

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/236219757>

Symptom Validity Testing: Unresolved Questions, Future Directions

Article in *Journal of Clinical and Experimental Psychopathology & Quarterly Review of Psychiatry and Neurology* · April 2013

DOI: 10.5127/jep.028312

CITATIONS

9

READS

468

3 authors, including:



April D Thames

University of Southern California

77 PUBLICATIONS 1,619 CITATIONS

[SEE PROFILE](#)



Ashley Watts

Emory University

69 PUBLICATIONS 1,176 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Psychopathy and Attitudes Toward Rape [View project](#)



Psychopathy and Patient SM [View project](#)



Symptom Validity Testing: Unresolved Questions, Future Directions

Scott O. Lilienfeld, Ph.D.^a, April D. Thames, Ph.D.^b, Ashley L. Watts, B.A.^a

^a Emory University

^b University of California, Los Angeles

Abstract

As the stimulating articles in this Special Issue demonstrate, symptom validity tests (SVTs) are alive and well in neuropsychology and allied fields. At the same time, a number of key unresolved issues regarding the construct validity and clinical utility of SVTs remain. In this commentary, we address six largely unanswered questions regarding SVTs: (1) Do SVTs possess clinical validity?; (2) Is malingering taxonic?; (3) Is there an overarching dimension of malingering and low effort?; (4) How should we combine information from different SVTs?; (5) Can the assessment of psychopathy supplement information from SVTs?; and (6) How do ethnicity and culture affect the interpretation of SVTs? We conclude that SVTs play an indispensable role in the detection of aberrant response sets in neuropsychology, although the precise meaning of scores on many SVTs requires clarification.

© Copyright 2012 Textrum Ltd. All rights reserved.

Keywords: neuropsychology, malingering, validity, psychopathy

Correspondence to: Scott O. Lilienfeld, Department of Psychology, Room 440, Emory University, 36 Eagle Row, Atlanta, Georgia 30322. Email: slilien@emory.edu

1. Department of Psychology, Room 440, Emory University, 36 Eagle Row, Atlanta, Georgia 30322.

2. Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, 760 Westwood Plaza C8-226, Los Angeles, CA 90095.

Received 10-Apr-2012; accepted 17-Apr-2012

Table of Contents

- Introduction
- Do SVTs Possess Clinical Validity?
- Is Malingering Taxonic?
- Is There an Overarching Dimension of Malingering and Low Effort?
- How Should We Combine Information From Different SVTs?
- Can the Assessment of Psychopathy Supplement Information from SVTs?
- How Do Ethnicity and Culture Affect the Interpretation of SVTs?
- Concluding Thoughts
- References

Introduction

As this Special Issue of the *Journal of Experimental Psychopathology* amply attests, symptom validity tests (SVTs) are alive and well in neuropsychology and allied domains, including the assessment of psychopathology. The lively, rich, and diverse set of papers in this Special Issue illustrate the clinical relevance of SVTs for detecting malingering, low effort, and other problematic response sets, and underscore several key unresolved issues in this domain.

Several cross-cutting themes emerge across the papers in this Special Issue. First, there has been increasing attention to symptom validity tests (SVTs) in neuropsychology, sparking a number of attempts to refine the construct validity of available diagnostic tools. SVTs stemmed from medico-legal evaluations of individuals reporting fallacious claims of disability, a use that later expanded to “effort” assessments in standard clinical neuropsychological evaluations. This Special Issue addresses the validity of SVT testing, particularly when assessing different types of malingering (cognitive vs. psychological), as well as challenges to their routine clinical use. In particular, instrument transparency and the potential low base rate of malingering in many samples (but see “Is Malingering Taxonic?” for a discussion of potential problems with the base rate concept in this literature) pose formidable methodological obstacles to appraising the validity of SVTs.

Indeed, SVTs are hardly without scientific controversy (for a different view, see Sweet & Guidotti Breting, in press). Nor have all important questions regarding their construct validity and clinical utility been addressed. In the interests of stimulating further debate and research that could help to place SVTs on even firmer scientific footing, in this concluding commentary we pose six key questions regarding the scientific status of SVTs. We believe that each of these questions points to a largely unresolved issue concerning SVTs that in turn offers fruitful directions for scientific investigation.

Do SVTs Possess Clinical Validity?

As we have seen, the “v” in SVTs stands for validity. The assumption underpinning SVTs (see Reynolds, 1998) is that administering them in conjunction with well-established neuropsychological indices will enhance the validity of the latter measures by detecting malingering and other aberrant response sets; some SVTs are embedded measures of effort within standard neuropsychological tests. Yet a recent and widely discussed article by McGrath, Mitchell, Kim, and Hough (2010; but see Rohling et al., 2011, for a response) called this longstanding assumption into question. Throwing down the gauntlet to neuropsychologists who use SVTs in their clinical practice and research, McGrath et al. argued that SVTs in neuropsychology (as well as in psychopathology) have, with few potential exceptions, yet to demonstrate clinical utility.

To justify their routine clinical use, McGrath et al. contended, neuropsychologists must demonstrate that SVTs enhance the convergent validity of neuropsychological measures, either as moderator variables, suppressor variables, or both. An SVT operating as a moderator variable would boost validity by identifying a subgroup of respondents (namely, those with high scores on the SVT) for whom a neuropsychological measure is significantly less valid than for other respondents; an SVT operating as a suppressor variable would boost validity by increasing the correlation between a neuropsychological measure and clinically relevant indicators (e.g., independently

established brain damage) when it is controlled statistically. The challenge posed by McGrath et al. is by no means limited to assessment in neuropsychology or psychopathology. In an article provocatively entitled “On the invalidity of validity scales,” Piedmont, McGrae, Riemann, and Angleitner (2000) concluded that widely used measures of response sets in personality research, such as standard indicators of negative or positive impression management, rarely enhance validity as either moderators or suppressors (but see Edens & Ruiz, 2006, for a different view).

It may be tempting to dismiss the critique of McGrath et al. (2010) on the grounds that their inclusion criteria for the validity of SVTs were overly restrictive (see Rohling et al., 2011). For example, McGrath et al. placed virtually exclusive emphasis on the enhancement of validity by means of moderation and suppression, according short shrift to voluminous evidence for the validity of SVTs in detecting malingering offered by simulation and known-groups designs. This criticism was voiced at a recent symposium at the American Psychology-Law Society (Ben-Porath, 2012), which featured a vigorous defense of SVTs in neuropsychology and psychopathology. To some extent, we are sympathetic to this concern.

Nevertheless, we regard McGrath et al.’s pointed challenge as well taken and believe that it needs to be answered straightforwardly by SVT proponents. After all, SVTs lose their meaning if the “v” in their name no longer refers to validity. Moreover, we agree with Merten and Merckelbach (in press) that simulation designs, although a prerequisite for demonstrating the validity of SVTs for malingering detection, provide an upper-bound - and almost always an overly sanguine - estimate of their validity. Neuropsychologists must be careful not to commit the logical error of affirming the consequent: The fact that A entails B does not imply that B entails A. Extending this logical error to the present body of literature, the fact that individuals asked to simulate malingering display statistically and clinically significant elevations on SVTs does not necessarily imply that individuals who exhibit statistically and clinically significant elevations on SVTs are malingering. Conversely, only a subset of malingerers produce below-chance response patterns (Biachini, Mathias, & Greve, 2001; Merten & Merckelbach, in press), so these measures surely yield many false-negatives as well.

Is Malingering Taxonic?

Several of the articles in this Special Issue refer to such terms as “base rate,” “prevalence,” “false positives,” and the like, in the context of malingering. In their helpful introduction to this Special Issue, Merten and Merckelbach observe that estimates of non-authentic responding, including malingering, range as high as 50% of claimants in certain settings. Yet if malingering is underpinned by a latent dimension rather than a taxon, such terms as base rate, false positive, and false negative – and more important, the concepts that underlie them – effectively lose their meaning, because they presume a categorical distinction between malingering and non-malingering.

A *taxon* is natural class or category, one that exists in the real world rather than merely in the minds of clinicians (Meehl, 2004; Meehl & Golden, 1982). Accumulating evidence from taxometric studies, which rely on sophisticated statistical procedures to ascertain whether an observed distribution is underpinned by a dimension (continuum) or a taxon, suggests that malingering is often or usually dimensional in nature (Frazier, Youngstrom, Naugle, Haggerty, & Busch, 2007; Walters, Berry, Rogers, Payne, & Granacher, 2009). Of course, there is no question that malingering *can* sometimes be taxonic, as when a subgroup of respondents in an experiment attempt consciously to simulate impairment. But the extant data suggest that such taxonicity may be rare or even nonexistent in many actual neuropsychological samples.

This dimensional view of malingering accords with clinical theorizing and recent research, which raise questions concerning whether malingering differs categorically from other forms of symptom endorsement, such as those observed in somatoform and factitious disorders (Delis & Wetter, 2007). Over 90 years ago, Rosanoff (1920) maintained that the difference between the patient with unexplained somatic complaints and the malingerer is one of degree. This conjecture gains support from an important study by Merckelbach, Jelicic, and Peters (2011), who found that participants asked to malingering cognitive and psychiatric symptoms on a well-validated measure, the Structured Interview of Malingered Symptomatology (SIMS; Merckelbach & Smith, 2003; Widows & Smith, 2005), continued to endorse more of these symptoms on a re-take of the SIMS approximately one hour later, *even though they were asked to respond honestly*. These intriguing results raise the possibility that the distinction between

conscious feigning and unconscious symptom reporting is blurrier than widely assumed. As Merckelbach et al. observed, in some cases “feigning may evolve into a less conscious form of symptom reporting” (p. 136).

If malingering is typically dimensional rather than taxonic, clinical psychologists, including neuropsychologists, may need to modify their traditional view of SVTs (see also Faust, 2003). Specifically, such dimensionality introduces the distinct possibility that many patients classified as malingerers by SVTs may be at least partly unaware of their intentions. Clinical psychologists may also need to come to view malingering as falling on a continuum of conscious intentionality of symptom endorsement, with malingering ostensibly anchoring one pole of this dimension and somatoform disorders anchoring the other.

Is There an Overarching Dimension of Malingering and Low Effort?

Many clinical psychologists, including neuropsychologists, assume that various forms of feigning are underpinned by a single higher-order dimension that cuts across measures in both neuropsychology and psychopathology (such as the MMPI-2 F scale). Some authors also include low effort in this broad category. Low effort refers to any instance of less than maximal exertion during a psychological evaluation (see Cronbach, 1960, for the distinction between maximal and typical performance). Certain psychiatric diagnoses, such as major depression, somatoform disorders, and some personality disorders, may increase the risk of sub-optimal effort, and such diminished effort should not be confused with malingering (Strauss, Sherman, & Spreen, 2006).

Yet as Merten and Merckelbach (this issue) note, research evidence offers at best weak support for the existence of a higher-order dimension of feigning. Even within neuropsychological assessment, the associations among SVTs tend to be modest (Merten, Bossink, & Schmand, 2007; but see Axelrod & Schutte, 2011). Furthermore, the correlations between SVTs in neuropsychology and SVTs embedded within psychopathology indices are generally even lower. For example, Haggerty, Frazier, Busch, and Naugle (2007) found only modest or negligible correlations between scores on the Victoria Symptom Validity Test (VSVT) and measures of feigning derived from the Personality Assessment Inventory (Morey, 2007; see also Greiffenstein, Gola, & Baker, 1995). The correlations between forced-choice SVTs for feigning cognitive impairment and SVTs designed to detect the feigning of psychosis are also low, as individuals who feign psychiatric malingering do not always underperform on cognitive malingering measures. In comparing measures such as the SIMS with the Medical Symptom Validity Test (MSVT) – a widely used memory test with a built-in effort measure – Dandachi-FitzGerald and Merckelbach (in press) report that the SIMS surpassed the MSVT in correctly identifying experimental malingerers, suggesting that memory-oriented SVTs, such as the MSVT, may be useful for detecting feigned memory deficits, but not other types of feigned psychopathology. In addition, a factor analysis by Nelson, Sweet, Berry, Bryant, and Granacher (2007) revealed that various feigning measures, including neuropsychological SVTs and validity indicators derived from the MMPI-2, loaded on four separable factors: underreporting of neurotic symptoms, overreporting of neurotic symptoms, low cognitive effort, and overreporting of psychotic or rarely endorsed symptoms, suggesting that these measures may not detect the same latent construct. Although Nelson et al. did not test the fit of a higher-order factor, it is worth noting that only one of their six factor intercorrelations (namely, that between overreporting of neurotic symptoms and over reporting of psychotic or rarely endorsed symptoms) were positive, and that several other factor intercorrelations were negligible.

These findings should serve as powerful reminder that different indices of malingering and low effort are not fungible, either within or across neuropsychological and psychopathological domains. Making matters more complicated, the extant neuropsychological research offers little guidance to practitioners concerning which of these indices possess higher construct validity for different clinical purposes. Although neuropsychologists generally agree that using multiple SVTs is preferable to using only one (Boone, 2007; Iverson & Franzen, 1996; Orey, Cragar, & Berry, 2000; Vickery et al., 2004; Victor, Boone, Serpa, Buehler, & Ziegler, 2009), the current literature offers scant information concerning how best to combine information derived from different SVTs. As Faust (2003) cautioned, we should only use data derived from different malingering measures when they exhibit incremental validity (Sechrest, 1963) above and beyond each other (see also “How Should We Combine Information from Different SVTs?”).

On the positive side, the low correlations among SVTs are ideal from the perspective of incremental validity, because nonredundant predictors can account for more variance in outcomes in multiple regression equations. It remains to be seen, however, whether clinicians can capitalize effectively on this nonshared variance in their judgments of malingering.

In retrospect, the low correlations among SVTs should not be terribly surprising. The oft-forgotten but still pertinent person-situation debate of the 1960s and 1970s (see Kenrick & Funder, 1988) taught us that *t data* (Block, 1977; also see Mischel, 1968), which are isolated “test” data typically derived from laboratory assessments, frequently display surprisingly low intercorrelations, even when ostensibly assessing the same construct. To a large extent, that is because *t data* tend to possess a high component of situational specificity (Epstein, 1979). As a consequence, relatively little of their variance assesses the construct of interest, in this case response sets. As Mischel’s (1974) classic findings on delay of gratification demonstrated (see also Block, 1977), even seemingly trivial differences in task demands often yield marked changes in the correlations among putatively similar measures (see also Kagan, 2012).

How Should We Combine Information From Different SVTs?

A widespread assumption in much of clinical psychology, including neuropsychology, is that more information is always superior to less information. As Faust (2003) pointed out,

“When attempting to identify malingering (or perform most any other diagnostic task), clinicians are commonly advised to combine or integrate all of the available data. They are also often urged to be particularly attentive to patterns or configurations in the data, such as whether the obtained results fit with a known pattern of disorder. This notion of combining all of the data is so frequently articulated that it has become almost like breathing, a component of our clinical consciousness that is so ensconced it draws minimal thought or evaluation; rather, it is a taken-for-granted and seemingly self-evident truth” (p. 108).

Yet the history of assessment teaches us that this presumption is often false (Faust, 2003; Garb, Wood, Lilienfeld, & Nezworski, 2005; Lilienfeld, Wood, & Garb, 2006). Sawyer’s (1966) classic review demonstrated that the addition of interview information to other psychological data often results in *decreases* in the validity of clinical judgments and predictions. Similarly, Whitehead (1985) found that adding the MMPI to the Rorschach resulted in significant and substantial increases in clinicians’ diagnostic accuracy, whereas adding the Rorschach to the MMPI resulted in slight (but nonsignificant) decreases in accuracy.

Research on the *dilution effect* (Nisbett, Zukier, & Lemley, 1981) may help to explain why. This research shows that presenting participants with accurate, but nondiagnostic, information regarding individuals often results in less accurate judgments, especially when this information is salient or attention-grabbing (e.g., unusual responses to a projective test). Compared with participants presented with relevant information only, participants presented with both relevant and irrelevant information often attend too heavily to the irrelevant information, resulting in a decrement in their judgments.

In this context, the frequently overlooked distinction between *statistical* incremental validity and *clinical* incremental validity is crucial. The former refers to the extent to which the addition of a new measure to extant information enhances validity when entered into a statistical formula, such as a multiple regression equation; the latter refers to the extent to which the addition of a new measure to extant information enhances the validity of clinical judgments, predictions, or both. Statistical incremental validity cannot be negative; at worst, a new predictor will merely be redundant with existing information and drop out of a predictive equation. In contrast, clinical incremental validity can be negative: clinicians’ judgments and predictions can become worse with new information, especially if this information is salient but of negligible or nonexistent validity.

Therefore, research is needed on how best to combine information from different SVTs to maximize the validity of clinical predictions of malingering and other response sets, and to ensure that these predictions are not worsened by the addition of new SVT information.

Can the Assessment of Psychopathy Supplement Information from SVTs?

Traditionally, neuropsychologists rely mostly or exclusively on SVTs to evaluate the likelihood of malingering, low effort, and allied response sets. They have accorded little or no attention to the possibility that psychopathy – a condition marked by a constellation of such attributes as superficial charm, egocentricity, guiltlessness, callousness, risk-taking, dishonesty, manipulativeness, and poor impulse control (Cleckley, 1941/1988; Hare, 1991/2003; Lilienfeld, 1994) – may play a helpful role in malingering detection in conjunction with SVTs. This conjecture may be broadly consistent with longstanding findings that psychopathy, or at least the impulsive and reckless behaviors often associated with it, is positively associated with the reporting of somatic symptoms (Lilienfeld, 1992; Lilienfeld & Hess, 2001), which as we have seen may fall on a continuum with certain forms of malingering.

Admittedly, the link between psychopathy and malingering has not been entirely consistent across studies (e.g., see Gacono, Meloy, Sheppard, Speth, & Roske, 1995, for evidence suggesting an association between psychopathy and malingering, and Cima, Merckelbach, Hollnack, & Knauer, 2003, and Poythress, Edens, & Watkins, 2001, for the opposite conclusion). Yet several recent studies, especially those using well-validated measures (e.g., the Psychopathy Checklist-Revised; Hare, 1991/2003; the Psychopathic Personality Inventory; Lilienfeld & Andrews, 1996) that assess the core interpersonal and affective features of psychopathy (e.g., dishonesty, absence of empathy) rather than its nonspecific antisocial and criminal features, suggest that although psychopathy is not associated with malingering success, it is associated with a somewhat greater willingness to malingering (as well as greater perceived ability to malingering; Edens, Buffington, & Tomicic, 2000; Kucharski, Duncan, Egan, & Falkenbach, 2006).

Although it is clear that psychopathy and malingering are far from equivalent (Rogers & Cruise, 2000), systematic assessments of psychopathy, especially those that target the core interpersonal and affective characteristics of this condition, might nonetheless offer clinically useful incremental validity above and beyond traditional SVTs in the detection of malingering. For example, because individuals who display low effort on neuropsychological tests may be heterogeneous (a minority may possess genuine psychopathology or may underperform for reasons of low motivation; Bianchini et al., 2001), incorporating well-validated psychopathy measures into neuropsychological batteries could help to identify respondents whose low effort stems from largely intentional efforts at deception. We call for systematic research on this possibility.

How Do Ethnicity and Culture Affect the Interpretation of SVTs?

Although not discussed explicitly in this Special Issue, a critical consideration in SVT evaluations is the potential impact of ethnicity and culture on neuropsychological performance. Certain ethnic minority groups may perform more poorly on standard neuropsychological tests than majority groups, risking misclassification of cognitive impairment among the former groups (Byrd et al., 2006; Manly, 2008; Rivera-Mindt, Saez, & Byrd, 2010). Such ethnic differences may reflect slope bias in neuropsychological tests arising from misunderstanding of the test instructions, inappropriate item content, differential prior exposure to test material, and poor rapport between examiner and examinee (see Reynolds, 1998). Nevertheless, they may also (or alternatively) reflect substantive differences in educational quality, genuine discrepancies in certain cognitive aptitudes, and other non-artifactual influences.

There is a pressing need for cross-cultural research on SVTs. Some work on SVTs has been conducted on Spanish speakers in Spain; for example, Ramirez, Chirivella-Garrido, Caballero, Ferri-Campos, and Noe-Sebastian (2004) demonstrated specificity of the Test of Memory Malingering (TOMM) with elder Spaniards. Similarly, in another Spanish sample, Burton, Vilar-Lopez, and Puente (2012) demonstrated the utility of the TOMM and Rey Fifteen Item Test in differentiating among individuals involved in litigation forensic non-capital cases, forensic capital cases, and clinical controls, whereas Dot Counting did not differ significantly across groups. In particular, individuals involved in litigation-forensic cases were more likely to be classified as malingerers than those involved in forensic capital murder cases or clinical controls.

Various influences other than intentional feigning, such as poor motivation, frustration, fatigue, and boredom, can produce suboptimal effort. Some these influences may themselves reflect cultural factors. Hence, examiners must be cognizant of possible cultural influences bearing on low effort and be prepared to take steps to minimize them. Several of these influences may be remedied by enhancing rapport between examiner and examinee, providing encouragement (when appropriate), allowing examinees to take frequent breaks, offering water/snacks throughout the evaluation, and the like.

Concluding Thoughts

The interpretation of neuropsychological assessment data hinges critically on valid data. The papers in this Special Issue should assist the field in moving closer to that important goal. At the same time, McGrath et al.'s (2010) review raises significant concerns regarding the clinical utility of SVTs in neuropsychology and related fields, including the detection of psychopathology. In our view, the onus of proof continues to fall on SVT proponents to demonstrate that these measures routinely improve the validity of their assessments.

Moreover, the effort to develop SVTs or combinations of SVTs that are superior to extant measures at detecting malingering and allied response sets should continue. Increasing the number of non-transparent (low face-valid) SVTs in a given evaluation may be one fruitful approach. The VSVT, which we mentioned earlier (Slick, Hoop, & Strauss, 1995; Slick, Hoop, Strauss, & Spellacy, 1996), is one example of such a measure; this test superficially appears to increase in difficulty across succeeding trials, although it does not. The VSVT has been widely used in forensic neuropsychological examinations, and studies reveal that both experimentally instructed simulators and compensation-seeking patients show significantly increased error rates and response latencies, especially on the purportedly "difficult" items, relative to control and non-compensation-seeking patients (Slick et al., 1996). Furthermore, the VSVT demonstrates discriminant validity, as evidenced by low correlations ($r = .29$ or less) with traditional neuropsychological measures.

The lengthy history of response bias measurement in the personality and psychopathology literatures teaches us that simplistic dichotomies are unlikely to capture the true state of nature when it comes to inaccurate responding. For example, once widespread assertions that response set or style indices embedded within omnibus psychopathology measures (e.g., the MMPI) reflect either purely substantive variance or purely stylistic variance (e.g., Edwards, 1963) have wisely been abandoned (Paulhus, 1991). Instead, the field has gradually come to accept that such measures are typically a complex mix of substantive and stylistic variance, with the nature of that mix dependent on individual differences, contextual factors, and other variables. The state of affairs is likely to be similar for many SVTs, especially those designed to detect low effort. Such tests almost surely assess variance relevant to both genuine psychopathology and response sets, so future research would do well to focus on parsing the differential contributions of these influences.

These noteworthy complexities aside, there remains no dispute that some respondents engage in malingering and other aberrant response sets on neuropsychological measures, so SVTs will always play an indispensable role in neuropsychological assessment, especially within the forensic context (see also Bush et al., 2005). The articles in this Special Issue underscore the crucial point that without SVTs, even seasoned neuropsychologists can be fooled – and fooled badly. As the Nobel-prize winning physicist Richard Feynman (1974) reminded us, "Science is a way of trying not to fool yourself. The first principle is that you must not fool yourself, and you are the easiest person to fool" (p. 12).

References

- Axelrod, B. N., & Schutte, C. (2011). Concurrent validity of three forced-choice measures of symptom validity. *Applied Neuropsychology, 18*, 27-33. <http://dx.doi.org/10.1080/09084282.2010.523369>
- Ben-Porath, Y. (Chair). (2012, March 16th). *Validity scales moderate the validity of scores on substantive measures in forensic assessments*. Symposium held at the annual convention of the American Psychology-Law Society, San Juan, Puerto Rico.
- Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist, 15*, 19-45. <http://dx.doi.org/10.1076/clin.15.1.19.1907>

- Block, J. (1977). Advancing the psychology of personality: Paradigmatic shift or improving the quality of research? In D. Magnusson & N.S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology* (pp. 37-63). New York: John Wiley & Sons.
- Boone, K. (2007). Commentary on 'Cogniform disorder and cogniform condition: Proposed diagnoses for excessive cognitive symptoms' by Dean C. Delis and Spencer R. Wetter. *Archives of Clinical Neuropsychology*, 22, 675-679. <http://dx.doi.org/10.1016/j.acn.2007.07.005>
- Burton, V., Vilar-Lopez, R., & Puente, A. E. (2012). Measuring effort in neuropsychological evaluations of forensic cases of Spanish speakers. *Archives of Clinical Neuropsychology*, 22, 675-679.
- Bush, S. S., Ruff, R. M., Troster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H.,... Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity - NAN policy & planning committee. *Archives of Clinical Neuropsychology*, 20, 419-426. <http://dx.doi.org/10.1016/j.acn.2005.02.002>
- Byrd, D. A., Miller, S., Reilly, J., Weber, S., Wall, T. L., & Heaton, R. K. (2006). Early environmental factors, ethnicity, and adult cognitive test performance. *The Clinical Neuropsychologist*, 20, 243-260. <http://dx.doi.org/10.1080/13854040590947489>
- Cima, M., Merckelback, H., Hollnack, S., & Knauer, E. (2003). Characteristics of psychiatric prison inmates who claim amnesia. *Personality and Individual Differences*, 35, 373-380. [http://dx.doi.org/10.1016/S0191-8869\(02\)00199-X](http://dx.doi.org/10.1016/S0191-8869(02)00199-X)
- Cleckley, H. H. (1941/1988). *The mask of sanity: An attempt to reinterpret the so-called psychopathic personality*. St. Louis, MO.: Mosby.
- Cronbach, L. J. (1960). *Essentials of psychological testing* (2nd ed.). Oxford England: Harper.
- Dandachi-FitzGerald, B. & Merckelbach, H., (2013). Feigning ≠ feigning a memory deficit: the medical symptom validity test as an example. *Journal of Experimental Psychopathology*, 4, 46-63. <http://dx.doi.org/10.5127/jep.025511>
- Delis, D. C., & Wetter, S. R. (2007). Cogniform disorder and condition: Proposed diagnoses for excessive cognitive symptoms. *Archives of Clinical Neuropsychology*, 22, 589-604. <http://dx.doi.org/10.1016/j.acn.2007.04.001>
- Edens, J. F., Buffington, J. K., & Tomicic, T. L. (2000). An investigation of the relationship between psychopathic traits and malingering on the Psychopathic Personality Inventory. *Assessment*, 7, 281-296. <http://dx.doi.org/10.1177/107319110000700307>
- Edens, J., & Ruiz, M. (2006). On the validity of validity scales: The importance of defensive responding in the prediction of institutional misconduct. *Psychological Assessment*, 18, 220-224. <http://dx.doi.org/10.1037/1040-3590.18.2.220>
- Edwards, A.L. (1963). A factor analysis of experimental social desirability and response set scales. *Journal of Applied Psychology*, 47, 308-316. <http://dx.doi.org/10.1037/h0039793>
- Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, 37, 1097–1126. <http://dx.doi.org/10.1037/0022-3514.37.7.1097>
- Faust, D. (2003). Holistic thinking is not the whole story: Alternative or adjunct approaches for increasing the accuracy of legal evaluations. *Assessment*, 10, 428-441. <http://dx.doi.org/10.1177/1073191103259534>
- Feynman, R. P. (1974). Cargo cult science. *Engineering and Science*, 37, 10-13.
- Frazier, T. W., Youngstrom, E. A., Naugle, R. I., Haggerty, K. A., & Busch, R. M. (2007). The latent structure of cognitive symptom exaggeration on the Victoria Symptom Validity Test. *Archives of Clinical Neuropsychology*, 22, 197-211. <http://dx.doi.org/10.1016/j.acn.2006.12.007>
- Gacono, C. B., Meloy, J., Sheppard, K., Speth, E., & Roske, A. (1995). A clinical investigation of malingering and psychopathy in hospitalized insanity acquittees. *Bulletin of The American Academy of Psychiatry & The Law*, 23, 387-397.
- Garb, H. N., Wood, J. M., Lilienfeld, S. O., & Nezworski, M. (2005). Roots of the Rorschach controversy. *Clinical Psychology Review*, 25, 97-118. <http://dx.doi.org/10.1016/j.cpr.2004.09.002>
- Greiffenstein, M. F., Gola, T., & Baker, W. (1995). MMPI-2 validity scales versus domain specific measures in detection of factitious traumatic brain injury. *Clinical Neuropsychologist*, 9, 230-240. <http://dx.doi.org/10.1080/13854049508400485>
- Haggerty, K. A., Frazier, T. W., Busch, R. M., & Naugle, R. I. (2007). Relationships among Victoria symptom validity test indices and personality assessment inventory validity scales in a large clinical sample. *The Clinical Neuropsychologist*, 21, 917-928. <http://dx.doi.org/10.1080/13854040600899724>

- Hare, R. D. (1991/2003). *The Hare Psychopathy Checklist-Revised (PCL-R)*: 2nd edition technical manual. Multi-Health Systems, Toronto, Canada.
- Iverson, G. L., & Franzen, M. D. (1996). Using multiple objective memory procedures to detect simulated malingering. *Journal of Clinical and Experimental Neuropsychology*, *18*, 38-51. <http://dx.doi.org/10.1080/01688639608408260>
- Kagan, J. (2012). *Psychology's ghosts: The crisis in the profession and the way back*. New Haven, CN: Yale University Press.
- Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. *American Psychologist*, *43*, 23-34. <http://dx.doi.org/10.1037/0003-066X.43.1.23>
- Kucharski, L.T., Duncan, S., Egan, S. S., & Falkenbach, D. M. (2006). Psychopathy and malingering of psychiatric disorder in criminal defendants. *Behavioral Sciences and the Law*, *24*, 311-322.
- Lilienfeld, S. O. (1992). The association between antisocial personality and somatization disorders: A review and integration of theoretical models. *Clinical Psychology Review*, *12*, 641–662. [http://dx.doi.org/10.1016/0272-7358\(92\)90136-V](http://dx.doi.org/10.1016/0272-7358(92)90136-V)
- Lilienfeld, S. O., & Hess, T. H. (2001). Psychopathic personality traits and somatization: Sex differences and the mediating role of negative emotionality. *Journal of Psychopathology and Behavioral Assessment*, *23*, 11-24. <http://dx.doi.org/10.1023/A:1011035306061>
- Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review*, *14*, 17-38. [http://dx.doi.org/10.1016/0272-7358\(94\)90046-9](http://dx.doi.org/10.1016/0272-7358(94)90046-9)
- Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment*, *66*, 488-524. http://dx.doi.org/10.1207/s15327752jpa6603_3
- Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2006). Why questionable psychological tests remain popular. *The Scientific Review of Alternative Medicine*, *10*, 6-15.
- Manly, J. J. (2008). Critical issues in cultural neuropsychology: Profit from diversity. *Neuropsychology Review*, *18*, 179-183. <http://dx.doi.org/10.1007/s11065-008-9068-8>
- McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, *136*, 450-470. <http://dx.doi.org/10.1037/a0019216>
- Meehl, P. E. (2004). What's in a taxon? *Journal of Abnormal Psychology*, *113*, 39-43. <http://dx.doi.org/10.1037/0021-843X.113.1.39>
- Meehl, P. E., & Golden, R. (1982). Taxometric methods. In P. Kendall & J. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 127-181). New York: Wiley.
- Merckelbach, H., & Smith, G. P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology*, *18*, 145-152.
- Merckelbach, H., Jellicic, M., & Peters, M. (2011). The residual effect of feigning: how intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical & Experimental Neuropsychology*, *33*, 131-139. <http://dx.doi.org/10.1080/13803395.2010.495055>
- 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. <http://dx.doi.org/10.1080/13803390600693607>
- Merten, T., & Merckelbach, H., (2013). Forced-choice tests as single-case experiments in the differential diagnosis of intentional symptom distortion. *Journal of Experimental Psychopathology*, *4*, 20-37. <http://dx.doi.org/10.5127/jep.023711>
- Mischel, W. (1968). *Personality and assessment*. Hoboken, NJ US: John Wiley & Sons Inc.
- Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 249–292). San Diego, CA: Academic Press.
- Morey, L.C. (2007). *The Personality Assessment Inventory professional manual*. Lutz, FL: Psychological Assessment Resources.
- Nelson, N. W., Sweet, J. J., Berry, D. R., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of The International Neuropsychological Society*, *13*, 440-449. <http://dx.doi.org/10.1017/S1355617707070373>

- Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, *13*, 248-277. [http://dx.doi.org/10.1016/0010-0285\(81\)90010-4](http://dx.doi.org/10.1016/0010-0285(81)90010-4)
- Orey, S. A., Cragar, D. E., & Berry, D. R. (2000). The effects of two motivational manipulations on the neuropsychological performance of mildly head-injured college students. *Archives of Clinical Neuropsychology*, *15*, 335-348.
- Paulhus, D.L. (1991). Measurement and control of response bias. In J.P. Robinson & P.R. Shaver (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Piedmont, R. L., McCrae, R. R., Riemann, R., & Angleitner, A. (2000). On the invalidity of validity scales in volunteer samples: Evidence from self-reports and observer ratings in volunteer samples. *Journal of Personality and Social Psychology*, *78*, 582–593. <http://dx.doi.org/10.1037/0022-3514.78.3.582>
- Poythress, N. G., Edens, J. F., & Watkins, M. (2001). The relationship between psychopathic personality features and malingering symptoms of major mental illness. *Law And Human Behavior*, *25*, 567-582. <http://dx.doi.org/10.1023/A:1012702223004>
- Ramirez, R.M., Chirivella-Garrido, J., Caballero, M.C., Ferri-Campos, J., & Noe-Sebastian, E. (2004). Intelligence, memory and malingering: Correlation between scales. *Revista de Neurologia*, *38*, 28-33.
- Reynolds, C.R. (1998). Common sense, clinicians, and actuarialism. In C.R. Reynolds (Ed.). *Detection of malingering during head injury litigation* (pp. 261-286). New York: Plenum.
- Rivera-Mindt, M., Saez, P., & Byrd, D. (2010). Primary challenges in the neuropsychological evaluation of ethnic minority individuals: A brief report. *NAN Bulletin*, *25*, 1-5.
- Rogers, R., & Cruise, K. R. (2000). Malingering and deception among psychopaths. In C. B. Gacono (Ed.), *The clinical and forensic assessment of psychopathy: A practitioner's guide* (pp. 269-284). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.
- 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. *Psychological Bulletin*, *137*, 708-712. <http://dx.doi.org/10.1037/a0023327>
- Rosanoff, A. (Ed.). (1920). *Manual of psychiatry*. (5th ed.). Oxford England: John Wiley.
- Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin*, *66*, 178-200. <http://dx.doi.org/10.1037/h0023624>
- Sechrest, L. (1963). Incremental validity: A recommendation. *Educational and Psychological Measurement*, *23*, 153-158. <http://dx.doi.org/10.1177/001316446302300113>
- Slick, D.J., Hoop, G., & Strauss, E. (1995). *The Victoria Symptom Validity Test*. Odessa, FL: Psychological Assessment Resources.
- Slick, D.J., Hoop, G., Strauss, E., & Spellacy, F. (1996). Victoria Symptom Validity Test: Efficiency for detection of feigned memory impairment and relationship to neuropsychological tests and MMPI-2 validity scales. *Journal of Clinical and Experimental Psychology*, *18*, 911-922.
- Strauss, E., Sherman, E. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd. ed). New York: Oxford University Press.
- Sweet, J. J., & Guidotti Breting, L. M., (2013). Symptom validity test research: status and clinical implications. *Journal of Experimental Psychopathology*, *4*, 6-19. <http://dx.doi.org/10.5127/jep.022311>
- Vickery, C. D., Berry, D. R., Dearth, C. S., Vagnini, V. L., Baser, R. E., Cragar, D. E., & Orey, S. A. (2004). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology*, *19*, 37-48.
- Victor, T. L., Boone, K. B., Serpa, J., Buehler, J., & Ziegler, E. A. (2009). Interpreting the meaning of multiple symptom validity test failure. *The Clinical Neuropsychologist*, *23*, 297-313. <http://dx.doi.org/10.1080/13854040802232682>
- Walters, G. D., Berry, D. T., Rogers, R., Payne, J. W., & Granacher, R. P., Jr. (2009). Feigned neurocognitive deficit: Taxon or dimension? *Journal of Clinical and Neuropsychology*, 1-10.
- Whitehead, W. C. (1985). *Clinical decision making on the basis of Rorschach, MMPI, and automated MMPI report data*. Unpublished doctoral dissertation, University of Texas Health Science Center at Dallas.
- Widows, M.R., & Smith, G.P. (2005). *Structured Inventory of Malingered Symptomatology professional manual*. Odessa, FL: Psychological Assessment Resources.