Larrabee, G.J., & Levin, H.S. (1986). Memory self-ratings and objective test performance in a normal elderly sample. *Journal of Clinical and Experimental Neuropsychology*, *8*, 275–284.

Larrabee, G.J., Millis, S.R., & Meyers, J.E. (2008). Sensitivity to brain dysfunction of the Halstead-Reitan vs. an ability-focused neuropsychological battery. *The Clinical Neuropsychologist*, *22*, 813–825.

Larrabee, G.J., Millis, S.R., & Meyers, J.E. (2009). 40 plus or minus 10, a new magical number: Reply to Russell. *The Clinical Neuropsychologist*, *23*, 746–753.

McCrea, M., Iverson, G.L., McAllister, T.W., Hammeke, T.A., Powell, M.R., Barr, W.B., & Kelly, J.P. (2009). An integrated review of recovery after mild traumatic brain injury (MTBI): Implications for clinical management. *The Clinical Neuropsychologist*, *23*, 1368–1390.

Merten, T., Bossink, L., & Schmand, B. (2007). On the limits of effort testing: Symptom validity tests and severity of neurocognitive symptoms in nonlitigant patients. *Journal of Clinical and Experimental Neuropsychology*, *29*, 308–318.

Meyers, J.E., & Volbrecht, M.E. (2003). A validation of multiple malingering detection methods in a large clinical sample. *Archives of Clinical Neuropsychology*, *18*, 261–276.

Mittenberg, W., Azrin, R., Millsaps, C., & Heilbronner, R. (1993). Identification of malingered head injury on the Wechsler Memory Scale-Revised. *Psychological Assessment*, *5*, 34–40.

Mittenberg, W., Rotholc, A., Russell, E., & Heilbronner, R. (1996). Identification of malingered head injury on the Halstead-Reitan Battery. *Archives of Clinical Neuropsychology*, *11*, 271–281.

Morgan, J.E., & Sweet, J.J. (Eds.), (2009) *Neuropsychology of malingering casebook*. New York: Psychology Press.

Rees, L.M., Tombaugh, T.N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology*, *16*, 501–506.

Rogers, R. (1988). Researching dissimulation. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 309–327). New York: Guilford Press.

Rohling, M.L., Larrabee, G.J., Greiffenstein, M.F., Ben-Porath, Y.S., Lees-Haley, P., Green, P., & Greve, K.W. (2011). A misleading review of response bias: Comment on McGrath, Mitchell, Kim, and Hough (2010). *Psychological Bulletin*, *137*, 708–712.

Ross, S.R., Millis, S.R., Krukowski, R.A., Putnam, S.H., & Adams, K.M. (2004). Detecting probable malingering on the MMPI-2: An examination of the Fake-Bad Scale in mild head injury. *Journal of Clinical and Experimental Neuropsychology*, *26*, 115–124.

Slick, D.J., Sherman, E.M.S., & Iverson, G.L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, *13*, 545–561.

Straus, S.E., Richardson, W.S., Glasziou, P., & Haynes, R.B. (2005). *Evidence-based medicine: How to practice and teach EBM* (3rd ed.). New York: Elsevier Churchill Livingstone.

Suhr, J.A., & Gunstad, J. (2005). Further exploration of the effect of "diagnosis threat" on cognitive performance in individuals with mild head injury. *Journal of the International Neuropsychological Society*, *11*, 23–29.

Tombaugh, T.N. (1996). *TOMM. Test of Memory Malingering*. New York: Multi-Health Systems.

Vickery, C.D., Berry, D.T.R., Inman, T.H., Harris, M.J., & Orey, S.A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology*, *16*, 45–73.

Victor, T.L., Boone, K.B., Serpa, J.G., Buehler, J., & Ziegler, E.A. (2009). Interpreting the meaning of multiple symptom validity test failure. *The Clinical Neuropsychologist*, *23*, 297–313.

Williams, J.M., Little, M.M., Scates, S., & Blockman, N. (1987). Memory complaints and abilities among depressed older adults. *Journal of Consulting and Clinical Psychology*, *55*, 595–598.

doi:10.1017/S1355617712000392

DIALOGUE RESPONSE

Response to Bigler

Glenn J. Larrabee

Bigler (this issue) and I apparently are in agreement about the importance of symptom validity testing, and my recommendation to adopt a new terminology of "performance validity" to address the validity of performance on measures of ability, and "symptom validity" to address the validity of symptom report on measures such as the MMPI-2. We appear to differ on issues related to false positives and the rigor of performance and symptom validity research designs.

The study by Locke, Smigielski, Powell, and Stevens (2008) is cited by Bigler as demonstrating potential false positive errors due to TOMM scores falling in a "near miss" zone just below cutoff. This interpretation suggests a continuum of performance. Review of Bigler's Figure 1 and Locke et al.'s Table 2 shows that the frequency distribution of TOMM scores does not, however, reflect a continuum but shows two discrete distributions: (1) a sample of 68 ranging from 45 to 50 (mean = 49.31, $SD = 1.16$) and (2) a sample of 19 ranging from 22 to 44 (mean = 35.11, $SD = 6.55$) [note Bigler interprets two distributions below 45, but the sample size is too small to establish this presence]. Clearly,

Locke et al. *did not* view TOMM failures as false positives in their sample. Although Locke et al. found that performance on neurocognitive testing was significantly lower in this group, TOMM failure was not related to severity of brain injury, depression or anxiety; only disability status predicted TOMM failure. They concluded: "This study suggests that reduced effort occurs outside forensic settings, is related to neuropsychometric performance, and urges further research into effort across various settings" (p. 273).

As previously noted in my primary review, several factors minimize the significance of false positive errors. First, scores reflecting invalid performance are atypical in pattern or degree for bona fide neurological disorder. Second, cutoff scores are typically set to keep false positive errors at or below 10%. Third, investigators are encouraged to specify the characteristics of bona fide clinical patients who fail PVTs representing "false positives," to enhance the clinical use of the PVT in the individual case. Fourth, appropriate use of PVTs in the individual case requires the presence of multiple abnormal scores on independent PVTs, occurring in the context of external incentive, with no compelling neurologic, psychiatric or developmental explanation for PVT failure, before one can conclude the presence of malingering (cf., Slick, Sherman, & Iverson, 1999).

Bigler also criticizes the research in this area as being, at best, Class III level research (American Academy of Neurology, AAN, Edlund, Gronseth, So, & Franklin, 2004), noting the research is typically retrospective, using samples of convenience, with study authors not blind to group assignment. Review of the AAN guidelines, however, shows that retrospective investigations using case control designs can meet Class II standards (p. 20). Moreover, there is no requirement for masked or independent assessment, if the reference standards for presence of disorder and the diagnostic tests are *objective* (italics added). The majority of studies cited in recent reviews (Boone, 2007; Larrabee, 2007; Morgan & Sweet, 2009) follow case control designs contrasting either non-injured simulators or criterion/known-groups of definite or probable malingerers, classified using objective test criteria from Slick et al. (1999), with groups of clinical patients with significant neurologic disorder (usually moderate/severe TBI) and/or psychiatric disorder (i.e., major depressive disorder). As such, these investigations would meet AAN Level II criteria.

In my earlier review in this dialog, I described a high degree of reproducibility of results in performance and symptom validity research. Additionally, the effect sizes generated by this research are uniformly large, for example, $d = -1.34$ for Reliable Digit Span (Jasinski, Berry, Shandera, & Clark, 2011); $d = .96$ for MMPI-2 FBS (Nelson, Sweet, & Demakis, 2006), replicated at $d = .95$ incorporating 43 new studies (Nelson, Hoelzle, Sweet, Arbisi, & Demakis, 2010); $d = 2.02$ for the two-alternative forced choice Digit Memory Test (Vickery, Berry, Inman, Harris, & Orey, 2001). These effect sizes exceed those reported for several psychological and medical tests (Meyer et al., 2001). Effect sizes of this magnitude are striking, considering that the discrimination is between feigned performance and legitimate neuropsychological abnormalities, rather than between feigned performance and normal performance. Reproducible results and large effect sizes cannot occur without rigorous experimental design.

REFERENCES

Bigler, E.D. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society*, *18*, 632–642.

Boone, K.B. (2007). *Assessment of feigned cognitive impairment. A neuropsychological perspective*. New York: Guilford.

Edlund, W., Gronseth, G., So, Y., & Franklin, G. (2004). *Clinical practice guideline process manual: For the Quality Standards Subcommittee (QSS) and the Therapeutics and Technology Assessment Subcommittee (TTA)*. St. Paul: American Academy of Neurology.

Jasinski, L.J., Berry, D.T.R., Shandera, A.L., & Clark, J.A. (2011). Use of the Wechsler Adult Intelligence Scale Digit Span subtest for malingering detection: A meta-analytic review. *Journal of Clinical and Experimental Neuropsychology*, *33*, 300–314.

Larrabee, G.J. (2007) (Ed.), *Assessment of malingered neuropsychological deficits*. New York: Oxford.

Locke, D.E.C., Smigielski, J.S., Powell, M.R., & Stevens, S.R. (2008). Effort issues in post-acute outpatient acquired brain injury rehabilitation seekers. *Neurorehabilitation*, *23*, 273–281.

Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.R., … Reed, G.M. (2001). Psychological testing and psychological assessment. A review of evidence and issues. *American Psychologist*, *56*, 128–165.

Morgan, J.E., & Sweet, J.J. (Eds.) (2009). *Neuropsychology of malingering casebook*. New York: Psychology Press.

Nelson, N.W., Hoelzle, J.B., Sweet, J.J., Arbisi, P.A., & Demakis, G.J. (2010). Updated meta-analysis of the MMPI-2 Symptom Validity Scale (FBS): Verified utility in forensic practice. *The Clinical Neuropsychologist*, *24*, 701–724.

Nelson, N.W., Sweet, J.J., & Demakis, G.J. (2006). Meta-analysis of the MMPI-2 Fake Bad Scale: Utility in forensic practice. *The Clinical Neuropsychologist*, *20*, 39–58.

Slick, D.J., Sherman, E.M.S., & Iverson, G.L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, *13*, 545–561.

Vickery, C.D., Berry, D.T.R., Inman, T.H., Harris, M.J., & Orey, S.A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology*, *16*, 45–73.