. We cannot simply accept Dr. Leong’s assertion that there are no rational reasons for changing the competency-to-stand-trial standard. The standard to measure competency must be understood and applied consistently by courts and by the forensic psychiatrists and psychologists who offer testimony on the matter. Clearly, that is not happening now.

From a methodological perspective, one would like to have evidence for both interrater and intrarater reliability before trying to prove that a construct is valid. In a sense, our study is a reliability study, because we asked (albeit obliquely) for respondents to indicate whether they thought the three standards of competence were identical or dissimilar. We predicted that subjects would see the differences among the standards and would apply and make judgments about vignettes based on those differences. We found that there was poor agreement on the meaning of those standards and poor agreement about how to apply them to different fact situations, which suggests that they are unreliable constructs. Furthermore, predictive validity was assessed in our study, in that we asked respondents to use the three standards to make judgments about the two vignettes, and we predicted how they would apply the standards. Essentially, respondents did a poor job (although better on the second vignette than on the first). This suggests that, at least for the vignettes we posed, there is no predictive validity when using these federal and state legal definitions, because they