

WPE WebPsychEmpiricist

Author's Update to "Too dumb to die"

R. K. McKinzey

Oakland, CA

2/20/08

Editor's Note: This update file comments upon events and research available after the publication of "Too dumb to die: Mental retardation meets the death penalty." This is the seventh update.

Abstract

Since the 2003 publication of "Too dumb to die: Mental retardation meets the death penalty" another relevant articles have become available and are reviewed. In the first update (9/20/03), the literature on the malingering formula for the WAIS-III is reviewed at more length. The second update (6/23/05) reviews a book on the diagnosis of MR, and summarizes two CA court opinions regarding Atkins. Specific issues include the SEM, part score, Flynn Effect, and adaptive behavior arguments. The third update (8/10/05) notes that a jury found Atkins not MR. The fourth update (9/8/05) corrects some false positive figures on page 2. The fifth update (9/16/06) notes that the VA Supreme Court has reversed the jury finding and ordered yet another jury hearing. The sixth update (6/5/07) reviewed another relevant CA Supreme Court decision. This update closes the Atkins case and reviews two new studies of faking with MR populations.

Author's Update to "Too dumb to die"
9/20/03

(Later updates follow)

In 1995, Wiley Mittenberg (Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995) published the experimental¹ validation of two formulas for the detection of malingering on the Wechsler Adult Intelligence Scale-Revised (WAIS-R). The formula using seven subtests (Seven Subtest Formula, or SSF) and a cutoff of >0 yielded 24% false negative rate² (FNR) and 18% false positive rate (FPR). The other formula used two subtests (Two Subtest Formula, or TSF) and a cutoff of >1.53 yielded a 27% FNR and 21% FPR. The TSF also added a probability table, which has advantages over cutoffs.

The formulas were cross-validated on the WAIS-R in 1998 (Millis, Ross, & Ricker, 1998). Using a differential prevalence design, the SSF yielded a 12% FNR and 14% FNR. Using a cutoff of >2 , the TSF yielded a 28% FNR and 21% FPR. The probability table was not considered. They specifically cautioned against use of either formula until the WAIS-R standardization sample was studied.

Both formulas were again studied in 1999 (Axelrod & Rawlings, 1999). Using patients with no reason to malingering ("true patients"), the SSF yielded a 10% FPR. The TSF yielded a 19% FPR. Since no malingerers were available, no FNR could be calculated.

However, the WAIS-R was replaced by the WAIS-III in 1997. Consequently, Mittenberg cross-validated the formulas (Mittenberg et al., 2001) using the WAIS-III. Using both experimental and differential prevalence designs, the SSF yielded a 17% FPR and 28% FNR for experimental malingerers and 56% FNR for suspected malingerers. The TSF yielded a 14% FPR and 44% FNR for experimental malingerers and 75% FNR for suspected malingerers. Furthermore, the authors note that, "Individuals who malingering intellectual impairment frequently do not simulate borderline mental retardation. Obvious exaggeration of this nature tends to produce relatively uniform global reductions in subtest scores that are better identified by methods other than pattern analysis." (page 444).

¹ Malingering studies use three different kinds of subject samples. One kind uses patients *suspected* of faking because they are either in litigation or have produced (in the opinion of the scientist) unbelievable results. This is called "differential prevalence." The second kind simply asks people to fake the test. This is variously known as "experimental," "analogue," or "simulation" design. The third design, concurrent validity, categorizes patients by using other malingering tests.

² For definitions of terms used in describing the accuracy of tests, see: <http://wpe.info/2x2table.pdf>

The formulas were again cross-validated on the WAIS-III (Greve, Bianchini, Mathias, Houston, & Crouch, 2003). Using a concurrent validity design (comparing the formulas to the outcome of the Portland Digit Recognition Test and/or the Test of Memory Malingering), the SSF yielded an 11% FPR and 42% FNR. The TSF yielded a 28% FPR and 59% FNR. The authors deemed the SSF results “acceptable,” with a positive outcome “associated with a diagnosis of malingering.” (page 256)

The TSF was studied (Iverson & Tulsky, 2003) on the WAIS-III standardization sample (N=2450 normals) and true patients (N=145). Using the validated cutoff of >2, the TSF yielded a 20% FPR in normals and 14% FPR in the true patient group.

NONE of these studies included a sample of people with Mental Retardation (MR). As Dr. Mittenberg observed, both formulas rely on wide variations in subtest scores. In people malingering MR, all WAIS-III subtests scores will be uniformly low, making the formulas useless. Until proven otherwise, I therefore deem both formulas useless in Atkins decisions. Given the TSF's 20% FPR, it would seem to be useless for any other condition as well. A decision on the SSF awaits its use in the WAIS-III standardization sample, since a positive score may turn out to be useful.

6/23/05

Judges facing Atkins cases are not the only ones who have to make life-changing decisions about whether or not a given person has MR. The Social Security Administration (SSA) faces the same decision routinely in deciding whether a claimant gets benefits. As a result, the SSA formed the Committee on Disability Determination for Mental Retardation to review the extant literature on operationalizing the diagnosis of MR—that is, which tests, which scores, and which cutoffs are to be used? The resulting review has been published (Reschly, Meyers, & Hartel, 2002), and the committee had similar findings to those I expressed in “Too dumb to die.”

The book's Introduction reviews the different and changing definitions of MR, noting one dates from 1500 BCE Egypt, and that some definitions end the developmental period at 18, while the SSA definition ends it at age 22. “It is important to note that the differences between the SSA definition of mental retardation and those of the major professional and health-related organizations derive from the purpose for which it is used. The SSA definition is used not for diagnostic purposes, but rather for purposes of program eligibility.” (p. 24)

Chapter Three reviews the concept and measurement of intelligence. The committee mentions three issues (two in passing) that will prove central in Atkins cases: The Flynn Effect, the SEM, and the use of part scores. The Flynn Effect (FE) refers to the easily seen fact that IQ tests have gotten harder over time (Flynn, 1983, 1984a, 1984b, 1985, 1987a, 1987b, 1990, 1998a, 1998b, 1998c, 1999, 2000; Flynn & Detterman, 1996; Flynn & Neisser, 1998) For example, (at age 35) getting 40 correct of Raven Standard Progressive Matrices item yields an IQ of 100 using the 1949 norms; using the 1984 norms, the same score gets an IQ of 80. Some psychologists will argue that the FE can be applied retroactively; that is, any old IQ score must be reduced by however much the new norms demand. As a result, goes the argument, a defendant who obtained a non-MR IQ score a decade ago can now be diagnosed as MR.

All tests have some statistical error. If a given test is given to a person twice, the exact score will vary to some degree, either more or less. This amount is termed the Standard Error of Measurement (SEM). Any test score should be considered the midpoint in the range of the SEM. That is, if the SEM of the WAIS-R FSIQ is 2.25 (in 25-34 year olds), there is a 95% chance of the true score being between +4.5 and -4.5. Citing this concern, the DSM-IV-TR recommends that IQs below 75 be considered MR.

The committee has a long discussion of the use of part scores in the determination of IQ. The WAIS variants sample intelligence using different kinds of tasks, such as defining vocabulary words, or solving a non-verbal puzzle. These tasks are then summed various ways to arrive at normed scores. For example, the verbal tasks are summed into the part score labeled Verbal IQ (VIQ) and the non-verbal tasks are summed into a part score labeled Performance IQ (PIQ). The VIQ and PIQ are then summed into a total score labeled the Full Scale IQ. The committee writes that the literature indicates that FSIQ is the best measure of *general* intelligence—and should be routinely used in the determination of MR. That is, total scores are preferable to part scores.

A footnote mentions that committee member Keith Widaman, Ph.D.³ dissented. He argued that, if any part score is under the cutoff, MR can be diagnosed, even if the other part (and total) scores are over the cutoff.

Chapter Four reviews adaptive behavior measures. “The presence of clinically significant maladaptive behavior *does not* meet the criterion of significant limitations in adaptive

³ <http://psychology.ucdavis.edu/faculty/pgms/page.cfm?PersonID=52>

functioning...refusal to perform a task that a person is capable of doing is also a reflection of problem behavior and should not be considered in relation to adaptive behavior” (p. 154; emphasis in original). The committee notes that the informant’s culture, biases, and expectation may change the results.

The committee confirmed that some adaptive functioning tests (specifically, the Vineland, AAMR Adaptive Behavior Scales, Comprehensive Test of Adaptive Behavior, Adaptive Behavior Inventory, Adaptive Behavior Evaluation Scale and Social Skills Rating Scales) lack any adult norms. The Comprehensive Test of Adaptive Behavior-R and Adaptive Behavior: Street Survival Skill Questionnaire have norms on adults with MR, but none on adults without MR. The Adaptive Behavior Assessment System (ABAS) and the Scales of Independent Behavior-R have adult norms on people without MR. The Independent Living Scales have gotten “generally negative” (p. 173) reviews.

When assessing the use of specific cutoffs, the committee reported that only the ABAS presented accuracy rates—and emphasizes the huge inaccuracy rates found therein. To avoid false negatives, they therefore argue for the use of a very liberal cutoff—one that makes 32% of the population eligible for the diagnosis of MR.

A recent California appellate Atkins decision (“Peo v Sup Ct of Tulare Co (Vidal, real party in interest),” 2005)⁴ considered the issues of adaptive functioning, the use of part scores, the Flynn Effect, and the SEM. Jorge Vidal is accused of kidnapping, torturing and killing Eric Jones on 1/24/01. The prosecutor is seeking the Death Penalty. Based on a psychological evaluation by Eugene Couture, Ph.D.⁵, Vidal sought an Atkins hearing.

Dr. Couture reported that Vidal (date of birth 10/23/69) failed 4th grade, had limited language skills, and forgot chores and household rules. Nevertheless, he married, had children, and became a field worker. The school records documented his limited academics and provided IQ testing. In 1980 (age 11), a WISC-R yielded VIQ of 59, PIQ of 109, and FSIQ of 81. A 1984 (age 15) WISC-R yielded a VIQ of 59 and PIQ of 126. A 1987 (age 17) WAIS-R yielded a VIQ of 77, PIQ of 119, and FSIQ of 92.

Dr. Couture did his own testing in 4/2003. A WAIS-R yielded a VIQ of 70, PIQ of 96, and FSIQ of 78. In his 4/10/03 report, he concluded that Vidal “clearly qualifies for a diagnosis

⁴ This case can be downloaded via: http://wpe.info/reprints_available.html

⁵ <http://www.neuroskills.com/index.shtml?main=/cns/couture.shtml>

of Borderline Intellectual Functioning, but not for a diagnosis of Mental Retardation.” He argued that the drop in PIQ scores was due to methamphetamine use and exposure to the manufacturing of it.

Dr. Couture returned for more testing in September 2003, this time doing a Wechsler Abbreviated Scale of Intelligence (WASI), which yielded a VIQ of 61, PIQ of 99, and FSIQ of 77. He also did a Vineland, using Vidal and his two sisters and ex-wife as informants. The tests were interpreted as showing impaired adaptive behavior, despite Vidal’s ex-wife’s telling a detective that Vidal never showed any signs of MR. The doctor now (report of 9/27/03) offered a diagnosis of MR.

I was called as rebuttal. I argued that CA’s new PC 1376 addressing Atkins (like many other similar laws) called for limitations in general IQ, not splinter skills. The best estimate of general IQ is the FSIQ, not one of the parts scores. Since Vidal’s FSIQ at age 17 was 92, he could not, under the SEM argument, be considered MR. Furthermore, even if the 1987 VIQ of 77 was considered useful, it is still outside the 2 SEM allowed by the various definitions.

I rejected the notion that the Flynn Effect allowed retrospective use of later norms, as doing so has never been shown to improve accuracy of diagnosis. Even if it were accepted, the 1987 FSIQ of 92 is still outside the MR range.

I argued that the Vineland scores were irrelevant, for reasons already covered. Furthermore, there is plenty of evidence in the records showing no impairments.

Dr. Widaman offered surrebuttal. He made the parts score argument that the various low VIQ scores were acceptable diagnostic indicators, since VIQ measures “crystallized” intelligence, and is more important for functioning. Since Vidal’s new adult scores were so low, a diagnosis of MR was acceptable. He pointed out that the Flynn Effect deducted 6 points, bringing the 2003 WAIS-R FSIQ of 78 down to 72, within the SEM of the MR range. He rejected the use of the WASI, but not the Vineland. He rejected the notion that Vidal might have malingered the 2003 testing, since the various test scores were consistent with each other and with the adaptive functioning reports. Furthermore, having some adaptive competencies, as I argued, was irrelevant, since only the presence of impairments was to be considered in the diagnosis. In describing impairments, only the behavior, not the ability⁶ is to be considered, since there is no way to measure lack of functioning due to choice. He noted that MR can “come and

⁶ Note that the rest of the SSA committee disagree with the doctor on this point.

go”. (p. 24). While he agreed that diagnostic disagreement was possible based on the 1987 scores, the current scores are conclusionary for MR.

The trial court agreed with the parts score, SEM, and Flynn arguments, finding that the 1987 VIQ of 77 was adequate for MR. The court also found the Vineland acceptable.

The appellate court disagreed with the parts score argument. The opinion notes the frequency with which the various definitions and PC1376 refer to general IQ. “General intellectual functioning is primarily determined by the defendant’s FSIQ score. It is this score which best represents the “functional’ or ‘operational’ IQ—the defendant’s overall general intellectual functioning...Although this does not mean a defendant’s VIQ and PIQ scores are irrelevant, it does mean a trial court is not free to disregard the FSIQ.” (p. 41) After all, the VIQ and PIQ scores are less comprehensive than the FSIQ.

The court also noted that a defendant need not be diagnosed as MR before age 18, but “there must be evidence that a diagnosis of mental retardation *could have been made* before age 18. Evidence merely of learning disabilities or mental impairment not rising to the level of mental retardation, is insufficient.” (p. 46: emphasis in original) While trial courts can use a wide range of such evidence, scores from near age 18 “must be given significant weight” (p. 46) and “may have the most probative value.” “If the onset of the disability occurs after age 18, the disability is not mental retardation.” (p. 47, fn 36). “We emphasize that we are not inserting a fixed cutoff or particular numerical score into the definition of mental retardation.” (p. 47) However, “there must be some parameters to the concept” of MR to avoid “arbitrary and capricious infliction of punishment which the Constitution forbids.” (p. 48)

“To summarize: A defendant’s FSIQ score is the primary determinant of whether his or her general intellectual functioning is significantly subaverage, as required by section 1376. the onset (origination) of mental retardation must occur before age 18, and lesser learning disabilities or mental impairments are insufficient.” (p. 48-49)

The court also rejected a purely behavioral interpretation of an adaptive behavior score: “clearly, the interpretation of such results must take into account factors such as Vidal’s cultural background and its effect on his motivation to perform the tested tasks....Test result interpretation must also take into account possible bias of informants.” (p. 52)

The appellate court concluded that the trial court “afforded insufficient significance to Vidal’s pre-age-18 FSIQ score” (p. 54) and remanded the case back for reconsideration.

The CA Supreme Court then offered an Atkins opinion in a different case ("In re Hawthorne," 2005). The court decided that post-conviction Atkins appeals should be treated the same as pre-conviction hearings. An Atkins hearing is triggered when a qualified expert files a declaration explaining "the basis for the assessment of mental retardation." The court refused to establish a cutting score of 70, since "IQ test scores are insufficiently precise to utilize a fixed cutoff," noting the SEM. In post-conviction hearings, the trier-of-fact will be a judge, who should consider the totality of the evidence regarding the defendant's mental retardation. The defendant bears the burden of proof, at the preponderance of the evidence level. A concurring opinion adds that "a person whose IQ score is over 75 is very likely not mentally retarded." The court did not address the role of the Flynn Effect or the validity of the part score argument, leaving the Vidal opinion a guiding one.

The CA Supreme Court then asked the 5th Appellate Court to reconsider the Vidal opinion in light of Hawthorne. The 5th Appellate Court reissued its opinion: "Hawthorne is not inconsistent with our determination that a defendant's FSIQ score is the primary factor in assessing his or her general intellectual functioning." (p. 43)

It is tempting to strive for a definition of mental retardation that is concrete and thus easily applied with uniformity. It does not appear, however, that the subject can be so readily condensed. While the professionals continue their research into the subject of subaverage intellectual functioning, the courts will of necessity face the inconsistencies of expert testimony and will have to trust in the ability of judges and juries to apply the law in a way that comports with the community standards upon which the decision in *Atkins* is based. (p. 55-56)

8/10/05

The U.S. Supreme Court sent the Atkins case back to Virginia for a hearing on whether or not Atkins is MR. On 8/5/05, a York county jury found that he was not. The AP story adds that he was given four IQ tests post-arrest, with scores of 59, 67, 74, and 76. Further appeals are expected. The URL for the story is at:

<http://www.wavy.com/Global/story.asp?S=3683967&nav=23iicxyI>

9/16/06

On June 8, 2006, the VA Supreme Court reviewed the death qualified jury finding that Atkins was not MR. The opinion explains that Dr. Samenow again testified for the prosecution, this time

giving the WAIS (presumably III). On voir dire, he said he had never given any test of adaptive functioning, and was therefore not an expert on such tests. The court noted that the VA legislature had included unambiguous language in their version of the Atkins law: the psychologist must be "skilled in the administration, scoring and interpretation of intelligence tests and measures of adaptive behavior." (p. 6) Since Dr. Samenow admittedly did not meet this requirement, the court reversed the jury's decision and sent the case back for a fourth jury. The opinion can be downloaded via:

<http://www.courts.state.va.us/opinions/opnscvwp/1052348.pdf#search='Daryl%20Atkins>

6/5/07

The CA Supreme Court has reviewed the Vidal case and sent it back to the Appellate court ("Peo v Sup Ct of Tulare Co (Vidal, real party in interest)," 2007). In Vidal 2004 (q.v.), an Atkins case, the CA Appellate court ruled that, since the FSIQ was the best measure of general intelligence, FSIQ is to be considered over other scores in deciding whether a defendant is mentally retarded. The Supreme court disagreed, refusing to decide on a bright line. "We are not faced here with a question of admissibility of disputed evidence but with the question whether, when both sides of a scientific dispute have been presented by expert testimony, an appellate court may declare the debate's winner as a matter of law...The question of how best to measure intellectual functioning in a given case is thus one of fact to be resolved in each case on the evidence, not by appellate promulgation of a new legal rule."

The opinion can be downloaded via: [http://wpe.info/reprints available.html](http://wpe.info/reprints_available.html)

2/20/08

The Atkins case has taken another, perhaps last turn, having been remanded back to a judge. Atkins' codefendant testified against him. The codefendant's attorney revealed that the testimony was coached by the defense attorney and prosecutor to better fit the evidence. Faced with this prosecutorial misconduct, the judge has vacated the death penalty and sentenced Atkins to LWOP, making the issue of Atkins' possible retardation moot. The story can be read at:

<http://www.nytimes.com/2008/01/19/us/19death.html?pagewanted=1&ei=5070&en=13df2b12ecc773c3&ex=1201410000&emc=eta1>

As explained in "Too dumb to die", the detection of malingering in Atkins evaluations has

been limited by the lack of adequately validated tools. One major problem is that any method of detection must be used in a known population of MR to determine the method's accuracy, especially its FPR. Two new studies have now done just that, with mixed results. Both can be requested from the authors via WPE's Reprints Available page.

Marshall & Happe (Marshall & Happe, 2007) gave two standalone tests and four within-test indicators⁷ that had been previously validated to 100 people with MR. None of them had any known reason to malingering. Both standalone tests failed the cross-validation, Rey's Fifteen Item Test had an FPR of 55%. Rey's Dot Counting Test had an FPR of 79%. Of the four within-test indicators, the WAIS-III TSF⁸ had an FPR of 2%, the WAIS-III Reliable Digit Span had an FPR of 69%, the CVLT-II Forced Choice Recognition subtest had an FPR of 11%, and the WMS-III Rarely Missed Index had an FPR of 9%.

Graue and colleagues (Graue et al., 2007) gave five standalone and four within-test indicators to 26 MR patients and demographically matched community volunteers, 25 of whom were asked to fake and 10 were asked to do the tests giving best effort. Table 1 presents the results.

Table 1: Results from Graue et (2007)

Test Name	Subtest	FNR	FPR
WAIS-III	Digit Span Scaled Score	32	81
	TSF	100	0
	SSF	52	35
	RDS	44	85
M-FAST	Total	32	50
SIMS	Total	12	77
	Low Intelligence	0	100
Digit Memory Test	Correct	24	15
Letter Memory Test	Correct	12	42
TOMM	Trial 2	20	31
	Trial 3	20	19

⁷ See TDTD for a review of malingering methodology.

⁸ See the first part of this paper for a review of the TSF's previous validation studies.

The SIMS Low Intelligence subtest was devised specifically to detect malingering of MR. It had an 100% FPR. The TSF had an 100% FNR. The SSF had a 52% FNR and 35% FPR. Given the uniformly disappointing results, the authors suggested adjusting the cutting scores for the DMT, LMT, and TOMM, but correctly noted that these new cutting scores would have to be cross-validated before being applied.

TSF and SSF continue to yield disappointing results. The TOMM has a 41% FPR in an MR population (Hurley & Deal, 2006), so cannot be used (at least until the new cutoffs are crossed). The CVLT-II has a 60% FNR (Donders & Boonstra, 2007). The WMS-III has a 75% FNR (Lange, Sullivan, & Anderson, 2005).

Marshall & Happe note the promise of the WMT⁹, as a German study of its use with MR yielded a remarkable 3% FPR (Brockhaus & Merten, 2004). However, they also report that the sample was rather carefully selected. It seems clear that a study of the oral WMT in a representative US sample is crucial!

References

- Axelrod, B., & Rawlings, D. (1999). Clinical utility of incomplete effort WAIS-R formulas: A longitudinal examination of individuals with traumatic brain injuries. *Journal of Forensic Neuropsychology, 1*, 15-27.
- Brockhaus, R., & Merten, T. (2004). Neuropsychologische Diagnostik suboptimalen Leistungsverhaltens mit dem Word Memory Test [Neuropsychological Assessment of Suboptimal Performance: The Word Memory Test]. *Nervenarzt, 75*(9), 882-887.
- Donders, J., & Boonstra, T. (2007). Correlates of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury, 21*(3), 319-326.
- Flynn, J. R. (1983). Now the great augmentation of the American IQ. *Nature, 301*(5902), 655.
- Flynn, J. R. (1984a). IQ gains and the Binet decrements. *Journal of Educational Measurement, 21*(3), 283.
- Flynn, J. R. (1984b). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*(1), 29.
- Flynn, J. R. (1985). Wechsler intelligence tests: Do we really have a criterion of mental retardation? *American Journal of Mental Deficiency, 90*(3), 236.

⁹ See my WPE paper for a review of the validations of the Word Memory Test (McKinzey, 2007, October 15).

- Flynn, J. R. (1987a). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*(2), 171.
- Flynn, J. R. (1987b). The rise and fall of Japanese IQ. *Bulletin of the British Psychological Society*, *40*, 459.
- Flynn, J. R. (1990). Massive IQ gains on the Scottish WISC: Evidence against Brand et al.'s hypothesis. *Irish Journal of Psychology*, *11*(1), 41.
- Flynn, J. R. (1998a). Israeli military IQ tests: Gender differences small; IQ gains large. *Journal of Biosocial Science*, *30*(4), 541.
- Flynn, J. R. (1998b). Rising IQ scores: Implications for the elderly. *Australasian Journal on Ageing*, *17*(3), 106.
- Flynn, J. R. (1998c). WAIS-III and WISC-III gains in the United States from 1972 to 1995: How to compensate for obsolete norms. *Perceptual & Motor Skills*, *86*(3, Pt 2), 1231.
- Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, *54*(1), 5.
- Flynn, J. R. (2000). IQ gains and fluid g. *American Psychologist*, *55*(5), 543.
- Flynn, J. R., & Detterman, D. K. (1996). What environmental factors affect intelligence: The relevance of IQ gains over time. In *The environment*. (pp. 17). Westport, CT, US: Ablex Publishing.
- Flynn, J. R., & Neisser, U. (1998). IQ gains over time: Toward finding the causes. In *The rising curve: Long-term gains in IQ and related measures*. (pp. 25). Washington, DC, US: American Psychological Association.
- Graue, L., Berry, D., Clark, J., Soliman, M., Cardi, M., Hopkins, J., et al. (2007). Identification of Feigned Mental Retardation Using the New Generation of Malingering Detection Instruments: Preliminary Findings. *Clinical Neuropsychologist*, *21*(6), 929-942.
- Greve, K., Bianchini, K., Mathias, C., Houston, R., & Crouch, J. (2003). Detecting malingered performance on the Wechsler Adult Intelligence Scale: Validation of Mittenberg's approach in traumatic injury. *Archives of Clinical Neuropsychology*, *18*, 245.
- Hurley, K., & Deal, W. (2006). Assessment Instruments measuring malingering used with individuals who have mental retardation: Potential problems and issues. *Mental Retardation*, *44*(2), 112-119.
- In re Hawthorne (35 Cal.4th 40 2005).

- Iverson, G. L., & Tulskey, D. S. (2003). Detecting malingering on the WAIS-III: Unusual Digit Span performance patterns in the normal population and in clinical groups. *Archives of Clinical Neuropsychology*, 18(1), 1-9.
- Lange, R. T., Sullivan, K., & Anderson, D. (2005). Ecological Validity of the WMS-III Rarely Missed Index in Personal Injury Litigation. *Journal of Clinical and Experimental Neuropsychology*, 27(4), 412.
- Marshall, P., & Happe, M. (2007). The performance of individuals with mental retardation on cognitive tests assessing effort and motivation. *Clinical Neuropsychologist*, 21(5), 826-840.
- McKinzey, R. K. (2007, October 15). The current accuracy rates of the Word Memory Test. *WebPsychEmpiricist* Retrieved October 15, 2007, from http://wpe.info/papers_table.html
- Millis, S. R., Ross, S. R., & Ricker, J. H. (1998). Detection of incomplete effort on the Wechsler Adult Intelligence Scale-Revised: A cross-validation. *Journal of Clinical and Experimental Neuropsychology*, 20(2), 167-173.
- Mittenberg, W., Theroux, S., Aguila-Puentes, G., Bianchini, K., Greve, K., & Rayls, K. R. (2001). Identification of malingered head injury on the Wechsler Adult Intelligence Scale--3rd edition. *The Clinical Neuropsychologist*, 15(4), 440-445.
- Mittenberg, W., Theroux-Fichera, S., Zielinski, R., & Heilbronner, R. L. (1995). Identification of malingered head injury on the Wechsler Adult Intelligence Scale--Revised. *Professional Psychology: Research & Practice*, 26(5), 491-498.
- Peo v Sup Ct of Tulare Co (Vidal, real party in interest) (CA App 5th 2005).
- Peo v Sup Ct of Tulare Co (Vidal, real party in interest) (CA Supreme 2007).
- Reschly, D., Meyers, T., & Hartel, C. (Eds.). (2002). *Mental Retardation: Determining eligibility for Social Security benefits*. Washington, D.C: National Academy Press.