

# Psychology, Public Policy, and Law

## **A META-ANALYTIC REVIEW OF COMPETENCY TO STAND TRIAL RESEARCH**

Gianni Pirelli, William H. Gottdiener, and Patricia A. Zapf

Online First Publication, January 17, 2011. doi: 10.1037/a0021713

### CITATION

Pirelli, G., Gottdiener, W. H., & Zapf, P. A. (2011, January 17). A META-ANALYTIC REVIEW OF COMPETENCY TO STAND TRIAL RESEARCH. *Psychology, Public Policy, and Law*. Advance online publication. doi: 10.1037/a0021713

# A META-ANALYTIC REVIEW OF COMPETENCY TO STAND TRIAL RESEARCH

Gianni Pirelli

The Graduate Center at John Jay College  
of Criminal Justice (CUNY)

William H. Gottdiener

John Jay College of Criminal Justice, The  
City University of New York (CUNY),  
and St. Luke's-Roosevelt Medical Center

Patricia A. Zapf

John Jay College of Criminal Justice, The City University of New York (CUNY)

The present study is a meta-analysis of competency to stand trial research. One meta-analysis was previously conducted in this area, but the large number of empirical studies that have been conducted since and the introduction of new instruments and revision of old instruments warranted updating and expanding upon the previously conducted study via contemporary meta-analytic methods. We meta-analyzed 68 studies published between 1967 and 2008 that compared competent and incompetent defendants on a number of demographic, psychiatric, and criminological variables. Categorical and continuous variables commonly investigated in competency research were coded and aggregated to generate cumulative effect sizes in the form of odds ratios and Cohen's *d* statistics, and moderation was tested via meta-*F* and meta-regression analyses. The most robust findings were that defendants diagnosed with a Psychotic Disorder were approximately eight times more likely to be found incompetent than defendants without a Psychotic Disorder diagnosis and the likelihood of being found incompetent was approximately double for unemployed defendants as compared to employed defendants. The likelihood of being found incompetent was also double for defendants with a previous psychiatric hospitalization compared to those without a hospitalization history. Comparative data on 12 competency assessment instruments and three traditional instruments were also explored and the effect sizes associated with the competency measures were substantially larger (i.e., approximately one Cohen's *d*-point) than those for the traditional measures. Limitations of the primary research and the previ-

---

Gianni Pirelli, The Graduate Center at John Jay College of Criminal Justice (CUNY); William H. Gottdiener, John Jay College of Criminal Justice, The City University of New York (CUNY), and St. Luke's-Roosevelt Medical Center; Patricia A. Zapf, John Jay College of Criminal Justice, The City University of New York (CUNY).

Dr. Pirelli is currently a Staff Clinical Psychologist 3 at Greystone Park Psychiatric Hospital in Morris Plains, NJ.

This manuscript is based on the first author's doctoral dissertation, which received the 1st Place Dissertation Award from the New Jersey Psychological Association (NJPA) and the 2nd Place Dissertation Award from the American Psychology-Law Society (AP-LS), Division 41 of the American Psychological Association. This research was supported by a Grant-in-Aid from AP-LS and it was completed when the first author was a doctoral student in the Forensic Psychology Ph.D. program of The Graduate Center at John Jay College of Criminal Justice, The City University of New York (CUNY). Thanks go to the dissertation committee for their role in this research: Drs. William H. Gottdiener, Patricia A. Zapf, Michele Galletta, Nancy L. Ryba, Barry Rosenfeld, and Robert Bornstein. Thanks also go to Grazyna Kusmierska for serving as the second coder on this project.

Correspondence concerning this article should be addressed to Gianni Pirelli, Psychology Department, Greystone Park Psychiatric Hospital, 59 Koch Avenue, Morris Plains, NJ 07950. E-mail: gianni.pirelli@gmail.com

ous and present meta-analyses are presented and future directions in this area are outlined.

*Keywords:* adjudicative competency, competency to stand trial, trial competence, forensic mental health assessment, meta-analysis

The principle in Western jurisprudence that a person must be competent to stand trial<sup>1</sup> has its roots in English common law dating back to the time of Edward I in the 14th century (Roesch & Golding, 1980) and is well documented in English case law and legal commentary (Blackstone, 1783; *Frith's Case*, 1790). The concept of competency may have stemmed from defendants who remained mute in lieu of making a plea in which case the English courts sought to determine whether their muteness was a function of “malice” or “by visitation of God” (Melton et al., 2007). The right to be competent to stand trial in American courts can be traced back to the early 19th century (*United States v. Lawrence*, 1835) and has been recognized as both a constitutional guarantee and essential to ensuring the integrity of our criminal justice system (*Drope v. Missouri*, 1975; *Youtsey v. United States*, 1899).

The current legal standard for competency to stand trial in the United States was set forth in *Dusky v. United States* (1960). In *Dusky*, the United States Supreme Court held:

It is not enough for the district judge to find that ‘the defendant is oriented to time and place and has some recollection of events’, but that the test must be whether he has sufficient present ability to consult with his lawyer with a reasonable degree of rational understanding – and whether he has a rational as well as factual understanding of the proceedings against him. (p. 402)

The *Dusky* holding has been criticized for both its brevity and ambiguity by mental health professionals and legal scholars alike. Despite these concerns, the *Dusky* standard, or some variation of it, has been adopted by every state in the United States (Favole, 1983).

### **Evaluating Competency**

Competency to stand trial evaluations have been regarded as “the most significant mental health inquiry pursued in the system of criminal law” (Stone, 1975, p. 200) with the number conducted throughout the United States each year estimated to be approximately 60,000 (Bonnie & Grisso, 2000). Over two decades ago Winick (1985) estimated that over \$185 million was spent in the United States annually for competency evaluations and related treatment (i.e., competency restoration). One decade later he suggested that this number may be closer to double or triple his initial estimate (Winick, 1996). Now that another decade has passed it is likely that this number is larger still.

In addition to monetary expenses, there are a number of costs associated with competency evaluations should they be conducted poorly. There is the potential

---

<sup>1</sup> The terms competency to stand trial, adjudicative competency, and fitness to stand trial are used interchangeably throughout the manuscript.

of violating a defendant's *due process* rights by allowing an incompetent defendant to stand trial; or of violating a defendant's civil rights by temporarily committing him or her to a forensic psychiatric facility for the purposes of competency restoration (typically via pharmacotherapy) when he or she is actually competent. These concerns are particularly salient when placed within the context of the base rates of incompetency.

Base rates from competency referrals and ultimate decisions of competency have been found to vary between and within jurisdictions and settings (Murrie, Boccaccini, Zapf, Warren, & Henderson, 2008; Nicholson & Kugler, 1991), but the modal jurisdictional estimate of incompetency for referred defendants has been thought to be 20% (Roesch, Zapf, Golding, & Skeem, 1999). Such a low base rate has major implications for the use of screening measures used to identify clearly competent defendants thereby avoiding the costs, time, and resources required to conduct full competency evaluations. These considerations have been recognized by psychologists for decades and hundreds of articles and numerous books have been published since the 1960s aimed at developing and refining practice standards in the competency arena (e.g., Ackerman, 1999; Bonnie, 1992, 1993; Goldstein, 2003, 2007; Grisso, 1986, 2003; Heilbrun, 2001; Heilbrun, Marczyk, & DeMatteo, 2002; Melton, Petrila, Poythress, & Slobogin, 1997; Melton et al., 2007; Zapf & Roesch, 2009).

Twelve<sup>2</sup> competency assessment instruments have been developed over the past 40 years intended to address a defendant's psycholegal abilities, ranging from informal checklists<sup>3</sup> to structured, criterion-based scoring instruments: the Competency Screening Test (CST; Lipsitt et al., 1971), the Competency to Stand Trial Assessment Instrument (CAI; Laboratory of Community Psychiatry, 1973), the Georgia Court Competency Test (GCCT/GCCT-MSH; Nicholson, Briggs, & Robertson, 1988), the Interdisciplinary Fitness Interview (IFI/IFI-R; Golding, 1993), the Fitness Interview Test (FIT/FIT-R; Roesch, Zapf, Eaves, & Webster, 1998), the Computer-Assisted Determination of Competency to Proceed (CADCOMP; Barnard et al., 1991), the Competence Assessment for Standing

---

<sup>2</sup> A new assessment measure was published after the present meta-analysis was completed and this manuscript was written: the Inventory of Legal Knowledge (ILK; Musick & Otto, 2010). According to the description on the Professional Assessment Resources (PAR) Website, "The ILK is *not* a test of adjudicative competence. It is solely a measure of response style; more specifically, it is a measure of a defendant's approach to inquiries about his or her legal knowledge."

<sup>3</sup> Ames Robey (1965) is credited with developing the first formal measure of competency—a checklist for psychiatrists. Robey's checklist consisted of three sections: Comprehension of Court Proceedings, Ability to Advise Counsel, and Susceptibility to Decompensation while awaiting or standing trial. Each section consisted of eight, seven, and five areas to explore, respectively. These sections were rated either, "OK," "Mental Illness," or "Intellectual Deficiency." Mental Illness included an evaluation of cognition, orientation, apperception, and judgment, and (generally) excluded character disorders. Intellectual Deficiency referred to obtaining a Wechsler Adult Intelligence Scale IQ score below 60. Robey's checklist has never been systematically studied (Roesch & Golding, 1980). Bukatman and colleagues (1971) followed with a series of interview questions designed to assess understanding of the current situation, as well as cooperation and participation in one's own defense. Although these checklists and interview questions are rarely used today, they were instrumental in providing a foundation for the assessment instruments that followed. Following these early efforts at developing checklists/interview questions, forensic psychologists began to develop instruments that were more psychometrically sound, and therefore, more clinically useful.

Trial for Defendants with Mental Retardation (CAST-MR; Everington & Luckasson, 1992), the Metropolitan Toronto Forensic Service (METFORS) Fitness Questionnaire (MFQ; Nussbaum, Mamak, Tremblay, Wright, & Callaghan, 1998), the MacArthur Competence Assessment Tool–Criminal Adjudication (MacCAT-CA; Poythress et al., 1999), the Mosley Forensic Competency Scale (MFCS; Mosley, Thyer, & Larrison, 2001), the Evaluation for Competency to Stand Trial–Revised (ECST-R; Rogers, Tillbrook, & Sewell, 2004), and the Test of Malingered Incompetence (TOMI; Colwell et al., 2008). For a full description of most of the aforementioned instruments, including instrument development, administration, scoring, and psychometric properties, readers are referred to other sources (Cooper & Grisso, 1997; Goldstein, 2003; Grisso, 1986, 1992, 2003; Melton et al., 2007; Mumley, Tillbrook, & Grisso, 2003; Pirelli, 2008; Roesch, Zapf, Golding, & Skeem, 1999; Zapf & Viljoen, 2003).

Traditional assessment instruments have also been utilized by competency examiners and researchers despite being designed to primarily measure broad psychological constructs (e.g., intelligence or personality). Although contemporary practice standards encourage the use of competency assessment instruments in evaluations (Grisso, 2003; Melton et al., 2007; Zapf & Roesch, 2009), many psychologists continue to rely heavily on traditional measures in forensic evaluations (Archer, Buffington-Vollum, Stredny, & Handel, 2006; Borum & Grisso, 1995; Nicholson & Norwood, 2000; Ryba, Cooper, & Zapf, 2003; Skeem & Golding, 1998). Three traditional measures most commonly researched in the competency arena are: the Minnesota Multiphasic Personality Inventory (MMPI/MMPI-2); the Wechsler Adult Intelligence Scales (WASI, WAIS, WAIS-R, WAIS-III); and the Brief Psychiatric Rating Scale (BPRS).

Notwithstanding the significance of the aforementioned measures' development, the notion of trial competency is socially constructed and represents an open-textured, context-specific construct and, therefore, cannot be reduced to a fixed set of psycholegal abilities (see Roesch & Golding, 1980). Thus, no instrument will ever be considered the "gold standard" for measuring competency, which complicates the evaluation process. Data from competency instruments represent only one piece of a comprehensive competency assessment and must be integrated with information obtained from clinical interviews, other relevant test data, and observations/reports from collateral sources. A number of questions vis-à-vis evaluating competency exist, including determining which variables are most closely related to findings of incompetency as well as which measures are best for use in competency evaluations. The present study was conducted, in part, to address such questions.

## **Types of Competency Research**

The competency to stand trial literature is comprised of three major areas of investigation: correlates of competency; performance of incompetent and competent defendants on traditional psychological tests; and performance of incompetent and competent defendants on specialized competency assessment measures. Studies on the correlates of competency have primarily investigated the relationship between competency status (i.e., incompetent or competent) and various demographic, psycholegal/criminological, and clinical variables. The

most commonly researched variables in this regard are: ethnicity, sex, employment status, and marital status (demographic); type of current criminal charge (e.g., violent or nonviolent) and competency evaluation history (psycholegal/criminological); and psychiatric diagnosis and psychiatric hospitalization history (clinical). Researchers examining the performance of defendants on traditional assessment measures (e.g., the WAIS) have compared the scores of incompetent and competent groups in addition to investigating associations between scores on these measures and competency status and/or the aforementioned demographic, psycholegal/criminological, and clinical variables. While some researchers have analyzed one or more of the 12 existing competency assessment instruments in the same way, most have conducted psychometric studies, whereby reliability and validity evidence for the competency measures was investigated.

Adjudicative competency research has been published steadily since the 1960s, but there is a dearth of review literature in the area. Grisso and colleagues have published three qualitative reviews/5-year research updates since 1992 (Cooper & Grisso, 1997; Grisso, 1992; Mumley, Tillbrook, & Grisso, 2003) and one meta-analysis was conducted by Nicholson and Kugler in 1991. Qualitative and quantitative reviews benefit psycholegal researchers and practitioners by providing summaries of acquired knowledge in the area of study, thereby facilitating conceptual and practical advancements (e.g., models, theories, standards of practice). Such reviews are particularly important in the competency arena because of the numerous empirical investigations conducted over the past 50 years.

### **Qualitative and Quantitative Reviews of Competency Research**

Grisso and colleagues (Cooper & Grisso, 1997; Grisso, 1992; Mumley et al., 2003) conducted three qualitative 5-year reviews over the past two decades. The reviews were divided into seven areas, which, according to Grisso, paralleled the competency assessment process: (a) the systemic context of competency to stand trial evaluations; (b) conceptual definitions of competence and models for competency to stand trial assessment; (c) research on competency assessment methods; (d) characteristics of incompetent defendants; (e) interpretation of competency evaluation data; (f) issues in competency assessment of special populations; and (g) treatment to restore competence. The authors provided the field with a template for competency research and commentary by delineating the aforementioned topics, which set the stage for theory formulation and spurred further research; however, there are limitations inherent to most, if not all, qualitative reviews (including book chapters).

First, no formal inclusion criteria typically exists; therefore, studies may not be formally vetted and they may be subsequently aggregated indiscriminately (i.e., the apples and oranges concept). Second, it is difficult to provide an overall summary of results and implications of research literature when the findings across studies are not completely consistent. Thus, authors typically engage in vote counting, such that evidentiary support is based on the number of studies with significant or nonsignificant findings rather than the magnitude of effect sizes. For example, three studies finding a nonsignificant relationship between two variables would likely be given more weight than one study with significant findings because effect size statistics are not calculated. Third, results across

reviews are compared rather than combined. Thus, authors are only able to make relative judgments (i.e., there is an increase in research in this area compared to the previous review) as opposed to aggregating results to support or fail to support specific hypotheses. As a result of these limitations, reviews may be of limited utility to the intended audience (i.e., forensic mental health professionals, judges, and lawyers).

A quantitative research synthesis, or meta-analysis, can address the limitations of qualitative research syntheses. Changes and differences in study outcomes across a large body of literature can be tracked via the calculation and analyses of effect size statistics. Furthermore, a meta-analyst can test hypotheses not previously evaluated in primary studies as well as those that cannot be tested by primary studies alone, including the investigation of potential moderator variables.

The only meta-analysis published in the adjudicative competency arena to date was conducted by Nicholson and Kugler (1991). They synthesized the findings of 27 studies<sup>4</sup> from 1967-1989 that compared competent and incompetent defendants and found the strongest correlates of incompetency to be: poor performance on competency assessment measures, a psychotic diagnosis, and psychiatric symptoms associated with severe psychopathology. Their findings are presented in greater detail below in the section entitled, Comparison with Nicholson and Kugler (1991).

Approximately 200 empirical investigations have been published and numerous competency assessment instruments have been developed and/or revised since Nicholson and Kugler's (1991) meta-analysis. The present meta-analysis was conducted to provide psycholegal researchers and practitioners with a summary of the cumulative knowledge gained over 50 years of research in this area and aimed to advance the state of knowledge in the field by testing hypotheses not previously tested in primary studies and those that cannot be tested by primary studies alone. Repetitive and/or ultimately uninformative studies may be conducted if a research literature is not meta-analyzed, as meta-analyses often serve as a new starting ground for research, practice, and policy in an area. As such, findings from the present meta-analysis should serve to close the door on some types of competency studies, while opening many new ones.

## Hypotheses

The following hypotheses were tendered based on the findings of the aforementioned qualitative reviews, previous meta-analysis, and primary research published in the competency arena:

*H*<sub>1</sub>: The mean base rate of incompetency will be between 20 and 30%.

*H*<sub>2</sub>: Demographic variables will relate to competency status; specifically, incompetency would be associated with ethnicity (i.e., Non-White); sex (i.e., Female); employment (i.e., Unemployed); and marital status (i.e., Not Married).

---

<sup>4</sup> Nicholson and Kugler (1991) reported the inclusion of 30 studies in their meta-analysis, and therefore, it has been cited as such over the years; however, only 27 independent studies were actually synthesized.

- a. Six study-level variables will moderate these relationships: (a) type of publication; (b) source of competency decision; (c) type of competent group; (d) setting; (e) country; and (f) recruitment method.
  - i. Note: *type of competent group* was developed for the purposes of this study. Five types of competent comparison groups were coded in the present study: referred defendants, purely competent defendants whose competency was never in question (e.g., inmates), those restored to competency, defendants who were initially deemed incompetent but then classified as competent by the researchers (i.e., study-competent), and a mixed group. *Recruitment* type was coded as either Archival/Retrospective (i.e., using data that has been previously collected, usually for clinical purposes) or Prospective (i.e., active recruitment of participants for the research study).

*H<sub>3</sub>*: Psychiatric and psycholegal variables (i.e., Psychotic Disorder diagnosis, previous psychiatric hospitalizations, previous competency evaluation history, and nonviolent current criminal charge) will relate to findings of incompetency.

- a. Six study-level variables will moderate these relationships: (a) type of publication; (b) source of competency decision; (c) type of competent group; (d) setting; (e) country; and (f) recruitment method.

*H<sub>4</sub>*: Scores on competency assessment measures and traditional measures (i.e., intellectual and personality assessment instruments) will both relate to competency status; however, larger effect sizes are anticipated for the relationship between scores on competency assessment instruments and such decisions.

## Method

The present meta-analysis included 68 studies published between 1967 and 2008 that compared competent<sup>5</sup> and incompetent defendants on a number of demographic, psychiatric, and criminological variables.

### Literature Search

A comprehensive search consisting of five methods was performed to identify empirical research studies in this area: (a) obtaining references of those found in acquired reports; (b) consulting with experts in the area; (c) searching electronic and print abstract databases; (d) incidental browsing of libraries and bookstores; and (e) searching citation indexes. Documents not available at local libraries were retrieved primarily through inter-library loan. A wide net was cast across five main electronic databases: (a) *PsycInfo*; (b) *PsycArticles*; (c) *Medline*; (d) Criminal Justice Periodicals Index 1981-2007; and (e) *National Criminal Justice*

---

<sup>5</sup> The term “competent defendants” is used throughout this paper. While the majority of studies conducted in this area have used referred defendants as their sample, a few studies have used competent participants which have included psychiatric patients or inmates whose competence was never questioned.



*Reference Service*, using three keywords: (1) *competenc\** to stand trial; (2) *adjudicative competenc\**; and (3) *trial competenc\** (Note: using an asterisk enables searching of various endings of the root word, such as *competency* and *competence*). Electronic searches yielded 154 potentially relevant reports and the other search methods yielded an additional 32 reports, equaling 186 potential reports. Although the first study considered for inclusion was published in 1965, the first study meeting inclusion criteria was published in 1967.

### **Publication Bias**

Retrieval of all studies ever conducted is impossible; however, publication and sampling bias was addressed by conducting a thorough literature search, which included dissertations, and via statistical methods. The *fail-safe N* statistic, developed by Rosenthal (1979), was computed contemporaneously with effect sizes to estimate the potential effects of studies not retrieved with null results or results in the opposite direction of the mean effect size. Put differently, the *fail-safe N* is an estimate of the number of unpublished studies finding null results to render a cumulative effect size nonsignificant.

### **Inclusion/Exclusion Criteria**

Although inclusion/exclusion criteria must be developed in an iterative manner, preliminary criteria were predetermined. Studies were considered for inclusion if they compared competent and incompetent defendants on at least one variable for which an effect size was calculable, if they were conducted in the United States or Canada, and if they included adult participants. Of the 186 potential reports reviewed, 88 met inclusion criteria. Of the 88, only 68 independent studies were identified and represented the total sample size ( $n = 68$ ) for the current meta-analysis (i.e., 20 reports were of redundant samples and added no new coding information). The reference list for all included studies is presented in Appendix A and the excluded study list is presented in Appendix B. Reports based on redundant samples are also included in each list when applicable. A study was typically excluded for one of four reasons. Of the excluded studies, (a) 35% did not utilize a competent comparison group; (b) 25% were solely competency restoration studies; (c) 16% met the main inclusion criteria but did not present sufficient data to code; and (4) 10% included only participants diagnosed with Mental Retardation. An additional 14% of the excluded studies were excluded for various other reasons (e.g., samples completely consisting of malingerers or coached simulators, a juvenile comparison group only, a case study, and an attorney survey).

### **Coding and Interrater Reliability**

Coding manuals and forms were developed iteratively and revised as needed. Two forms/manuals were used in the present study: one for study-level variables and one for continuous outcomes (e.g., scores on a competency instrument), both of which are available from the first author. These forms were created in FileMaker Pro, per the suggestion of Lipsey and Wilson (2001), to facilitate citation retrieval and coding, and for the maintenance of records of retrieved reports. FileMaker Pro is particularly useful for meta-analysis research because

coding is completed directly on the computer and data can be exported into Excel and other statistical software programs.

Study-level variables were defined as those related to sample characteristics as well as study design. Variables associated with sample characteristics were: sample sizes, age, ethnicity, sex, education, employment status, marital status, psychiatric diagnosis, psychiatric and competency evaluation history, and legal history. Coding of the aforementioned variables mainly consisted of rates and proportions based on categorical data. Variables associated with study design were: type of publication, publication year, source of competency decision used for comparison, type of competent comparison group, setting of study, sample's country of origin, method of participant recruitment, and type of sample (i.e., matched or random).

Continuous outcome variables were coded for data derived from scores on both traditional and competency assessment measures. Specifically, scores on 11 published competency assessment instruments were recorded, as well as scores on intelligence and personality assessment instruments. The validation study of the Test of Malingered Incompetence (TOMI; Colwell, Colwell, Perry, Wasieleski, & Billings, 2008) was not included in the present meta-analysis because it was published after data collection and coding for this study was completed.

All reports were coded by the first author, a fifth-year doctoral student at the time of the coding, and approximately 20% (i.e., 13) of the reports were coded by a second psychology doctoral student experienced in conducting meta-analyses who was also in her fifth year. The second-coding procedure consisted of a number of steps. An initial training session was conducted to review the coding manual and to provide an overview of the competency literature. This session was followed by the practice coding of 10 studies chosen via an online random number generator by both the first author and second coder. The coders met to address inconsistencies once practice coding was complete. The coding manual was subsequently revised to address all concerns elicited during the practice-coding step. A second training session was provided focusing on the revisions implemented in the coding manual. Finally, a systematic random selection procedure was used to generate interrater reliability statistics; specifically, every third study from the possible 68 studies (listed alphabetically) was chosen for inclusion in the second coding procedure.

A total of 1,194 coding decisions (i.e., each variable coded was characterized as a decision) were made across 13 studies of which the first author and second coder demonstrated strong agreement. An interrater reliability analysis was conducted vis-à-vis the coding of 1,025 continuous variables and high interrater reliability and significant statistical agreement and was found: intraclass correlation coefficient  $r = .94$  (0.94–0.95),  $p < .001$ . A kappa statistic ( $\kappa$ ) was computed to determine the level of agreement between the coders on 169 categorical variables; it was .76, and the agreement rate was approximately 84%. Although the interpretation of the kappa statistic has been debated over the years, existing benchmarks would classify a kappa of .76 as an overall high level of agreement. This kappa statistic is considered “substantial” based on Landis and Koch's (1977) classification, “good” per Altman (1991); and “excellent” per Fleiss, Levin, and Paik (2003).

## Calculation of Effect Sizes and Statistical Modeling

**Odds ratios.** The majority of data presented in the competency research literature can be conceptualized in the context of  $2 \times 2$  tables because it frequently involves an investigation of the relationship between competency status (i.e., competent/incompetent) and another dichotomous variables (e.g., psychotic/not psychotic); therefore, odds ratios (ORs) were calculated as effect sizes for these categorical data. ORs and their statistical variants (e.g., log-ORs, logit models, logistic regression models) are the recommended statistics for meta-analyses that utilize  $2 \times 2$  tables (Fleiss, 1981; Haddock, Rindskopf, & Shadish, 1998; Sanchez-Meca, Marin-Martinez, & Chacon-Moscoso, 2003; Sandercocock, 1989; Schumacker, 2005).

In the present meta-analysis, ORs were calculated to investigate the relationship between competency decision (i.e., incompetent/competent) and eight categorical variables: (a) ethnicity; (b) sex; (c) employment status; (d) marital status; (e) psychiatric diagnosis; (f) psychiatric hospitalization history; (g) competency evaluation history; (h) current criminal charge. Each variable was dichotomized in the following manner: ethnicity was analyzed as Non-White (yes/no); sex as Female (yes/no); employment as Unemployed (yes/no); marital status as Not Married (yes/no); psychiatric diagnosis as Psychotic Disorder (yes/no); psychiatric hospitalization history as Previous Psychiatric Hospitalization (yes/no). Competency evaluation history as Previous Competency Evaluation (yes/no); and, current criminal charge as Current Violent Charge (yes/no).

Using ORs as effect sizes and dichotomizing the aforementioned variables enabled a straightforward interpretation from which the actual level of *likelihood* was elicited (e.g., “Female defendants are X times more likely to be found Incompetent”). While ORs are used in the initial analyses because they are easier to interpret from a descriptive standpoint (i.e., levels of likelihood), log-ORs are easier to interpret than ORs in the context of meta-regression analyses using categorical antecedent variables because they are centered at 0, whereas ORs are centered at 1. After each variable was dichotomized and analyzed in relation to competency status, the following statistics were calculated: ORs and their associated 95% confidence intervals (CIs),  $z$ - and  $p$ -values, study weights, the cumulative random effects ORs (i.e., combined effect size of included studies on a particular variable), the median OR, *fail-safe N* (a publication bias statistic), and  $Q$  (a homogeneity statistic).

All effect size calculations were performed with Comprehensive Meta-Analysis (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2005) a widely used meta-analysis software package. Once all effect sizes were calculated, they were weighted by the inverse of their variance and summed to generate an overall mean effect size statistic (i.e., the cumulative OR); a process that also controls for sampling error (Hedges & Olkin, 1985). Cumulative ORs were generated via a random effects model, which assumes that “each observed effect size differs from the population mean by subject-level sampling error *plus* a value that represents other sources of variability assumed to be randomly distributed” (Lipsey & Wilson, 2001, p. 119). The decision to use a random effects model, rather than a fixed effects model, is subjective and is based on the analyst’s perspective on the included studies. Cooper and Hedges (1994) recommended using a random effects

model if the analyst conceptualizes the studies as different from each other in ways too complex to account for by only a few study characteristics, and if the intent of the meta-analysis is “to make inferences about a universe of such diverse studies” (p. 526). In short, there is simply too much potentially uncontrolled variance in this research area to use a fixed effects model.

**Standardized and unstandardized mean differences.** The main effect size statistics used in the present meta-analysis to investigate the differences between competent and incompetent defendants on continuous outcome measures (i.e., scores on competency and traditional assessment instruments) were unstandardized and standardized mean differences, also calculated with the use of the CMA software. The unstandardized mean difference was calculated for data generated from the same exact measure or scale across studies (e.g., the MMPI-2). The standardized mean difference (a Cohen’s *d* statistic) was calculated when the same construct was measured across studies by a different measure or scale (e.g., Verbal IQ scores measured by the WAIS, WAIS-III, and WASI). Cohen (1977, 1988) set forth the following widely accepted interpretive ranges for standardized mean difference effect sizes:  $\leq .20$  = Small;  $.50$  = Medium;  $\geq .80$  = Large. These ranges can serve as useful guidelines, but they were not empirically derived and interpretations of effect sizes are dependent on the area of study (e.g., a Medium effect size according to Cohen’s ranges may be considered Large in some domains); therefore, these statistics were converted into ORs for interpretive purposes.

**Moderator analyses.** Three steps were taken to formally test for moderation related to the analyses of categorical variables: subgroup analyses, meta *F*-tests, and meta-regression analyses (see Lipsey & Wilson, 2001, pp. 208-220, for the SPSS macros used). Subgroup analyses consisted of calculating ORs for each level of six study-level variables hypothesized to serve as potential moderators: (a) type of publication; (b) source of competency decision; (c) type of competent group; (d) setting; (e) country; and (f) recruitment. The meta *F*-test represents an analog to ANOVA, whereby each moderator is formally tested for statistically significant differences between its levels. For example, a meta *F*-test analysis of *type of publication* within the *marital status* variable would entail a calculation of the ORs elicited from journal articles, dissertations, and books for which not married (yes/no) was coded to determine if the effect sizes for the marital status variable statistically differ across types of publication. Bonferroni-type corrections were used to account for the potential of inflated Type I error. Meta-regression analyses are conceptually equivalent to multiple regression analyses insofar as predictive models are tested; however, in meta-regression analyses, the effect size serves as the outcome variable and the moderators being explored serve as the antecedent variables (i.e., predictors).

### Assessment of Study Quality

Assessment of study quality is an important process in meta-analysis and can be investigated empirically by an investigation of the abovementioned moderators hypothesized to relate to study quality. The frequencies and percentages of each potential moderator across all 68 studies are presented in Table 1.

Table 1  
*Study Descriptors of All Included Studies (n = 68)*

Descriptor	Number of studies (%)
Type of publication	
Article	56 (82.4)
Dissertation	10 (14.7)
Book	2 (2.9)
Source of competency decision	
Psychiatrist(s)	20 (29.4)
Mixed <sup>a</sup>	17 (25.0)
Mental Health Professional Team <sup>a</sup>	14 (20.6)
Court	13 (19.1)
Psychologist(s)	2 (2.9)
Not reported	2 (2.9)
Type of competent group	
Referred	59 (86.8)
Pure	4 (5.9)
Restored	3 (4.4)
Study-competent	1 (1.5)
Mixed	1 (1.5)
Setting	
Inpatient	46 (67.6)
Mixed	11 (16.2)
Outpatient	9 (13.2)
Other	1 (1.5)
Not reported	1 (1.5)
Country	
USA	52 (76.5)
Canada	16 (23.5)
Recruitment	
Archival/retrospective	40 (58.8)
Prospective	28 (41.2)
Type of setting	
Random/convenience	59 (86.8)
Matched	8 (11.8)
Other	1 (1.5)

<sup>a</sup> A Mental Health Professional Team was characterized as two or more mental health professionals working together to arrive at one decision, whereas a Mixed decision referred to one that was based on various independent decisions.

## Homogeneity Analysis

Homogeneity analyses were conducted on each of the eight categorical variables after descriptive and effect size statistics were calculated, producing a  $Q$  statistic. A significant  $Q$  indicates that the variability among effect sizes is greater than expected from sampling error alone. Formal moderation analyses were conducted if homogeneity statistics indicated significant variance across studies not because of sampling error (Hedges & Olkin, 1985). The impact of the aforementioned moderators on each effect size was investigated via meta  $F$ -tests and meta-regression analyses to determine the extent to which various aspects of study design effect or predict the calculated effect sizes.

## Results

The results are presented based on each hypothesis below.

### H<sub>1</sub>: The Mean Base Rate of Incompetency Will Be Between 20 and 30%

The first hypothesis was supported, as the base rate of incompetency was 27.5% across 59 nonmatched samples (*Median* = 25.3, *Mode* = 10); a one-sample *t*-test was conducted to produce a 95% CI around the mean estimate (25.7–33.4). The following is a presentation of additional descriptive statistics related to the included studies. Sample characteristics for all study participants (incompetent and competent) across all 68 included studies are presented in Table 2 (*n* = 26,139). It is noteworthy that only approximately half of the studies included female participants in their samples.

Characteristics of the incompetent (*n* = 6,428) and competent (*n* = 19,711) sub-samples of participants across all included studies are presented in Table 3. Although the sample studies are relatively large when combined, most of the data was derived from few studies (as is illustrated in the second column) and, therefore, the following descriptive statistics should be considered in that context. In

Table 2  
*Sample Characteristics for All Included Studies (n = 68)*

Characteristic	Number of studies	Mean	Range
<b>Study/sample</b>			
Date of publication	68	1989.9	1967–2007
Sample size ( <i>n</i> )	68	384.5 (median = 176)	21–8,416
% Incompetent	59	27.5 (median = 25.3) (mode = 10)	7–70
<b>Demographics</b>			
Age	22	33.4	29.8–37.6
% Male	41	83.0	0–100 (50–100) <sup>a</sup>
Included Females	37	—	—
% White	22	53.4	17–84
% Not Married	10	80.7	54–92
% Unemployed	8	64.5	24–88
Education level (years)	14	10.4	7.8–12
<b>Diagnosis</b>			
% Psychotic disorder	25	44.4	20–82
% Personality disorder	16	18.3	0–47
% Substance use disorder	16	17.8	0–72
% Mood disorder	15	13.4	0–32
% Mental retardation	16	6.3	0–23
<b>Psychiatric history</b>			
% Prev. psych. hospitalization(s)	5	46.1	22–56
<b>Competency history</b>			
% w/Prev. competency eval(s)	3	31.5	16–52
<b>Criminal history</b>			
% w/Prior arrest(s)	5	61.5	29–75
% Current violent crime	18	52.9	25–75

<sup>a</sup> Reflects the range in mixed-sex samples.

Table 3  
*Incompetent and Competent Sub-Sample Characteristics Across Studies (n = 68)*

Characteristic	Incompetent defendants (n = 6,428)			Competent defendants (n = 19,711)		
	Number of studies	Mean	Range	Number of studies	Mean	Range
<b>Demographics</b>						
Age	22	35.0	29-40	22	31.8	27-38
% Male	41	84.1	0-100 (45-100) <sup>a</sup>	41	81.9	0-100 (53-100) <sup>b</sup>
% White	23	47.7	0-83	22	56.9	12-91
% Not Married	10	84.0	56-100	10	77.3	52-86
% Unemployed	8	70.8	29-96	8	58.2	19-80
Education level (years)	14	10.4	7-12	14	10.5	8-12
<b>Diagnosis</b>						
% Psychotic disorder	25	66.5	30-100	25	22.2	0-64
% Personality disorder	16	8.2	0-31	17	27.9	0-73
% Substance use disorder	16	13.0	0-60	17	22.0	0-84
% Mood disorder	15	13.4	0-45	15	13.4	0-38
% Mental retardation	16	7.5	0-25	16	5.2	0-27
<b>Psychiatric history</b>						
% Prev. psych. hospitalization(s)	5	53.4	18-79	6	32.3	0-50
<b>Competency history</b>						
% w/Prev. competency eval(s)	4 (0-14) <sup>c</sup>	23.6	9-40	4	25.9	0-80
<b>Criminal history</b>						
% w/Prior arrest(s)	4	59.6	18-78	4	63.4	40-78
% Current violent crime	18	50.8	25-83	18	55.1	20-75

<sup>a</sup> When female-only samples were removed (n = 3). <sup>b</sup> When female-only samples were removed (n = 3) and one study with only 1% of competent males. <sup>c</sup> When one study with 80% was removed.

contrast to their competent counterparts, incompetent defendants were slightly older (35 years old vs. 31.8), predominantly Non-White (52.3 vs. 43.1%), had a higher unemployment rate (70.8 vs. 58.2%), and a greater percentage were not married (84 vs. 77.3%). The biggest differences between incompetent and competent defendants were on psychiatric variables. Most incompetent defendants were diagnosed with a Psychotic Disorder (66.5%) and had a previous psychiatric hospitalization (53.4%), while few incompetent defendants were diagnosed with a Personality Disorder (8.2%). This breakdown is different from the competent group, wherein only 22.2% were diagnosed with a Psychotic Disorder, 32.3% had a previous psychiatric hospitalization, and 27.9% were diagnosed with a Personality Disorder.

While there were a number of distinct differences, incompetent and competent defendants were characteristically similar across some variables. Specifically, the vast majority of all defendants were male (84.1% incompetent, 81.9% competent); had a prior arrest history (59.6% incompetent, 63.4% competent); approximately half had a current violent criminal charge (50.8% incompetent, 55.1% competent); and both groups had a mean of approximately 10 years of education (10.4 years for the incompetent group, 10.5 years for the competent group).

### **H<sub>2</sub>/H<sub>2a</sub>: Demographic Variables Will Relate to Competency Status and These Relationships Will Be Moderated by Six Study-Level Variables**

The relationship between competency status and four categorical variables was investigated: (a) ethnicity (i.e., Non-White); (b) sex (i.e., Female); (c) employment (i.e., Unemployed); and (d) marital status (i.e., Not Married).

**Ethnicity.** Across studies that presented ethnicity data ( $n = 22$ ), the cumulative OR was 1.39 (95% CI: 1.08, 1.77, *Median* = 1.38). As such, Non-White defendants were approximately one and a half times more likely to be found incompetent than White defendants, and the *fail-safe N* was 133 (i.e., there would need to be 133 unpublished studies reporting null results to reduce the OR to nonsignificance). Homogeneity analysis was conducted to determine if the variability across ORs is larger than expected from sampling error alone. Significant heterogeneity was found:  $Q(21) = 95.1, p < .01$ ; therefore, subgroup analysis was conducted to explore such variability across potential moderators. *F*-tests were conducted to formally determine if the ORs statistically differ across the levels of each moderator. ORs for two of the six moderator groups (i.e., type of competent group and recruitment) significantly differed at a .01 alpha level. Specifically, the effect size for the *pure* competent comparison group ( $OR = 2.33$ ) was significantly larger than both the *referred* ( $OR = 1.20$ ) and *restored* groups ( $OR = 1.53$ ); and, the OR for studies using a *prospective* recruitment method ( $OR = 1.77$ ) was significantly larger than those utilizing an *archival/retrospective* method ( $OR = 1.17$ ). The meta-regression model including the six moderators as predictors and the cumulative OR as the outcome was nonsignificant.

**Sex.** The cumulative OR was 1.12 (95% CI: 0.86, 1.50, *Median* = 1.15) for studies with available sex data ( $n = 18$ ), such that female defendants were essentially equally as likely as male defendants to be found incompetent (*fail-safe N* = 0). Significant heterogeneity was found:  $Q(17) = 48.5, p < .01$ . ORs for three of the six moderator groups significantly differed: type of publication, country,



and recruitment. The effect size for the relationship between sex and competency status presented in the book by Roesch and Golding (1980) was significantly lower ( $OR = 0.001$ ) than those found in the journal articles ( $OR = 1.19$ ) and dissertations ( $OR = 0.98$ ). This finding is not particularly compelling, however, because Roesch and Golding's research only included two female participants, both of whom were deemed competent. A more salient finding was the difference in effect sizes between studies conducted in the United States and Canada. Female defendants were twice as likely ( $OR = 2.03$ ) to be found incompetent than males in the four studies conducted in Canada for which sex data was available (i.e., Crocker, Favreau, & Caulet, 2002; Robertson, Gupton, McCabe, & Bankier, 1997; Roesch, Eaves, Sollner, Normandin, & Glackman, 1981; Rogers, Gillis, McMain, & Dickens, 1998) as compared to the 14 studies conducted in the United States, whereby the finding was neutral ( $OR = 1.10$ ). Last, the cumulative OR for the 14 studies using an archival/retrospective sample recruitment method was significantly larger than the four studies utilizing prospective sampling; however, both effect sizes were relatively neutral ( $ORs = 1.20$  and  $0.77$ ). The predictive meta-regression model was analyzed and found to be nonsignificant.

**Employment.** Eight studies ( $n = 8$ ) included information relevant to employment (i.e., frequencies of unemployed competent and incompetent defendants) and the cumulative OR was 2.07 (95% CI: 1.38, 3.10, *Median* = 1.77); therefore, unemployed defendants were twice as likely to be found incompetent as are employed defendants (*fail-safe*  $N = 54$ ). Significant heterogeneity was found:  $Q(7) = 15.8$ ,  $p < .05$ ; however, meta  $F$ -tests for type of publication and setting were not computable because of the limited variability within the moderator groups and no significant differences were found for comparisons across the other four moderators. A meta-regression model with only three predictors (i.e., type of competent group, country, recruitment) was investigated because of the lack of variability within the other moderator groups, but the model was not significant.

**Marital status.** The cumulative OR was 1.43 (95% CI: 1.09, 1.89, *Median* = 1.65) from studies in which marital status data was available ( $n = 10$ ), such that defendants who were not married were approximately one and a half times more likely than married defendants to be found incompetent (*fail-safe*  $N = 15$ ). Significant heterogeneity was not found:  $Q(9) = 6.43$ ,  $p > .05$ ; therefore, neither meta  $F$ -tests nor meta-regression analyses were conducted.

### **H<sub>3</sub>/H<sub>3a</sub>: Psychiatric and Psycholegal Variables Will Relate to Competency Status and These Relationships Will Be Moderated by Six Study-Level Variables**

**Psychiatric diagnosis.** The cumulative OR was 7.96 (95% CI: 5.99, 10.60, *Median* = 9.28) for studies with diagnostic data ( $n = 25$ ), such that defendants diagnosed with a Psychotic Disorder were nearly eight times more likely to be found incompetent than those without such a diagnosis. These odds are considerably larger than any of those found throughout the present study. There would need to be 5,901 unpublished studies reporting null results to reduce the OR to nonsignificance (i.e., *fail-safe*  $N$ ). Significant heterogeneity was found:  $Q(24) = 119.1$ ,  $p < .01$ ; therefore, a subgroup analysis was conducted. Only levels within the type of competent comparison group significantly differed, studies using *pure*

competent comparison groups found defendants much more likely to be found incompetent when diagnosed with a Psychotic Disorder as compared to studies consisting of *referred* ( $OR = 7.94$ ), *mixed* ( $OR = 12.95$ ), and *restored* ( $OR = 1.71$ ) competent defendants. This finding must be interpreted with caution because only two studies used *pure* competent comparison groups (i.e., Hoge et al., 1996, 1997a); in addition, the OR (33.92) is somewhat misleading when aggregated. The OR calculated from the Hoge et al. (1997a) study was 11.3, as ~65% of incompetent defendants were diagnosed with a Psychotic Disorder (i.e., 103 of 159) and 14% of competent defendants had such a diagnosis (i.e., 29 of 207). The OR generated from Hoge et al. (1996) is 207.4 because 30 of the 42 incompetent defendants were diagnosed with a Psychotic Disorder, while none of the competent defendants ( $n = 42$ ) had been given that diagnosis. A meta-regression analysis resulted in a nonsignificant model ( $p > .05$ ).

**Psychiatric hospitalization history.** The cumulative OR was 1.86 (95% CI: 1.09, 3.20, *Median* = 1.58) for studies in which psychiatric hospitalization history was available ( $n = 5$ ), such that defendants who had a previous psychiatric hospitalization were nearly twice as likely as defendants without such history to be found incompetent (*fail-safe*  $N = 48$ ). Significant heterogeneity was found:  $Q(4) = 15.5$ ,  $p < .01$ ; however, neither meta  $F$ -tests nor meta-regression analyses could be conducted. As noted, only five studies made psychiatric hospitalization history data available, and as such, virtually no variability across moderator groups existed.

**Competency evaluation history.** Only three studies ( $n = 3$ ) presented data on defendants' competency evaluation history. For interpretive purposes, competency rather than incompetency was used as the criterion in this analysis because of the nature of the data. The cumulative OR was essentially neutral ( $OR = 1.07$ ; 95% CI: 0.10, 11.20; *Median* = 2.47); that is, defendants who had a prior competency evaluation were no more likely to be found competent (or incompetent) than those who did not have such an evaluation (*fail-safe*  $N = 0$ ). Although heterogeneity was found:  $Q(2) = 90.04$ ,  $p < .01$ , neither  $F$ -tests nor meta-regression analyses were conducted because of the invariability within moderator groups.

**Current criminal charge.** Competency, rather than incompetency, was used once again as the criterion in this analysis because of the nature of the data. Defendants with a current violent criminal charge were 1.25 times more likely to be found competent than those with a current nonviolent charge ( $OR = 1.25$ ; 95% CI: 1.00, 1.60; *Median* = 0.87; *fail-safe*  $N = 60$ ) across 18 studies that presented data on type of current criminal charge (i.e., violent or nonviolent). Significant heterogeneity was found:  $Q(17) = 74.94$ ,  $p < .01$ ; meta  $F$ -tests for all potential moderator variables were conducted (with the exception of type of competent group), but no significant differences were found within groups. A meta-regression model was tested, but it was found to be nonsignificant.

#### **H<sub>4</sub>: Scores on Competency Assessment Measures and Traditional Measures Will Relate to Competency Decisions, With Larger Effect Sizes Associated With Scores on Competency Assessment Instruments**

The final hypothesis tested in this study was that scores on traditional measures (i.e., intellectual and personality assessment instruments) would be related to competency decisions, but larger effect sizes were anticipated for the

relationship between scores on competency assessment instruments and such decisions. Scores on competency assessment instruments and traditional measures were coded as continuous outcome data in the present meta-analysis. The following two sections present descriptive information and effect size data generated from studies using competency assessment instruments and those using traditional measures.

**Competency assessment instruments.** The research studies conducted on each competency instrument, including their respective total sample sizes and the inclusion/exclusion and coding status in the present meta-analysis, are presented in Table 4. (Note: dashes (–) are used to symbolize missing information related to specific sample sizes, and the plus/minus symbol ( $\pm$ ) is used to acknowledge approximate total sample sizes.) As mentioned earlier, a number of studies incorporated competency assessment measures into their designs; however, most of the published reports did not present data from which effect sizes were calculable.

Although numerous studies have incorporated competency measures into their designs, only eight independent studies have compared scores of competent and incompetent defendants on such measures to the extent that an effect size was calculable. Furthermore, sufficient data was only available for five of these measures (i.e., CST, GCCT-MSH, FIT, MFQ, MFCS), and only the CST and GCCT-MSH have such data from more than one independent study.<sup>6</sup> Descriptive and effect size data are presented in the following tables; however, neither meta-*F* nor meta-regression analyses were performed because of insufficient variability across various levels of the moderators (e.g., type of setting).

Two studies that investigated the CST presented data sufficient to calculate effect sizes. Competent defendants ( $M = 17.0$ ,  $SD = 8.8$ ) scored approximately 10 points higher than their incompetent counterparts ( $M = 26.7$ ,  $SD = 8.3$ ) on the CST (unstandardized mean difference = 9.8). The small and disproportionate sample sizes in the two studies are noteworthy (i.e., incompetent group total sample size,  $n = 26$ ; competent group sample size,  $n = 131$ ).

Four studies investigated the GCCT-MSH and reported sufficient data. Competent defendants ( $M = 81.3$ ,  $SD = 16.9$ ) scored approximately 25 points higher than their incompetent counterparts ( $M = 55.6$ ,  $SD = 25.9$ ) on the GCCT-MSH (unstandardized mean difference = 25.76, Median = 27.1). The small sample sizes across studies are noteworthy (i.e., incompetent group total sample size,  $n = 102$ ; competent group sample size,  $n = 335$ ).

The FIT, MFQ, and MFCS were used in one included study each; therefore, those data were combined with the data on the CST and GCCT-MSH to investigate the difference between competent and incompetent defendants on competency measures, in general. Standardized mean difference statistics (i.e., Cohen's *d*) were calculated rather than unstandardized differences to account for the use of different measures used to operationalize the same construct (e.g., competency to stand trial). The nine sources from which the effect sizes were generated are

---

<sup>6</sup> Two studies (Gothard, Rogers, & Sewell, 1995; Otto et al., 1998; Rogers, Sewell, Grandjean, & Vitacco, 2002) presented data on specific scales of competency measures. Otto and colleagues (1998) reported data on the Understanding, Reasoning, and Appreciation scales of the MacCAT-CA, and Gothard et al. (1995) and Rogers et al. (2002) presented data on the GCCT's Atypical Presentation Scale (APS) developed by Gothard and colleagues in 1995.

Table 4  
*Inclusion Status and Sample Sizes for Studies Using Competency Instruments*  
*(n = 59)*

Study	Included (Yes/No)	Total, <i>n</i>	Incompetent, <i>n</i>	Competent, <i>n</i>
<b>Competency Screening Test (CST)</b>				
Lipsitt et al. (1971) <sup>a</sup>	Y	43	19	24
Shatin (1979)	Y	21	9	12
Shatin & Brodsky (1979)				
Roesch & Golding (1980) <sup>a</sup>	Y	128	5	123
Nottingham & Mattson (1981) <sup>a</sup>	Y	50	4	46
Randolph et al. (1981) <sup>a</sup>	Y	25	15	10
Randolph et al. (1982) <sup>a</sup>	Y	39	10	29
Nicholson (1988)	Y	132	11	121
Nicholson, Briggs, & Robertson (1988) <sup>a</sup>				
Nicholson, Robertson, et al. (1988) <sup>a</sup>				
Bagby et al. (1992)	Y	311	121	190
Chellsen (1986) <sup>a</sup>	N	25	25	0
Paramesh (1987)	N	260	—	—
Schreiber et al. (1987) <sup>a</sup>	N	120	—	—
Roach (1994)	N	72	30	42
Smith & Hudson (1995)	N	55	—	—
Smith (1996)				
Ustad et al. (1996)	N	111	111	0
<b>Total</b>	<b>15 = 8Y/7N</b>	<b>1,392</b>	<b>358±</b>	<b>597±</b>
<b>Competency to Stand Trial Assessment Instrument (CAI)</b>				
Roesch (1978)	N	30	4	26
Roesch & Golding (1980)				
Schreiber et al. (1987)	N	120	—	—
Siegel & Elwork (1990)	N	41	41	0
Robbins et al. (1997)	N	60	17	43
Bertsch et al. (2002)	N	20	10	10
<b>Total</b>	<b>6 = 0Y/6N</b>	<b>331</b>	<b>132±</b>	<b>79±</b>
<b>Georgia Court Competency Test (GCCT/GCCT-MSH)</b>				
Nicholson (1988)	Y	132	11	121
Nicholson, Briggs, & Robertson (1988) <sup>a</sup>				
Nicholson, Robertson, et al. (1988) <sup>a</sup>				
Johnson et al. (1990) <sup>a</sup>	Y	120	9	111
Nicholson & Johnson (1991)				
Wildman et al. (1990) <sup>a</sup>	Y	100	52	48
Bagby et al. (1992)	Y	311	121	190
Gothard (1993)	Y	108	23	85
Gothard, Rogers, & Sewell (1995)				
Gothard, Viglione, et al. (1995)				
Rogers et al. (1996)	N	125	20	105
Roach (1994)	N	72	30	42
Ustad et al. (1996)	N	111	111	0
Bertman (2000)	N	26	26	0

(table continues)

Table 4 (continued)

Study	Included (Yes/No)	Total, <i>n</i>	Incompetent, <i>n</i>	Competent, <i>n</i>
Bertman et al. (2003)				
Manguno-Mire et al. (2007)	N	21	12	0
Total	10 = 5Y/5N	1,126	415	702
Interdisciplinary Fitness Interview (IFI/IFI-R)				
Golding et al. (1984) <sup>a</sup>	Y	75	17	58
Barnard et al. (1991) <sup>a</sup>	N	50	50	0
Barnard et al. (1992)		99	99	0
Holmes (1991)	N	—	—	—
Total	3 = 1Y/2N	174±	116±	58±
Fitness Interview Test (FIT/FIT-R)				
Bagby et al. (1992)	Y	311	121	190
McDonald et al. (1991)	Y	243	99	144
Viljoen et al. (2003)	Y	96	13	83
Whittemore et al. (1997)	Y	236	26	210
Zapf & Roesch (1998)	Y	178	20	158
Zapf et al. (2001)	Y	100	10	90
Menzies et al. (1983) <sup>a</sup>	N	270	—	—
Zapf & Roesch (1997)	N	57	—	—
Zapf (1999)	N	100	—	—
Zapf & Roesch (2001, 2005)				
Viljoen et al. (2002)	N	212	—	—
Viljoen & Zapf (2002)		160	—	—
Total	10 = 6Y/4N	1,924	309±	1,033±
Computer-Assisted Determination of Competency to Proceed (CADCOMP)				
Barnard et al. (1991)	N	50	50	0
Barnard et al. (1992)		99	99	0
Nicholson et al. (1994)	N	133	133	0
Roach (1994)	N	72	30	42
Buigas (1996)	N	74	—	—
Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST-MR)				
Everington (1989, 1990)	N	93	11	82
Everington & Dunn (1995)	N	35	20	15
Peacock (2005)	N	68	9	59
Bennett (2006)	N	60	60	0
Everington et al. (2007)	N	95	0	95
Stoops et al. (2007)	N	1	1	0
Total	6 = 0Y/6N	352	101	251
Metropolitan Toronto Forensic Service METFORS) Fitness Questionnaire (MFQ)				

(table continues)

Table 4 (continued)

Study	Included (Yes/No)	Total, <i>n</i>	Incompetent, <i>n</i>	Competent, <i>n</i>
Nussbaum et al. (1998)	Y	44	15	29
Nussbaum & Amaral (2001)	N	144	—	—
Total	2 = 1Y/1N	188	15±	29±
MacArthur Competence Assessment Tool — Criminal Adjudication (MacCAT-CA)				
Otto et al. (1998)	Y	729	283	446
Poythress et al. (1999)				
Zapf et al. (2005)				
Tillbrook (2001)	Y	70	33	37
Viljoen et al. (2003)	Y	96	13	83
Zapf (1999)	N	100	—	—
Zapf & Roesch (2001, 2005)				
Redlich et al. (2003)	N	17	0	17
Ryba (2005)	N	77	37	0
Bennett (2006)	N	60	60	0
Pinals et al. (2006)	N	—	—	—
Total	8 = 3Y/5N	1,149±	426±	583±
Mosley Forensic Competency Scale (MFCS)				
Mosley et al. (2001)	Y	75	19	56
Evaluation for Competency to Stand Trial-Revised (ECST-R)				
Grandjean (2004)	Y	48	30	18
Rogers et al. (2003)				
Jackson et al. (2005)	Y	137	41	96
Tillbrook (2001)	Y	70	33	37
Rogers et al. (2002, 2003)				
Rogers et al. (2004)	N	129	42	87
Gabel (2007)	N	100	—	—
Vitacco et al. (2007)				
Total	5 = 3Y/2N	484±	146±	238±
Test of Malingered Incompetence				
Colwell et al. (2008)	N	392	30	362

presented in Table 5. Incompetent and competent defendants' scores across competency assessment instruments significantly differed and a rather large effect size was found: standardized effect size ( $d$ ) = 1.4 (1.1, 1.7),  $p < .001$ , which is equal to an OR of 2.5.<sup>7</sup> The relatively small total sample sizes for both incompetent ( $n = 214$ ) and competent defendants ( $n = 574$ ) are noteworthy in this analysis.

**Traditional assessment instruments.** The three traditional measures most commonly researched in the competency arena are presented in Table 6: the MMPI/MMPI-2; the WASI, WAIS, WAIS-R, WAIS-III; and the BPRS. A num-

<sup>7</sup> ORs are presented for continuous data whereby standardized mean difference statistics were calculated to maintain continuity throughout the paper (Tables 24, 27–32). The formula provided by Borenstein (2009) was used to convert the  $d$  statistic to an OR, whereas the OR and its variance are:  $[\ln(o)] = \pi d \sqrt{3} V_{[\ln(o)]} = \pi^2 v_d 3$ .

Table 5  
Standardized Mean Difference (Cohen's *D*) Statistics for Competency Measures ( $n = 8$ )

Study	Measure	Incompetent group,		Competent group,		Cohen's $d^a$ (CI)	Odds ratio
		$M$ (SD)	$n$	$M$ (SD)	$n$		
Randolph et al. (1981)	CST	19.9 (9.3)	15	29.0 (3.8)	10	1.2 (0.6,1.9)	2.18
Nicholson (1988)	CST	14.1 (8.1)	11	24.4 (8.5)	121	1.2 (0.3,2.1)	2.18
Nicholson (1988)	GCCT-MSH	51.3 (24.8)	11	81.6 (18.3)	121	1.6 (0.9,2.3)	2.90
Johnson et al. (1990)	GCCT-MSH	66.0 (25.9)	9	79.4 (16.5)	111	0.8 (0.1,1.5)	1.45
Wildman et al. (1990)	GCCT-MSH	44.4 (30.5)	52	79.8 (21.5)	48	1.3 (0.9,1.8)	2.36
Gothard et al. (1995)	GCCT-MSH	60.5 (15.4)	30	84.5 (7.3)	55	2.2 (1.7,2.8)	3.99
McDonald et al. (1991)	FIT	42.9 (26.2)	63	28.0 (6.6)	144	0.9 <sup>b</sup> (0.7,1.3)	1.63
Nussbaum et al. (1991)	MFQ	9.8 (3.8)	15	15.9 (3.0)	29	1.9 (1.1,2.6)	3.45
Mostey et al. (2001)	MFCs	10.2 (7.4)	19	17.0 (4.1)	56	1.3 (0.8,1.9)	2.36
Overall		n/a	214 $\pm$	n/a	574 $\pm$	1.4 (1.1,1.7)	2.54

<sup>a</sup> A positive difference denotes higher scores for the *Competent* group. <sup>b</sup> The FIT standardized mean difference score was calculated as a positive integer to maintain overall consistency with the model because it is the only measure presented in the table whereby higher scores are associated with "incompetency".

$\pm$  Total sample sizes reflect independent samples only (i.e., the Nicholson, 1988, sample was included once in the calculations).

ber of other measures have been included by researchers over the years; however, most have only been included in single studies (see Table 7). Each of these measures consists of numerous scales and subscales, but most of the studies conducted in this area have only presented data on total scores or scores on a few scales. The available data for each measure is presented in the following tables. Effect sizes were calculated in the form of standardized mean differences, but neither meta-*F* nor meta-regression analyses were conducted because of the insufficient variability across the various levels of moderators.

The Wechsler instruments have been included in 36 competency studies, but very few authors have presented comparative data on competent and incompetent defendants across three main indices of cognitive functioning: Full Scale IQ, Performance IQ, and Verbal IQ. The following results are based on analysis of data from few studies with relatively small sample sizes.

Three studies investigated the Full Scale IQ (FSIQ) score differences between competent and incompetent defendants. Nestor, Daggett, Haycock, and Price (1999) and Otto and colleagues (1998) utilized the WAIS-R, while Grandjean (2004) incorporated the WASI. Competent defendants ( $M = 86.8$ ,  $SD = 14.0$ ) scored approximately six Full Scale IQ points higher than their incompetent counterparts ( $M = 80.6$ ,  $SD = 14.1$ ). The standardized mean difference (0.32, *Median* = 0.42) can be classified as small to medium and is equal to an OR of 0.58.

Four studies presented comparative Performance IQ (PIQ) data from the Wechsler scales; specifically, the WASI (Grandjean, 2004), the WAIS-R (Lesser, 1990; Nestor et al., 1999), and the WAIS-III (Shields, 2005). Competent defendants ( $M = 84.9$ ,  $SD = 14.0$ ) scored approximately five Performance IQ points greater than incompetent defendants ( $M = 79.6$ ,  $SD = 13.4$ ), which is also associated with a small to medium effect size statistic (*standardized mean difference* = 0.38, *Median* = 0.27) and is equivalent to an OR of 0.69.

Seven studies investigated differences between competent and incompetent defendants on Wechsler verbal indices, which included Verbal IQ (VIQ) scores as well as those on a Verbal Cognitive Functioning (VCF) index. The VCF index was calculated in the three included MacArthur studies (i.e., Hoge et al., 1996, 1997a; Poythress et al., 1998) using the Vocabulary, Similarities, and Digit Span subtests of the WAIS-R. Five of the seven studies utilized the WAIS-R, while the WASI and WAIS-III were included by one study each. Comparable to the aforementioned FSIQ and PIQ findings, competent defendants ( $M = 87.2$ ,  $SD = 13.5$ ) scored approximately five IQ points higher than incompetent defendants ( $M = 82.1$ ,  $SD = 12.3$ ), which translates into a small to medium effect size (*standardized mean difference* = 0.37, *Median* = 0.36) and is equal to an OR of 0.67.

The MMPI and/or the MMPI-2 have been included in 13 studies; however, the reported data is limited to validity and clinical scale scores. Furthermore, only two studies (i.e., Maxson & Neuringer, 1970; Sachsenmaier, 1991) have presented comparative data of competent and incompetent defendants on three scales for which effect sizes are calculable. Although it was not explicated in her dissertation, it seems Sachsenmaier (1991) reported scale raw scores; therefore, standardized mean difference statistics were computed in the present study's analysis.



Table 6  
*Inclusion Status and Sample Sizes for Studies Using Traditional Instruments*  
*(n = 46)*

Study	Included (Yes/No)	Total, <i>n</i>	Incompetent, <i>n</i>	Competent, <i>n</i>
<b>Minnesota Multiphasic Personality</b>				
Interview (MMPI/MMPI-2)				
Pfeiffer et al. (1967) <sup>a</sup>	Y	89	34	55
Cooke (1969) <sup>a</sup>	Y	215	93	122
Maxson & Neuringer (1970) <sup>a</sup>	Y	594	56	538
Cooke et al. (1974)	Y	325	126	199
Rogers et al. (1988) <sup>a</sup>	Y	459	56	403
Johnson et al. (1990) <sup>a</sup>	Y	120	9	111
Lesser (1990)	Y	136	52	83
Wildman et al. (1990) <sup>a</sup>	Y	100	52	48
Sachsenmaier (1991)	Y	445	97	348
Otto et al. (1998)	Y	729	283	446
Carbonell et al. (1992)	N	152	—	—
Miller (2004)	N	50	50	0
Wygant et al. (2007)	N	87	—	—
Total	13 = 10Y/3N	3,501	908±	2,353±
<b>Wechsler Abbreviated Test of</b>				
Intelligence (WASI)				
Grandjean (2004)	Y	48	30	18
<b>Wechsler Adult Intelligence Scale</b>				
(WAIS)				
Pfeiffer et al. (1967) <sup>a</sup>	Y	89	34	55
Cooke (1969) <sup>a</sup>	Y	215	93	122
Heller et al. (1981, 1983) <sup>a</sup>	Y	410	106	304
Laczko et al. (1970) <sup>a</sup>	Y	421	104	317
Shatin (1979)	Y	21	9	12
Shatin & Brodsky (1979) <sup>a</sup>				
Smith & Broughton (1994)	N	160	—	—
Smith & Hudson (1995)	N	55	—	—
Smith (1996)				
Vernon et al. (1999)	N	28	—	—
Total	8 = 5Y/3N	1,399	346±	810±
<b>Wechsler Adult Intelligence Scale-</b>				
Revised (WAIS-R)				
Johnson et al. (1990) <sup>a</sup>	Y	120	9	111
Lesser (1990)	Y	136	52	83
Sachsenmaier (1991)	Y	445	97	348
Hoge et al. (1996)	Y	84	42	42
Hoge et al. (1997a)	Y	366	159	207
Otto et al. (1998)	Y	729	283	446
Poythress et al. (1998)	Y	106	38	68
Nestor et al. (1999)	Y	181	53	128
Pierrel (1986)	N	73	—	—
Gannon (1990)	N	50	50	0
Carbonell et al. (1992)	N	152	—	—

(table continues)

Table 6 (continued)

Study	Included (Yes/No)	Total, <i>n</i>	Incompetent, <i>n</i>	Competent, <i>n</i>
Everington & Dunn (1995)	N	35	20	15
Jones (1995)	N	271	271	0
Redding (1997)	N	29	—	—
Bertman (2000)	N	26	26	0
Bertman et al. (2003)				
Nicholson, Briggs, & Robertson (1988) <sup>a</sup>	N	132	11	121
Nicholson, Robertson, et al. (1988) <sup>a</sup>				
Nicholson & Johnson (1991)				
Anderson (1999)	N	75	75	0
Anderson & Hewitt (2002)				
Viljoen et al. (2002)	N	212	—	—
Viljoen & Zapf (2002)		160	—	—
Everington et al. (2007)	N	95	—	—
Total	19 = 8Y/11N	3,316	1,186±	1,569±
Wechsler Adult Intelligence Scale- Third Edition (WAIS-III)				
Shields (2005)	Y	218	35	183
Bertsch et al. (2002)	N	20	10	10
Ryba (2005)	N	77	37	0
Patterson (2005)	N	617	617	0
Peacock (2005)	N	68	9	59
Bennett (2006)	N	60	60	0
Everington et al. (2007)	N	95	—	—
Total	7 = 1Y/6N	1,155	768±	252±
Brief Psychiatric Rating Scale (BPRS)				
Johnson et al. (1990) <sup>a</sup>	Y	120	9	111
Hoge et al. (1996)	Y	84	42	42
Hoge et al. (1997a)	Y	366	159	207
Otto et al. (1998)	Y	729	283	446
Poythress et al. (1998)	Y	106	38	68
Zapf et al. (2001)	Y	100	10	90
Viljoen et al. (2003)	Y	96	13	83
Quinsey et al. (1975)	N	56	24	0
Roach (1994)	N	72	30	42
Redding (1997)	N	29	—	—
Bertman (2000)	N	26	26	0
Bertman et al. (2003)				
Ryba (2005)	N	77	37	0
Total	12 = 7Y/5N	1,861	671±	1,089±

<sup>a</sup> Included in Nicholson & Kugler (1991).

Both Maxson and Neuringer (1970) and Sachsenmaier (1991) found incompetent defendants to produce higher scores across all three scales. The standardized mean difference effect sizes for the MMPI F scale, scale 6, and scale 8 were 0.33, 0.39, and 0.33, which are all considered small to medium. The associated OR statistics were 0.59, 0.71, and 0.59.

Twelve studies utilized the BPRS in their design, four of which investigated the total score differences between competent and incompetent defendants and presented data from which an effect size could be generated. Each study found

Table 7  
*Other Measures Used Across Studies (n = 14)*

Study	Measure(s) used
Pfeiffer et al. (1967)	Rorschach, Thematic Apperception Test (TAT), Bender Gestalt
Heller et al. (1981, 1983)	Western Personnel Test (IQ)
Simon (1987)	Quick Test, Proverbs Test
Lesser (1990)	Bender Gestalt, Psychopathy Checklist (PCL), Schedule for Affective Disorders and Schizophrenia (SADS), Structured Clinical Interview for DSM-III (SCID), Michigan Alcoholism Screening Test (MAST)
Wildman et al. (1990)	Peabody Picture Vocabulary Test
Hoge et al. (1997a)	Perceived Criminal Injustice Scale
Whittemore et al. (1997)	Test of Charter Comprehension (ToCC), Structured Clinical Interview for DSM-III –R – Patient Version (SCID-P)
Nussbaum et al. (1998)	Wechsler Memory Scales (WMS), Rey Complex Figure Test, Trail Making Test, Controlled Oral Word Association Test FAS, Common Item Estimation Test
Matthews (1999)	Scale to Assess Unawareness of Mental Disorder (SUMD), Positive and Negative Syndrome Scale (PANNS), Test of Nonverbal Intelligence-2 (TONI-2)
Nestor et al. (1999)	Wechsler Memory Scale-Revised (WMS-R), Trail Making Test, Wisconsin Card Sort Test (WCST), Wide Range Achievement Test-Revised (WRAT-R)
Zapf et al. (2001)	Structured Clinical Interview for DSM-III –R – Patient Version (SCID-P)
Grandjean (2004)	Wechsler Memory Scale (WMS), Controlled Oral Word Association Test, Stroop Color and Word Test, Continuous Performance Test, Coglab, Wisconsin Card Sort Test (WCST), Social Knowledge Questionnaire, Insight scale for Psychosis, Rey 15-item Memory Test
Jackson et al. (2005)	Miller Forensic Assessment of Symptoms Test (MFAST)
Ryba (2005)	Brief Test of Attention, Trail Making Test

higher BPRS total scores for incompetent defendants compared to competent defendants, which relates to more severe psychiatric symptomatology. The overall discrepancy was 7.4 points (*Median* = 4.9), such that incompetent defendants scored approximately 40 ( $M = 39.6$ ,  $SD = 9.2$ ), whereas competent defendants' mean BPRS total score was 32.0 ( $SD = 7.9$ ).

### **Comparison With Nicholson and Kugler (1991)**

The conduct and writing of meta-analyses have become more sophisticated and standardized in practice than they were in the late 1980s when Nicholson and Kugler conducted their study. Nevertheless, it is important to compare and contrast their study with the present meta-analysis in terms of both study design and results.

## Inclusion Criteria

The main inclusion criteria for both the present and previous meta-analysis are identical; namely, included studies must have compared competent and incompetent groups on at least one variable such that an effect size could be calculated. Despite this similarity, five studies used by Nicholson and Kugler (1991) were not included in the present meta-analysis because they were deemed to not meet inclusion criteria (see Appendix B). Furthermore, some data analyzed by Nicholson and Kugler was deemed unsuitable for coding in the present meta-analysis.<sup>8</sup>

## Calculation and Analysis of Effect Sizes

Nicholson and Kugler (1991) calculated Pearson correlation coefficients as measures of effect size in their meta-analysis, which is the same method Nicholson (1986) used in his meta-analysis investigating the correlates of civil commitment published 5 years earlier. Phi coefficients ( $\Phi$ ) were generated for  $2 \times 2$  categorical analyses (e.g., the relationship between competency status and sex) and point-biserial correlations were calculated to investigate the relationship between categorical and continuous variables (e.g., competency status and years of education). Utilizing correlation coefficients, especially the phi coefficient, presents major limitations for use with these data.

Not only do phi coefficients underestimate the population correlation coefficient (Sanchez-Meca et al., 2003), they can wildly vacillate as a function of marginal proportions and/or sample sizes alone, whereas ORs cannot (Kline, 2004). The use of correlation coefficients is particularly problematic for competency data because of the base-rate of incompetency, which is  $\sim 20\text{--}30\%$ . As Lipsey and Wilson (2001) articulated:

---

<sup>8</sup> The validation study of the Competency Screening Test (CST) conducted by Lipsitt and colleagues (1971), which was included in both meta-analyses, illustrates the point that initial inclusion of a study does not equal inclusion of all of its presented data. Lipsitt and colleagues incorporated six samples in their study: (a) defendants referred to Bridgewater State Hospital for a competency evaluation ( $n = 43$ ); (2) defendants for whom no question of competency was posed ( $n = 11$ ); (3) patients civilly committed to Bridgewater State Hospital ( $n = 47$ ); (4) patients civilly committed to Boston State Hospital ( $n = 19$ ); (5) college undergraduates ( $n = 13$ ); and, (6) a men's breakfast club associated with a church ( $n = 28$ ). The authors presented the CST means and standard deviations for each of the six aforementioned groups; however, note that none of these groups represented an incompetent group, and therefore, relevant effect sizes cannot be calculated in a comparative meta-analysis. Lipsitt and colleagues subsequently divided the group of 43 referred defendants (i.e., Bridgewater experimental group) into two groups based on their CST scores: Low ( $n = 23$ ) and High ( $n = 20$ ). The CST ranges from 0 to 44, such that higher scores relate to competency and lower scores are associated with incompetency. Lipsitt and colleagues classified a score of 21 or higher as "High"; however, this cutoff score was not derived empirically, but rather because the "research staff found that a qualitative difference in responses appeared at about a score of 20" (p. 106). Furthermore, only sample sizes (not scores) were presented for the bifurcated referred group. To summarize, the validation study of the CST conducted by Lipsitt and colleagues provides no comparative CST data. The CST means and standard deviations initially presented are those of five purely competent groups and one group referred for competency evaluations, which are inappropriate comparisons for this study; and, although the referred defendants were subsequently divided into competent and incompetent groups, CST data was not reported.

... the maximum possible phi value for a  $2 \times 2$  table with a 90–10 split on one variable and 50–50 split on the other is .33; considerably less than 1. For an 80–20 split the maximum only climbs to .5. The odds-ratio, on the other hand, is insensitive to changes in the marginal proportions, that is, the proportion split for each dichotomy, and is therefore well suited to represent low frequency events. (pp. 60–61)

In addition to being statistically problematic, use of correlation coefficients to handle  $2 \times 2$  analyses can be conceptually confusing. A correlation coefficient is a measure of association, however, and it cannot directly speak to actual levels of likelihood. Nicholson and Kugler (1991) attempted to circumvent the problems associated phi coefficients by subsequently calculating Cohen's *h* statistics. Nevertheless, this statistic remains problematic because it is a member of the *d* family and, therefore, represents the difference between two proportions rather than a measure of likelihood (Rosenthal, 1994).

Using correlation coefficients as measures of effect size to investigate the relationship between competency status and continuous outcome data is also problematic because point-biserial correlation coefficients are subject to the same range restriction problems previously outlined for categorical data (Lipsey & Wilson, 2001) and they are also conceptually limiting for these data. As such, effect size data in the form of difference-statistics is more consistent with the aims of a comparative study. Furthermore, while correlation coefficients are easily interpretable and likely more familiar to most researchers and practitioners, they are actually more convoluted and abstract than difference-statistics for these data.

## Results

The present study consisted of data from 88 total manuscripts, 68 of which represent independent studies, including 11 dissertations, published between 1967 and 2007. Nicholson and Kugler's meta-analysis was based on 27 independent studies published between 1967 and 1989. The mean total sample size found in the present study neared 400 (i.e.,  $M = 384.5$ ) with a median size of 176, whereas the previous study found a mean total sample size of 272.3. The base rate of incompetency found in the present meta-analysis was 27.5% as compared to Nicholson and Kugler's finding of 30.6%. It is not possible to compare the meta-analyses with respect to the descriptive statistics of incompetent and competent subsamples because Nicholson and Kugler only presented such data for total samples rather than bifurcating the incompetent and competent sample data.

The reporting of total sample size statistics is not applicable in the context of comparative competency research because such data is reflective of the characteristics of referred defendants rather than that of incompetent and competent defendants. For example, presenting the percent of a total sample of defendants diagnosed with a Psychotic Disorder is misleading without associated subsample statistics. In the present study, 44.4% of the total sample was diagnosed with a Psychotic Disorder; however, when divided, the groups were starkly different. Specifically, 66.5% of incompetent defendants carried such a diagnosis as compared to only 22.2% of their competent counterparts. The interested reader can compare and contrast the descriptive findings from both meta-analyses by refer-

encing Table 3 of the present study and Table 1 of Nicholson and Kugler's meta-analysis.

Nicholson and Kugler found that 86.7% of studies utilized mental health professional decisions as the ultimate competency criterion compared to 77.9% found in the present study. Approximately three-quarters (73.3%) of their studies were conducted in inpatient settings; in the present meta-analysis, 67.6% of studies were conducted inpatient-only and 16.2% were conducted in a combination of inpatient and outpatient settings.

Both meta-analyses investigated the relationship between defendants' background characteristics and competency status, as well as the association between performance on traditional and competency assessment instruments and competency status. Nicholson and Kugler coded for six demographic variables (i.e., age, gender, race, marital resources, education, and employment status). These variables were all included in the present study, although effect sizes were not calculated for age or education level because  $2 \times 2$  tables were used in the present study to generate ORs. Nicholson and Kugler found statistically significant, yet small relations between findings of incompetency and three of the four remaining demographic variables: female gender ( $n = 12$ ,  $r = .09$ ), minority race ( $n = 12$ ,  $r = .09$ ), and marital status/married ( $n = 5$ ,  $r = -.08$ ). A negative, but nonsignificant, association was found between being found incompetent and being employed ( $n = 4$ ,  $r = -.06$ ). Relatively consistent results were found in the present study. The ORs (i.e., the odds of being found incompetent if "x") for each of the aforementioned variables neared neutral in the present meta-analysis with the exception of employment status: Female ( $n = 18$ ,  $OR = 1.12$ ), Non-White ( $n = 22$ ,  $OR = 1.39$ ), Not Married ( $n = 10$ ,  $OR = 1.43$ ), and Unemployed ( $n = 8$ ,  $OR = 2.07$ ).

Nicholson and Kugler coded for three variables related to legal and psychiatric history (i.e., type of offense, previous legal involvement, previous hospitalization). The correlation between a finding of incompetency and nonviolent offense was virtually nonexistent across 12 studies ( $r = .01$ ), whereas the relationship between incompetency and having no previous legal involvement ( $n = 4$ ,  $r = .17$ ) and having a previous psychiatric hospitalization ( $n = 5$ ,  $r = .26$ ) were significant, yet small. Once again, interpretation of the present study's findings are similar, such that the odds of being found incompetent if currently charged with a violent crime was approximately neutral ( $OR = 1.3$ ) across 18 studies, and those with a psychiatric hospitalization history ( $OR = 1.86$ ) were nearly twice as likely to be found incompetent than those without such history.

Nicholson and Kugler also coded for a number of "psychiatric characteristics," including diagnoses of Mental Retardation and Psychosis, as well as eight types of psychiatric symptoms (e.g., disorientation). Diagnosis of Mental Retardation and the eight psychiatric symptom categories were also coded in the present study; however, effect sizes were only calculated for diagnosis of a Psychotic Disorder (yes/no). As was the case in Nicholson and Kugler (1991), the effect size associated with a psychotic diagnosis was the highest in the present meta-analysis. Nicholson and Kugler found the association with psychosis and incompetency to be relatively large ( $r = .45$ ) across 17 studies and, in the present study, those diagnosed with a Psychotic Disorder were nearly eight times more likely to be found incompetent than those without the diagnosis ( $n = 25$ ).

With respect to psychological test performance, Nicholson and Kugler examined the relationship between scores on four competency assessment instruments and competency status. Moderately sized, negative correlations were found for each measure, such that poor performance was related to incompetency. Effect sizes for the Competency Screening Test (CST) were derived from data from 11 studies ( $r = -.37$ ), four studies for the Georgia Court Competency Test (CGGT;  $r = -.42$ ), two studies for the Competency Assessment Instrument (CAI;  $r = -.52$ ), and one study for the Interdisciplinary Fitness Interview (IFI;  $r = -.42$ ). The data utilized by Nicholson and Kugler to calculate the aforementioned correlations are unknown. Neither the first author nor the second-coder of the present study was able to replicate Nicholson and Kugler's findings while maintaining the parameters of the inclusion criteria; therefore, meaningful comparisons between the findings of the present and previous meta-analyses are not possible.

In the present study, sufficient data were only available for five competency instruments (i.e., CST, GCCT-MSH, FIT, MFQ, MFCS), and only the CST and GCCT-MSH had such data from more than one independent study. Competent defendants scored approximately 10 points higher than incompetent defendants on the CST across two studies. Four studies utilized the GCCT-MSH and found competent defendants to score almost 26 points higher than incompetent defendants on average. The FIT, MFQ, and MFCS were used in one included study each and, when data on these measures was combined with data on the CST and GCCT-MSH, the standardized cumulative effect size was relatively large ( $d = 1.4$ ).

Nicholson and Kugler also investigated defendants' performance on IQ<sup>9</sup> and four MMPI scales. The authors reported a small, negative relationship between findings of incompetency and intelligence test scores ( $r = -.16$ ) across eight studies. In the present study, standardized mean difference statistics were calculated to represent the relationship between competency status and Wechsler FSIQ, PIQ, and VIQ/VCF indices. Competent defendants scored approximately 5 to 6 points greater than their incompetent counterparts across all three indices, and all three effect sizes were small to medium, ranging from 0.32-0.38. Nicholson and Kugler found small, yet significant correlations representing the association between incompetency and four MMPI scales across five studies; specifically, positive relationships were found for the *F* scale ( $r = .08$ ), and scales 5 ( $r = .05$ ), 6 ( $r = .08$ ), and 8 ( $r = .08$ ). Effect sizes for the *F* scale and scales 6 and 8 were calculated in the present meta-analysis. Incompetent defendants evidenced higher scores on average than competent defendants, and small to medium standardized effect sizes were found for each scale (i.e., 0.33, 0.39, and 0.33).

## Discussion

The main objective of the present study was to quantitatively synthesize the comparative research on competent and incompetent defendants via contemporary meta-analytic methods and statistical procedures; namely, to address specific research questions, including determining which variables are related to a defen-

---

<sup>9</sup> Nicholson and Kugler (1991) presented effect sizes related to "IQ score," but the intelligence measure(s) used in the calculations were not mentioned.

dant's competency status as well as the utility of various traditional and competency assessment measures in differentiating between competent and incompetent defendants.

Researchers have investigated the relationship between competency status and demographic, psycholegal, criminological, and clinical variables consistently since the 1960s. Eight of the most commonly researched variables were investigated in the present meta-analysis: ethnicity; sex; marital status; employment status; psychiatric diagnosis; psychiatric hospitalization history; competency evaluation history; and current criminal charge, most of which do not have a compelling association with findings of incompetency. In fact, only two of the eight variables (i.e., employment and psychiatric diagnosis) evidenced ORs above 2.0, such that unemployed defendants were approximately twice as likely to be found incompetent than those who are employed, and those diagnosed with a Psychotic Disorder were approximately eight times more likely to be found incompetent than those without a psychotic diagnosis.

Two of the six other variables, sex and competency evaluation history, produced neutral ORs (i.e., no difference in levels of likelihood), but the confidence interval surrounding the effect size for competency evaluation history is quite large and, therefore, to suggest a neutral relationship between that variable and competency status would be inappropriate. Only three studies were included in the analysis of competency evaluation history and their ORs differed tremendously: 0.23 (Hoge et al., 1997a); 0.40 (Reich & Wells, 1985), and 13.56 (Rosenfeld & Ritchie, 1998). The confidence intervals surrounding the ORs of the four remaining variables (i.e., ethnicity, marital status, psychiatric hospitalization history, current criminal charge) all included a neutral OR (i.e., 1.0). These data, coupled with the *fail-safe N* statistics associated with each OR, provide modest support for the relationship between competency status and most of these variables; however, this interpretation is made cautiously in light of the relatively few studies included in some of these analyses (e.g., psychiatric hospitalization and competency evaluation history analyses included five and three studies, respectively).

An additional issue to consider is that ORs were calculated for variables independently in the present study despite their actual interdependence; therefore, conditional, or joint, probabilities were not estimated. In other words, what is the associated likelihood of being found incompetent to stand trial for an unemployed, minority female who has never been married and who has been diagnosed with Schizophrenia? Unfortunately, this type of question cannot actually be answered with any scientific certainty in the present study because it would have required that primary research data be presented in a factorial manner with numerous variable combinations resulting in many cells.

For example, 256 possible defendant characteristic combinations would exist if a researcher were to collect data dichotomously (i.e., yes/no) for the eight categorical variables included in the present study (i.e.,  $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2$ ). Such data collection would obviously be overwhelming and lead to useless information; however, important information may be gained in the future if some variable combinations were investigated; specifically, it would be useful to first investigate the combination of a Psychotic Disorder diagnosis, unemployment, and a psychiatric hospitalization history, as the largest effect sizes found in



this study were associated with these variables. An essential aspect of such an analysis would be to investigate the level of multicollinearity among these variables because of their close association to one another. Perhaps a more efficient and useful method would be to conduct sequential or hierarchical logistic regressions that examine the incremental validity of the aforementioned and other legally relevant variables.

Researchers and commentators have also investigated the use of traditional and competency assessment instruments in the competency context since the 1960s and, while effect sizes corresponding to the relationship between competency status and scores on competency measures were found to be larger than those pertaining to scores on traditional measures by the present and the previous meta-analysts, comparative data related to defendants' performance on these measures remains scant. It was initially surprising that data from so few studies was available for effect size analyses pertaining to traditional and competency assessment instruments in light of the numerous publications in this area over the past 50 years; however, studies are likely to be more heavily scrutinized by meta-analysts than they are by those authoring books/chapters, journal articles, and qualitative reviews. Even when authors of such publications are scrupulous, they usually do not use formal inclusion criteria for studies incorporated in their literature review. In addition, limitations and relevant interpretive caveats are usually only briefly mentioned and may be overshadowed by a paper's main points (e.g., significant findings).

There is no way to know how many times traditional and competency assessment instruments have actually been incorporated into research studies but not reported or how much associate unpublished data exists. Nevertheless, empirical knowledge is based on available research and, as such, there is insufficient empirical evidence to support the complete superiority of competency assessment instruments over traditional measures at this time. Nonetheless, insufficient empirical evidence is not equivalent to a lack of empirical evidence, and the conclusions drawn from the present meta-analysis should not be interpreted to mean traditional measures have utility comparable to competency measures for use in competency evaluations. Competency assessment instruments, as a class of forensic assessment instruments, have been developed to address specific psycholegal questions related to the psycholegal construct of adjudicative competency; therefore, they are conceptually appropriate for use in competency evaluations. Still, the debate on the use of forensic versus traditional assessment instruments is an oversimplified characterization of the often-complicated decisions associated with choosing appropriate assessment measures for evaluations.

Substantially larger effect sizes have been found for the relationship between competency status and competency measures than for the association between competency status and traditional measures; however, these findings do not negate the fact that traditional assessment instruments can be useful in competency evaluations for specific reasons, although Nicholson and Kugler (1991) suggested otherwise:

Defendants with lower IQ scores and defendants with higher scores on Scales *F*, 5, 6, and 8 of the MMPI were more often judged incompetent. However, these correlations were small at best (all  $r_s \leq .08$ ), confuting the argument that

instruments such as the MMPI can be used for screening purposes in competency evaluation. (cf. Maxson & Neuringer, 1970, p. 363)

The reality is the MMPI and MMPI-2 are very robust measures encompassing hundreds of scales that may never be fully explored in the adjudicative competency arena. Comparative research on the nine validity scales and 10 additional primary clinical scales is needed before definitive recommendations regarding its utility for use in competency evaluations can be proffered. It is far-reaching to dismiss the potential utility of the MMPI/MMPI-2 in competency evaluations based on data from five studies on four of its scales. Personality measures can be useful in establishing the existence of a mental illness (a threshold issue) or evaluating malingering. The MMPI-2, for example, is constantly evolving and is comprised of a multitude of scales and subscales, which may have utility in this regard. Nevertheless, when measures are used inappropriately, they can be useless or misleading. Thus, the question should not be *if* traditional measures should be used in competency evaluations but rather, *when* and *how*. Unfortunately, edited books devoted to the use of traditional measures in forensic evaluations provide virtually no guidance in this regard.

Archer's (2006) book on the forensic uses of clinical assessment instruments includes whole chapters on the MMPI-2, PAI, MCMI-III, PCL-R, and Rorschach; however, the use of these measures in competency evaluations is only addressed in two places: one paragraph in the MMPI-2 chapter authored by Sellbom and Ben-Porath and three paragraphs on the Rorschach in Weiner's chapter. Gacono, Evans, Kaser-Boyd, and Gacono's (2008) recently edited a handbook on forensic Rorschach assessment, which contains a full chapter on its use in trial competency evaluations written by Gray and Acklin. Although the authors presented cogent arguments for using the Rorschach in competency evaluations, they offered no empirical support for its utility in such evaluations, likely because none exists. There are currently no published studies available that present Rorschach data from incompetent and competent defendant samples and only two studies even mention the incorporation of the Rorschach in their designs (i.e., Laboratory of Community Psychiatry, 1974; Pfeiffer, Einstein, & Dabbs, 1967). It is noteworthy that only one study mentioned the use of the Thematic Apperception Test (TAT) and Draw-a-Person Test with incompetent/competent samples, but no data were presented (i.e., Vernon, Steinberg, & Montoya, 1999).

Nicholson and Kugler (1991) found a small effect size ( $r = -.16$ ) to represent the relationship between IQ score and competency status; however, they did not provide an explanation nor interpretation of this finding. As a result, it can easily be taken out of context and misinterpreted, thereby supporting a notion that traditional measures are not useful in the assessment of adjudicative competency. All of the studies included in the present and previous meta-analysis have presented data on three main indices of intellectual functioning: Full Scale IQ, Verbal IQ or VCF, and Performance IQ. It is noteworthy that the VIQ and PIQ indices are no longer calculated by the new Wechsler adult intelligence measure (WAIS-IV). In addition, Full Scale IQ score in and of itself provides insufficient information to the competency evaluator, particularly without knowing whether it should be interpreted in the first place (i.e., the presence of a significant difference split in scores between verbal and performance indices). The findings of the

present and past meta-analysis should not be interpreted as evidence against the use of traditional measures in competency evaluations. In fact, the assessment of cognitive functioning is quite important.

Zapf (1999) argued, “what makes an individual competent is . . . *cognitive organization*” (p. 78), which she operationalized as an ability to understand, process, and express or communicate information. Zapf further argued that cognitive organization was necessary, but not sufficient, for trial competency and suggested defendants must not possess any context-specific impairment, or thought processes not based in reality, specific to the context in which the question of competency has arisen, such as a delusional belief regarding the forthcoming criminal procedures. These abilities comprise the legal standard of competence à la *Dusky*; as such, evaluating cognitive abilities is an essential component of the functional assessment of a defendant’s competency (Grisso, 2003; Zapf, 2009) or, as Skeem and Golding (1998) articulated, “Thus, the fundamental task for a forensic examiner is to relate any psychopathological or cognitive difficulties to possible impairments in the defendant’s psycholegal abilities” (p. 358).

Research on the cognitive/neuropsychological assessment of competent and incompetent samples has increased in recent years. Nussbaum and colleagues (1998) administered a cognitive battery in addition to the METFORS Fitness Questionnaire and concluded, “Empirically we have provided initial evidence that the legal fitness concept appears grounded within a cognitive psychological foundation” (p. 59). Nestor and colleagues (1999) administered a comprehensive neuropsychological battery to a sample of 181 patients committed to Bridgewater State Hospital who had undergone competency evaluations between 1987 and 1995. They found the greatest differences between competent and incompetent defendants on tests of memory, particularly verbal memory, “which assess the ability to acquire, encode, retain, and retrieve new verbal information” (p. 407), but found no differences on standardized tests of academic skills. Grandjean (2004) found competent defendants differed from incompetent defendants in four cognitive domains (i.e., verbal memory, verbal comprehension, social judgment, and executive functioning), but not on measures of visual memory, visual spatial skills, or attention. The significance of cognitive abilities in the competency context is further appreciable by the very existence of the literature pertaining to juvenile defendants and those diagnosed with Mental Retardation, wherein such abilities are directly implicated.

In summary, larger effect sizes corresponding to the relationship between competency status and competency assessment instruments have been found as compared to those associated with traditional measures; however, the data from which these effect sizes have been calculated are limited for a number of reasons and additional research on both classes of measures is needed before empirically supported conclusions can be made. Still, some evidence exists for the ability of certain measures and scales to reliably differentiate between competent and incompetent defendants. Competency assessment instruments are useful because they address competence-related abilities directly per the relevant legal standard (i.e., *Dusky*), but evaluators must be mindful when choosing which measure to use because of variability in their utility (see Grisso, 2003; Melton et al., 2007; Zapf & Viljoen, 2003). Traditional assessment instruments can be useful in competency evaluations; however, research and commentary to date has not adequately

addressed *when* and *how* they may be used most effectively. Additional research incorporating these measures is needed, particularly those designed to assess cognitive functioning.

### Limitations

The present study provides the field with a timely quantitative review of the competency to stand trial research; however, limitations remain. The first limitation is related to moderation analysis. Moderation was formally investigated in the present meta-analysis (via meta *F*-tests and meta-regression models), but these analyses were considerably constrained because of the lack of variability within moderator groups across studies. Although this is a limitation of the primary data rather than the present meta-analysis, it is noteworthy in this context.

Second, data from some groups was combined for the purposes of comparative analysis (i.e., incompetent versus competent defendants). When applicable, questionably fit groups' data was combined with that of unfit groups, and data from groups whose competency was never in question (e.g., inmates) was combined with that of competent defendant groups.<sup>10</sup> Combining data in this way is not unprecedented (see, e.g., Robertson et al., 1997; Rogers et al., 1998), but it is a method in need of further attention because, for certain variables, questionably fit groups have been found to be more similar to fit groups than to those who were unfit (e.g., McDonald, Nussbaum, & Bagby, 1991) and competent inmates receiving mental health treatment have been found to be more comparable to incompetent defendants than mentally healthy inmates (e.g., Hoge et al., 1997a; Otto et al., 1998; Poythress et al., 1998).

Last, some studies included defendants who have had a previous competency evaluation (e.g., Reich & Wells, 1985; Robertson et al., 1997), while others excluded such defendants (e.g., Bluestone & Melella, 1978), and aggregating data from these studies may have had an impact on a portion of the findings. Specifically, the relationship between having undergone a previous competency evaluation and the defendant's present competency status may be spurious in nature. For instance, in one of the few studies wherein the issue of multiple competency evaluations was investigated, Reich and Wells (1986a) compared those with previous evaluations, or "repeaters," to those who had not been previously evaluated and they found repeaters were more likely to be diagnosed with Schizophrenia and Affective Disorders, were less educated, and were found competent less often than their nonrepeater counterparts. In addition to potential within-group discrepancies among incompetent defendants, researchers who include repeaters are forced to make an arbitrary decision as to which evaluation to include in their data set. For example, Robertson and colleagues (1997) included data from the initial assessment of two participants with multiple evaluations during their data collection period, but they automatically excluded the initial evaluation data of two other participants that was gathered before the data collection period.

---

<sup>10</sup> Averaging group means is statistically correct, but doing so with standard deviations is not; therefore, a pooled standard deviation statistic was calculated to represent the "average" of two or more standard deviations.

## Conclusion

The present study was a synthesis of approximately 50 years of comparative competency research and represents the current state of knowledge with respect to differences between incompetent and competent defendants on various demographic, psycholegal, criminological, and clinical variables. The major findings of this meta-analysis are that defendants diagnosed with a Psychotic Disorder, those who are unemployed, and those with a psychiatric hospitalization history are most likely to be found incompetent to stand trial. In addition, there is a significantly greater discrepancy in scores between competent and incompetent defendants on competency assessment instruments as compared to traditional psychological measures; however, this finding is based on limited data. Although these findings have been incorporated into competency practice recommendations and standards over the years, many questions remain. Perhaps the greatest strength of this meta-analysis is its stimulation of such questions and the discontinuation of others, which will hopefully inspire novel research and commentary in this area.

There is room for improvement and growth in all areas of study and the competency arena is no different; however, that reality should not overshadow the significant advancements in this field since the 1960s. The quality of a meta-analysis is, in part, dependent on the quality of the primary research literature, which is largely reliant on our ability to learn from and improve upon earlier studies. Sustained developments in the adjudicative competency arena are also largely dependent on the continued commitment of authors to conceptualize their work as part of a field of research rather than as isolated investigations (e.g., consideration of if and how their data could be coded by future meta-analysts)- an integral step toward bridging gaps between research, practice, and policy.

## References

References marked with an asterisk indicate studies included in the meta-analysis.

- Ackerman, M. J. (1999). *Essentials of forensic psychological assessment*. New York: John Wiley & Sons, Inc.
- Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.
- Archer, R. P. (2006). *Forensic uses of clinical assessment instruments*. Mahwah, NJ: Lawrence Erlbaum.
- Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment*, 87, 84–94. doi:10.1207/s15327752jpa8701\_07
- \*Barnard, G. W., Thompson, J. W., Freeman, W. C., & Robbins, L. (1991). Competency to stand trial: Description and initial evaluation of a new computer-assisted assessment tool (CADCOMP). *Bulletin of the American Academy of Psychiatry & the Law*, 19(4), 367–381.
- Blackstone, W. (1783). *Commentaries on the laws of England* (9th ed.). London: W. Strahan.
- \*Bluestone, H., & Mellela, J. (1978). A study of criminal defendants referred for competency to stand trial in New York City. *Bulletin of the American Academy of Psychiatry & the Law*, 7(2), 166–178.
- Bonnie, R. J. (1992). The competence of criminal defendants: A theoretical reformulation. *Behavioral Sciences and the Law*, 10, 291–316. doi:10.1002/bsl.2370100303

- Bonnie, R. J. (1993). The competence of criminal defendants: Beyond *Dusky* and *Drope*. *University of Miami Law Review*, 47, 539–601.
- Bonnie, R. J., & Grisso, T. (2000). Adjudicative competence and youthful offenders. In T. Grisso & R. G. Schwartz, (Eds.), *Youth on trial: A developmental perspective on juvenile justice* (pp. 73–103). Chicago, IL: University of Chicago Press.
- Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (2nd ed., pp. 221–235). New York: Russell Sage.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive Meta-analysis* Version 2. Englewood, NJ: Biostat.
- Borum, R., & Grisso, T. (1995). Psychological test use in criminal forensic evaluations. *Professional Psychology: Research and Practice*, 26, 465–473. doi:10.1037/0735-7028.26.5.465
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Colwell, K., Colwell, L. H., Perry, A. T., Wasieleski, D., & Billings, T. (2008). The Test of Malingered Incompetence (TOMI): A forced-choice instrument for assessing cognitive malingering in competence to stand trial evaluations. *American Journal of Forensic Psychology*, 26(3), 17–42.
- Cooper, D., & Grisso, T. (1997). Five year research update (1991–1995): Evaluations for competence to stand trial. *Behavioral Sciences and the Law*, 15, 167–180. doi:10.1002/(SICI)1099-0798(199722/06)15:3<347::AID-BSL270>3.0.CO;2-K
- Cooper, H., & Hedges, L. V. (1994). Potentials and limitations of research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 521–529). New York: Russell Sage Foundation.
- Crocker, A. G., Favreau, O. E., & Caulet, M. (2002). Gender and fitness to stand trial: A 5-year review of remands in Québec. *International Journal of Law and Psychiatry*, 25, 67–84. doi:10.1016/S0160-2527(01)00089-9
- Drope v. Missouri*, 420 U.S. 162 (1975).
- Dusky v. United States*, 362 U.S. 402 (1960).
- Everington, C., & Luckasson, R. (1992). *Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST-MR)*. Worthington, OH: IDS Publishing.
- Favole, R. J. (1983). Mental disability in the American criminal process: A four issue survey. In J. Monahan & H. J. Steadman (Eds.), *Mentally disordered offenders: Perspectives from law and social science* (pp. 247–295). New York: Plenum.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.
- Fleiss, J. L., Levin, B. A., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). New York: John Wiley and Sons.
- Frith's Case, 22 How. St. Tr. 307 (1790).
- Gacono, C. B., Evans, F. B., Kaser-Boyd, N., & Gacono, L. A. (2008). *The handbook of forensic Rorschach assessment*. New York: Routledge/Taylor & Francis.
- Golding, S. L. (1993). *Interdisciplinary Fitness Interview - Revised* (Training Manual and Interview Procedure). Unpublished monograph, State of Utah Division of Mental Health.
- Goldstein, A. M. (2003). *Handbook of psychology: Forensic psychology* (Vol. 11). Hoboken, NJ: John Wiley & Sons, Inc.
- Goldstein, A. M. (2007). Forensic psychology: Toward a standard of care. In A. M. Goldstein (Ed.), *Forensic psychology: Emerging topics and expanding roles* (pp. 3–41). Hoboken, NJ: John Wiley & Sons, Inc.

- \*Grandjean, N. R. (2004). *Neuropsychological predictors of incompetency to stand trial in defendants referred for competency restoration*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2006-99002-240)
- Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments*. New York: Kluwer Academic/Plenum Publishers.
- Grisso, T. (1992). Five-year research update (1986–1990): Evaluations for competence to stand trial. *Behavioral Sciences and the Law*, *10*, 353–369. doi:10.1002/bsl.2370100306
- Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2nd ed.). New York: Kluwer Academic/Plenum Publishers.
- Haddock, C. K., Rindskopf, D., & Shadish, W. R. (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. *Psychological Methods*, *3*, 339–353. doi:10.1037/1082-989X.3.3.339
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.
- Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York: Kluwer Academic/Plenum Publishers.
- Heilbrun, K., Marczyk, G. R., & DeMatteo, D. (2002). *Forensic mental health assessment: A casebook*. New York: Oxford University Press.
- \*Hoge, S. K., Bonnie, R. J., Poythress, N., Monahan, J., Eisenberg, M., & Feucht-Haviar, T. (1997a). The MacArthur adjudicative competence study: Development and validation of a research instrument. *Law and Human Behavior*, *21*, 141–179. doi: 10.1023/A:1024826312495
- \*Hoge, S. K., Poythress, N., Bonnie, R., Eisenberg, M., Monahan, J., Feucht-Haviar, T., & Oberlander, L. (1996). Mentally ill and non-mentally ill defendants' abilities to understand information relevant to adjudication: A preliminary study. *Bulletin of the American Academy of Psychiatry & the Law*, *24*(2), 187–197.
- Kline, R. B. (2004). *Beyond significance testing: Reforming data Analysis methods in behavioral research*. Washington, DC: American Psychological Association.
- Laboratory of Community Psychiatry, Harvard Medical School. (1974). *Competency to stand trial and mental illness*. New York: Jason Aronson, Inc.
- Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. Retrieved from <http://jstor.org/stable/2529310>
- \*Lesser, F. I. (1990). *Formulating a prediction model for regaining competency to stand trial*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1991–51161-001)
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Applied Social Research Methods Series (Vol. 49). Thousand Oaks, CA: SAGE Publications.
- \*Lipsitt, P. D., Lelos., D., & McGarry, A. L. (1971). Competency for trial: A screening instrument. *American Journal of Psychiatry*, *128*(1), 105–109.
- \*Maxson, L. S., & Neuringer, C. (1970). Evaluating legal competency. *Journal of Genetic Psychology*, *117*(2), 267–273.
- \*McDonald, D. A., Nussbaum, D. S., & Bagby, R. M. (1991). Reliability, validity and utility of the Fitness Interview Test. *Canadian Journal of Psychiatry*, *36*(7), 480–484.
- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (1997). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (2nd ed.). New York: Guilford Press.
- Melton, G. B., Petrila, J., Poythress, N. G., Slobogin, C., Lyons, P. M., Jr., & Otto, R. K. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York: Guilford Press.
- Mosley, D., Thyer, B. A., & Larrison, C. (2001). Development and preliminary validation

- of the Mosley Forensic Competency Scale. *Journal of Human Behavior in the Social Environment*, 4, 41–48. doi:10.1300/J137v04n01\_03
- Mumley, D. L., Tillbrook, C. E., & Grisso, T. (2003). Five year research update (1996–2000): Evaluations for competence to stand trial (adjudicative competence). *Behavioral Sciences and the Law*, 21, 329–350. doi:10.1002/bsl.534
- Murrie, D. C., Boccaccini, M., Zapf, P. A., Warren, J. I., & Henderson, C. E. (2008). Clinician variation in findings of competence to stand trial. *Psychology, Public Policy, & Law*, 14, 177–193. doi:10.1037/a0013578
- Musick, J. E., & Otto, R. K. (2010). *Inventory of Legal Knowledge (ILK)*. Lutz, FL: Psychological Assessment Resources.
- \*Nestor, P. G., Daggett, D., Haycock, J., & Price, M. (1999). Competence to stand trial: A neuropsychological inquiry. *Law and Human Behavior*, 23, 397–412. doi:10.1023/A:1022339130582
- Nicholson, R. A. (1986). Correlates of commitment status in psychiatric patients. *Psychological Bulletin*, 100, 241–250. doi:10.1037/0033-2909.100.2.241
- \*Nicholson, R. A. (1988). Validation of a brief form of the Competency Screening Test. *Journal of Clinical Psychology*, 44(1), 87–90.
- \*Nicholson, R. A., Briggs, S. R., & Robertson, H. C. (1988). Instruments for assessing competency to stand trial: How do they work? *Professional Psychology: Research and Practice*, 19, 383–394. doi:10.1037/0735-7028.19.4.383
- Nicholson, R. A., & Kugler, K. (1991). Competent and incompetent criminal defendants: A quantitative review of comparative research. *Psychological Bulletin*, 109, 355–370. doi:10.1037/0033-2909.109.3.355
- Nicholson, R. A., & Norwood, S. (2000). The quality of forensic psychological assessments, reports, and testimony: Acknowledging the gap between promise and practice. *Law and Human Behavior*, 24, 9–44. doi:10.1023/A:1005422702678
- \*Nussbaum, D., Mamak, M., Tremblay, H., Wright, P., & Callaghan, J. (1998). The METFORS Fitness Questionnaire (MFQ): A self-report measure for screening competency to stand trial. *American Journal of Forensic Psychology*, 16(3), 41–65.
- \*Otto, R. K., Poythress, N. G., Nicholson, R. A., Edens, J. F., Monahan, J., Bonnie, R. J., Hoge, S. K., & Eisenberg, M. (1998). Psychometric properties of the MacArthur Competence Assessment Tool-Criminal Adjudication. *Psychological Assessment*, 10, 435–443. doi:10.1037/1040-3590.10.4.435
- \*Pfeiffer, E., Einstein, R. B., & Dabbs, E. G. (1967). Mental competency evaluation for the federal courts: I. Methods and results. *Journal of Nervous & Mental Disease*, 144, 320–328. doi:10.1097/00005053-196704000-00011
- Pirelli, G. (2008). *The use of arbitrary metrics in competence to stand trial assessment instruments*. (Unpublished doctoral thesis). The Graduate Center at John Jay College of Criminal Justice (The City University of New York). New York, NY.
- Poythress, N., Nicholson, R., Otto, R. K., Edens, J. F., Bonnie, R. J., Monahan, J., & Hoge, S. K. (1999). *The MacArthur Competence Assessment Tool—Criminal Adjudication: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- \*Reich, J., & Wells, J. (1985). Psychiatric diagnosis and competency to stand trial. *Comprehensive Psychiatry*, 26, 421–432. doi:10.1016/0010-440X(85)90079-3
- Reich, J., & Wells, J. (1986a). Defendants with repeated competency evaluations. *Journal of Nervous and Mental Disease*, 174, 120–122. doi:10.1097/00005053-198602000-00009
- \*Robertson, R. G., Gupton, T., McCabe, S. B., & Bankier, R. G. (1997). Clinical and demographic variables related to 'fitness to stand trial' assessments in Manitoba. *Canadian Journal of Psychiatry*, 42(2), 191–195.
- Robey, A. (1965). Criteria for competency to stand trial. A checklist for psychiatrists. *American Journal of Psychiatry*, 122, 616–623.



- \*Roesch, R., Eaves, D., Sollner, R., Normandin, M., & Glackman, W. (1981). Evaluating fitness to stand trial: A comparative analysis of fit and unfit defendants. *International Journal of Law & Psychiatry*, 4, 145–157.
- Roesch, R., & Golding, S. L. (1980). *Competency to stand trial*. Urbana: University of Illinois Press.
- Roesch, R., Zapf, P. A., Golding, S. L., & Skeem, J. L. (1999). Defining and assessing competency to stand trial. In A. K. Hess & I. B. Weiner (Eds.), *The handbook of forensic psychology* (2nd ed., pp. 327–349). New York: John Wiley & Sons, Inc.
- \*Rogers, R., Gillis, J. R., McMain, S., & Dickens, S. E. (1988). Fitness evaluations: A retrospective study of clinical, criminal, and sociodemographic characteristics. *Canadian Journal of Behavioural Science Revue canadienne des Sciences du comportement*, 20, 192–200. doi:10.1037/h0079925
- Rogers, R., Tillbrook, C. E., & Sewell, K. W. (2004). *Evaluation of Competency to Stand Trial-Revised (ECST-R) and professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.
- \*Rosenfeld, B., & Ritchie, K. (1998). Competence to stand trial: Clinician reliability and the role of offense severity. *Journal of Forensic Sciences*, 43(1), 151–157.
- Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin*, 86, 638–641. doi:10.1037/0033-2909.86.3.638
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–260). New York: Russell Sage Foundation.
- Ryba, N. L., Cooper, V. G., & Zapf, P. A. (2003). Juvenile competence to stand trial evaluations: A survey of current practices and test usage among psychologists. *Professional Psychology: Research and Practice*, 34, 499–507. doi:10.1037/0735-7028.34.5.499
- Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Mascoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, 8, 448–467. doi:10.1037/1082-989X.8.4.448
- \*Sachsenmaier, S. J. (1991). *The relationship between traditional psychometric test scores and the determination of criminal responsibility and competency to stand trial* (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1992–71233-001)
- Sandercock, P. (1989). The odds ratio: A useful tool in neurosciences. *Journal of Neurology, Neurosurgery, and Psychiatry*, 52, 817–820. doi:10.1136/jnnp.52.7.817
- Schumacker, R. E. (2005). Effect size and confidence intervals in general linear models for categorical data analysis. *Multiple Linear Regression Viewpoints*, 31(1), 42–45.
- \*Shields, C. L. (2005). *Competence to stand trial: Cognitive-related abilities and other factors*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2005–99016-192)
- Skeem, J., & Golding, S. (1998). Community examiners’ evaluations of competence to stand trial: Common problems and suggestions for improvement. *Professional Psychology: Research and Practice*, 29, 357–367. doi:10.1037/0735-7028.29.4.357
- Stone, A. (1975). *Mental health and law: A system in transition*. (DHEW Pub. No. ADM 75–176). Rockville, MD: National Institute of Mental Health.
- United States v. Lawrence, 26 F. Cas. 887 D. C. Cir. (1899).
- Vernon, M., Steinberg, A. G., & Montoya, L. A. (1999). Deaf murderers: Clinical and forensic issues. *Behavioral Sciences & the Law*, 17, 495–516. doi:10.1002/(SICI)1099-0798(199910/12)17:4<495::AID-BSL361>3.0.CO;2-6
- Winick, B. J. (1985). Restructuring competency to stand trial. *UCLA Law Review*, 32, 921–985.
- Winick, B. J. (1996). Incompetency to proceed in the criminal process: Past, present, and

- future. In D. B. Wexler & B. J. Winick (Eds.), *Law in a therapeutic key: Developments in therapeutic jurisprudence* (pp. 77–111). Durham, NC: Carolina Academic Press.
- Youtsey v. United States, 97 F. 937, 940 (6th Cir. 1899).
- Zapf, P. A. (1999). *An investigation of the construct of competence in a criminal and civil context: A comparison of the fit, the MacCAT-CA, and the MacCAT-T*. (adjudicative competence, treatment competence, Macarthur competence assessment tool). (Doctoral Dissertation) Retrieved from Dissertation Abstracts International. (1999–95024-036)
- Zapf, P. A., & Roesch, R. (2009). *Evaluation of competence to stand trial. (Best Practices in Forensic Mental Health Assessment)*. New York: Oxford.
- Zapf, P. A., & Viljoen, J. L. (2003). Issues and considerations regarding the use of assessment instruments in the evaluation of competency to stand trial. *Behavioral Sciences and the Law*, 21, 351–367. doi:10.1002/bsl.535

## Appendix A

### Reference List for Included Studies

- ◆ Used by Nicholson & Kugler (1991)  
 † Not used in Nicholson & Kugler, but eligible  
 ± Dissertation
- † Aubrey, M. (1988). Characteristics of competency referral defendants and nonreferred criminal defendants. *Journal of Psychiatry & Law*, 16(2), 233–245.
- Bagby, R. M., Nicholson, R. A., Rogers, R., & Nussbaum, D. (1992). Domains of competency to stand trial: A factor analytic study. *Law and Human Behavior*, 16, 491–507. doi:10.1007/BF01044620
- Bittman, B. J., & Convit, A. (1993). Competency, civil commitment, and the dangerousness of the mentally ill. *Journal of Forensic Sciences*, 38(6), 1460–1466.
- ◆ Bluestone, H., & Mellela, J. (1978). A study of criminal defendants referred for competency to stand trial in New York City. *Bulletin of the American Academy of Psychiatry & the Law*, 7(2), 166–178.
- Brown, D. L., Felthous, A. R., Barratt, E. S., & Stanford, M. (1994). The incompetent defendant: Support systems help avoid future legal problems. *Journal of Forensic Sciences*, 39(4), 1057–1068.
- Caldwell, R. M., Mandracchia, S. A., Ross, S. A., & Silver, N. C. (2003). Competency to stand trial and criminal responsibility: An examination of racial and gender differences among African American and Caucasian pretrial defendants. *American Journal of Forensic Psychology*, 21(3), 5–19.
- Chaimowitz, G. A., & Ferencz, J. (1999). Cost savings associated with fitness-to-stand-trial assessments in detention centres: A pilot program. *The Canadian Journal of Psychiatry*, 44(8), 808–810.
- ± Chatfield, G. G. (1998). *Repeat competence to stand trial evaluations: A marker for severe and persistent mental illness*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1998-95009-182)

(Appendices continue)

- Chellsen, J. A. (1986). Trail competency among mentally retarded offenders: Assessment techniques and related considerations. *Journal of Psychiatry and Law, 14*(1-2), 177-185.
- Cochrane, R. E., Grisso, T., & Frederick, R. I. (2001). The relationship between criminal charges, diagnoses, and psycholegal opinions among federal pretrial defendants. *Behavioral Sciences & the Law, 19*, 565-582. doi:10.1002/bsl.454
- ◆Cooke, G. (1969). The court study unit: Patient characteristics and differences between patients judged competent and incompetent. *Journal of Clinical Psychology, 25*, 140-143. doi:10.1002/1097-4679(196904)25:2<140::AID-JCLP2270250207>3.0.CO;2-J
- ◆Cooke, G., Pogany, E., & Johnston, N. G. (1974). A comparison of blacks and whites committed for evaluation of competency to stand trial on criminal charges. *Journal of Psychiatry and Law, 2*(3), 319-337.
- Cooper, V. G., & Zapf, P. A. (2003). Predictor variables in competency to stand trial decisions. *Law and Human Behavior, 27*, 423-436. doi:10.1023/A:1024089117535
- Cox, M. L., & Zapf, P. A. (2004). An investigation of discrepancies between mental health professionals and the courts in decisions about competency. *Law & Psychology Review, 28*, 109-132. Retrieved from <http://law.ua.edu/lawpsychology/>
- Crocker, A. G., Favreau, O. E., & Caulet, M. (2002). Gender and fitness to stand trial: A 5-year review of remands in Québec. *International Journal of Law and Psychiatry, 25*, 67-84. doi:10.1016/S0160-2527(01)00089-9
- †Fitzgerald, J. F., Peszke, M. A., & Goodwin, R. C. (1978). Competency evaluations in Connecticut. *Hospital & Community Psychiatry, 29*(7), 450-453.
- ◆Golding, S. L., Roesch, R., & Schreiber, J. (1984). Assessment and conceptualization of competency to stand trial: Preliminary data on the Interdisciplinary Fitness Interview. *Law & Human Behavior, 8*, 321-334. doi:10.1007/BF01044699
- †Goldstein, R. L. (1973). "The fitness factory," part I: The psychiatrist's role in determining competency. *The American Journal of Psychiatry, 130*(10), 1144-1147.
- ±Gothard, S. (1994). *The detection of malingering in competency to stand trial evaluations*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1997-70278-001)
- Gothard, S., Rogers, R., & Sewell, K. W. (1995). Feigning incompetency to stand trial: An investigation of the Georgia Court Competency Test. *Law and Human Behavior, 19*, 363-373. doi:10.1007/BF01499137
- Gothard, S., Viglione, D. J., Meloy, J. R., & Sherman, M. (1995). Detection of malingering in competency to stand trial evaluations. *Law and Human Behavior, 19*, 493-505. doi:10.1007/BF01499340
- ±Grandjean, N. R. (2004). *Neuropsychological predictors of incompetency to stand trial in defendants referred for competency restoration*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2006-99002-240)
- †Haney, C. A., & Michielutte, R. (1968). Selective factors operating in the adjudication of incompetency. *Journal of Health and Social Behavior, 9*, 233-242. doi:10.2307/2948408
- Haney, C. A., & Miller, K. S. (1970). Definitional factors in mental incompetency. *Sociology & Social Research, 54*(4), 520-532.
- Haney, C. A., Miller, K. S., & Michielutte, R. (1969). The interaction of petitioner and deviant social characteristics in the adjudication of incompetency. *Sociometry, 32*, 182-193. doi:10.2307/2786262

(Appendices continue)

- Hart, S. D., & Hare, R. D. (1992). Predicting fitness to stand trial: The relative power of demographic, criminal, and clinical variables. *Forensic Reports*, 5(1), 53–65.
- ◆ Heller, M. S., Traylor, W. H., Ehrlich, S. M., & Lester, D. (1981). Intelligence, psychosis and competency to stand trial. *Bulletin of the American Academy of Psychiatry & the Law*, 9(4), 267–274.
- ◆ Heller, M. S., Traylor, W. H., Ehrlich, S. M., & Lester, D. (1983). Competent and incompetent defendants referred to a court psychiatric clinic: A clinical comparison. *Journal of Forensic Sciences*, 28(2), 415–422.
- Hoge, S. K., Bonnie, R. J., Poythress, N., Monahan, J., Eisenberg, M., & Feucht-Haviar, T. (1997a). The MacArthur adjudicative competence study: Development and validation of a research instrument. *Law and Human Behavior*, 21, 141–179. doi:10.1023/A:1024826312495
- Hoge, S. K., Poythress, N., Bonnie, R., Eisenberg, M., Monahan, J., Feucht-Haviar, T., & Oberlander, L. (1996). Mentally ill and non-mentally ill defendants' abilities to understand information relevant to adjudication: A preliminary study. *Bulletin of the American Academy of Psychiatry & the Law*, 24(2), 187–197.
- Hubbard, K. L., Zapf, P. A., & Ronan, K. A. (2003). Competency restoration: An examination of the differences between defendants predicted restorable and not restorable to competency. *Law and Human Behavior*, 27, 127–139. doi:10.1023/A:1022566328781
- Jackson, R. L., Rogers, R., & Sewell, K. W. (2005). Forensic Applications of the Miller Forensic Assessment of Symptoms Test (MFAST): Screening for Feigned Disorders in Competency to Stand Trial Evaluations. *Law and Human Behavior*, 29, 199–210. doi:10.1007/s10979-005-2193-5
- ◆ Johnson, W. G., Nicholson, R. A., & Service, N. M. (1990). The relationship of competency to stand trial and criminal responsibility. *Criminal Justice and Behaviour*, 17, 169–185. doi:10.1177/0093854890017002002
- ◆ Laczko, A. L., James, J. F., & Alltop, L. B. (1970). A study of four hundred and thirty-five court-referred cases. *Journal of Forensic Sciences*, 15(3), 311–323.
- ± Lesser, F. I. (1990). *Formulating a prediction model for regaining competency to stand trial*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1991-51161-001)
- ◆ Lipsitt, P. D., Lelos, D., & McGarry, A. L. (1971). Competency for trial: A screening instrument. *American Journal of Psychiatry*, 128(1), 105–109.
- Martell, D. A., Rosner, R., & Harmon, R. B. (1994). Homeless mentally disordered defendants: Competency to stand trial and mental status findings. *Bulletin of the American Academy of Psychiatry & the Law*, 22(2), 289–295.
- ± Matthews, B. (1999). *Competency to stand trial and insight in mentally ill defendants*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1999-95004-052)
- ◆ Maxson, L. S., & Neuringer, C. (1970). Evaluating legal competency. *Journal of Genetic Psychology*, 117(2), 267–273.
- McDonald, D. A., Nussbaum, D. S., & Bagby, R. M. (1991). Reliability, validity and utility of the Fitness Interview Test. *The Canadian Journal of Psychiatry*, 36(7), 480–484.

(Appendices continue)

- Mosley, D., Thyer, B. A., & Larrison, C. (2001). Development and preliminary validation of the Mosley Forensic Competency Scale. *Journal of Human Behavior in the Social Environment, 4*, 41–48.
- ±Mueller, C. L. (2007). *An examination of predictor variables for competency to stand trial determinations in a multicultural setting*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2007-99016-061)
- Nestor, P. G., Daggett, D., Haycock, J., & Price, M. (1999). Competence to stand trial: A neuropsychological inquiry. *Law and Human Behavior, 23*, 397–412. doi:10.1023/A:1022339130582
- †Nicholson, R. A. (1988). Validation of a brief form of the Competency Screening Test. *Journal of Clinical Psychology, 44*(1), 87–90.
- ◆Nicholson, R. A., Briggs, S. R., & Robertson, H. C. (1988). Instruments for assessing competency to stand trial: How do they work? *Professional Psychology: Research and Practice, 19*, 383–394. doi:10.1037/0735-7028.19.4.383
- ◆Nicholson, R. A., Robertson, H. C., Johnson, W. G., & Jensen, G. (1988). A comparison of instruments for assessing competency to stand trial. *Law and Human Behavior, 12*, 313–321. doi:10.1007/BF01044387
- ◆Nottingham, E. J., & Mattson, R. E. (1981). A validation study of the competency screening test. *Law and Human Behavior, 5*, 329–336. doi:10.1007/BF01044948
- Nussbaum, D., Mamak, M., Tremblay, H., Wright, P., & Callaghan, J. (1998). The METFORS Fitness Questionnaire (MFQ): A self-report measure for screening competency to stand trial. *American Journal of Forensic Psychology, 16*(3), 41–65.
- Ohayon, M. M., Crocker, A., St-Onge, B., & Caulet, M. (1998). Fitness, responsibility, and judicially ordered assessments. *The Canadian Journal of Psychiatry, 43*(5), 491–495.
- Otto, R. K., Poythress, N. G., Nicholson, R. A., Edens, J. F., Monahan, J., Bonnie, R. J., Hoge, S. K., & Eisenberg, M. (1998). Psychometric properties of the MacArthur Competence Assessment Tool-Criminal Adjudication. *Psychological Assessment, 10*, 435–443. doi:10.1037/1040-3590.10.4.435
- ◆Pfeiffer, E., Einstein, R. B., & Dabbs, E. G. (1967). Mental competency evaluation for the federal courts: I. Methods and results. *Journal of Nervous & Mental Disease, 144*, 320–328. doi:10.1097/00005053-196704000-00011
- Poythress, N. G., Hoge, S. K., Bonnie, R. J., Monahan, J., Eisenberg, M., & Feucht-Haviar, T. (1998). The competence-related abilities of women criminal defendants. *Journal of the American Academy of Psychiatry and the Law, 26*, 215–222. doi:10.1007/s10979-006-9005-4
- ◆Randolph, J. J., Hicks, T., & Mason, D. (1981). The Competency Screening Test: A replication and extension. *Criminal Justice and Behavior, 8*, 471–482. doi:10.1177/009385488100800405
- ◆Randolph, J. J., Hicks, T., Mason, D., & Cuneo, D. J. (1982). The Competency Screening Test: A validation study in Cook County, Illinois. *Criminal Justice & Behavior, 9*(4), 495–500.
- ◆Reich, J., & Wells, J. (1985). Psychiatric diagnosis and competency to stand trial. *Comprehensive Psychiatry, 26*, 421–432. doi:10.1016/0010-440X(85)90079-3
- ±Riley, S. E. (1997). *Gender and competency adjudications in a southwestern urban court system*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1997-95015-101)

(Appendices continue)

- Riley, S. E. (1998). Competency to stand trial adjudication: A comparison of female and male defendants. *Journal of the American Academy of Psychiatry and the Law*, 26(2), 223–240.
- Robertson, R. G., Gupton, T., McCabe, S. B., & Bankier, R. G. (1997). Clinical and demographic variables related to “fitness to stand trial” assessments in Manitoba. *The Canadian Journal of Psychiatry*, 42(2), 191–195.
- ±Roesch, R. M. (1977). Competency to stand trial: An analysis of legal/mental health issues and procedures and a proposal for change. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1979-00918-001)
- ◆Roesch, R. (1978a). A brief, immediate screening interview to determine competency to stand trial: A feasibility study. *Criminal Justice & Behavior*, 5, 241–248. doi: 10.1177/009385487800500304
- ◆Roesch, R. (1978b). Competency to stand trial and court outcome. *Criminal Justice Review*, 3, 45–56.
- ◆Roesch, R. (1979). Determining competency to stand trial: An examination of evaluation procedures in an institutional setting. *Journal of Consulting & Clinical Psychology*, 47, 542–550. doi:10.1037/0022-006X.47.3.542
- ◆Roesch, R., Eaves, D., Sollner, R., Normandin, M., & Glackman, W. (1981). Evaluating fitness to stand trial: A comparative analysis of fit and unfit defendants. *International Journal of Law & Psychiatry*, 4, 145–157.
- ◆Roesch, R., & Golding, S. L. (1980). *Competency to stand trial*. Urbana: University of Illinois Press.
- Roesch, R., Ogloff, J. R. P., Hart, S. D., Dempster, R. J., Zapf, P. A., & Whittemore, K. E. (1997). The impact of Canadian criminal code changes on remands and assessments of fitness to stand trial and criminal responsibility in British Columbia. *The Canadian Journal of Psychiatry*, 42(5), 509–514.
- ◆Rogers, R., Gillis, J. R., McMain, S., & Dickens, S. E. (1988). Fitness evaluations: A retrospective study of clinical, criminal, and sociodemographic characteristics. *Canadian Journal of Behavioural Science Revue canadienne des Sciences du comportement*, 20, 192–200. doi:10.1037/h0079925
- Rosenfeld, B., & Ritchie, K. (1998). Competence to stand trial: Clinician reliability and the role of offense severity. *Journal of Forensic Sciences*, 43(1), 151–157.
- Rosenfeld, B., & Wall, A. (1998). Psychopathology and competence to stand trial. *Criminal Justice and Behavior*, 25, 443–462. doi:10.1177/0093854898025004003
- ±Sachsenmaier, S. J. (1991). *The relationship between traditional psychometric test scores and the determination of criminal responsibility and competency to stand trial*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1992-71233-001)
- Schreiber, J. (1982). Professional judgment in the assessment of competency to stand trial: Report of an evaluation study. *International Journal of Law and Psychiatry*, 5, 331–340. doi:10.1016/0160-2527(82)90027-9
- Schreiber, J., Roesch, R., & Golding, S. L. (1987). An evaluation of procedures for assessing competency to stand trial. *Bulletin of the American Academy of Psychiatry and the Law*, 15(2), 187–203.
- ◆Shatin, L., & Brodsky, S. H. (1979). Competency for trial: The Competency Screening Test in an urban hospital forensic unit. *The Mount Sinai Journal of Medicine*, 46(2), 131–134.

(Appendices continue)

- ±Shields, C. L. (2005). *Competence to stand trial: Cognitive-related abilities and other factors*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2005-99016-192)
- ◆Sikorski, C. R., & Benedek, E. P. (1977). Competency to stand trial: The female offender. *Journal of the American Medical Women's Association*, 32(4), 150–154.
- †Simon, M. J. (1987). Use of the Proverbs Test in the assessment of competency to stand trial. *Psychological Reports*, 60(3), 1166.
- Stafford, K. P., & Wygant, D. B. (2005). The role of competency to stand trial in mental health courts. *Behavioral Sciences & the Law*, 23, 245–258. doi:10.1002/bsl.649
- ±Tillbrook, C. E. (2001). *Competency to proceed: A comparative appraisal of approaches to assessment*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2001-95006-456)
- Viljoen, J. L., Zapf, P. A., & Roesch, R. (2003). Diagnosis, current symptomatology, and the ability to stand trial. *Journal of Forensic Psychology Practice*, 3, 23–37. doi: 10.1300/J158v03n04\_02
- Warren, J. I., Fitch, W. L., Dietz, P. E., & Rosenfeld, B. D. (1991). Criminal offense, psychiatric diagnosis, and psycholegal opinion: An analysis of 894 pretrial referrals. *Bulletin of the American Academy of Psychiatry & the Law*, 19(1), 63–69.
- Warren, J. I., Murrie, D. C., Stejskal, W., Colwell, L. H., Morris, J., Chauhan, P., et al. (2006). Opinion formation in evaluating the adjudicative competence and restorability of criminal defendants: A review of 8,000 evaluations. *Behavioral Sciences & the Law*, 24, 113–132. doi:10.1002/bsl.699
- Warren, J. I., Rosenfeld, B., & Fitch, W. L. (1994). Beyond competence and sanity: The influence of pretrial evaluation on case disposition. *Bulletin of the American Academy of Psychiatry & the Law*, 22(3), 379–388.
- ◆Webster, C. D., Menzies, R. J., & Jackson, M. A. (1982). *Clinical assessment before trial*. Toronto, Ontario, Canada: Butterworths.
- Whittemore, K. E., Ogloff, J. R. P., & Roesch, R. (1997). An investigation of competency to participate in legal proceedings in Canada. *The Canadian Journal of Psychiatry*, 42(8), 869–875.
- ◆Wildman, R. W., White, P. A., & Brandenburg, C. E. (1990). The Georgia Court Competency Test: The base-rate problem. *Perceptual and Motor Skills*, 70, 1055–1058.
- Zapf, P. A., Hubbard, K. L., Cooper, V. G., Wheelles, M. C., & Ronan, K. A. (2004). Have the courts abdicated their responsibility for determination of competency to stand trial to clinicians? *Journal of Forensic Psychology Practice*, 4, 27–44. doi:10.1300/J158v04n01\_02
- Zapf, P. A., & Roesch, R. (1997). Assessing fitness to stand trial: A comparison of institution-based evaluations and a brief screening interview. *Canadian Journal of Community Mental Health*, 16(1), 53–66.
- Zapf, P. A., & Roesch, R. (1998). Fitness to stand trial: Characteristics of remands since the 1992 Criminal Code amendments. *The Canadian Journal of Psychiatry*, 43(3), 287–293.
- Zapf, P. A., Roesch, R., & Viljoen, J. L. (2001). *Assessing Fitness to Stand Trial: The Utility of the Fitness Interview Test (Revised Edition)*. *The Canadian Journal of Psychiatry*, 46(5), 426–432.

(Appendices continue)

## Appendix B

### Reference List for Excluded Studies

◆ Used by Nicholson & Kugler (1991)

† Not used in Nicholson & Kugler, but eligible

± Dissertation

- Addington, D. E., & Holley, H. L. (1987). Pre-trial assessment of schizophrenics on remand. *The Canadian Journal of Psychiatry*, *32*(9), 773–776.
- ± Amadeo, A. M. (2005). *The development of a “restoration to competence” program for patients found incompetent to stand trial*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2005-99018-341)
- ± Anderson, S. D. (1999). *The effect of competency restoration training on defendants with mental retardation found not competent to proceed*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1999-95002-083)
- Anderson, S. D., & Hewitt, J. (2002). The effect of competency restoration training on defendants with mental retardation found not competent to proceed. *Law and Human Behavior*, *26*, 343–352. doi:10.1023/A:1015328505884
- Barnard, G. W., Nicholson, R. A., Hankins, G. C., & Raisani, K. K. (1992). Itemmetric and scale analysis of a new computer-assisted competency assessment instrument (CADCOMP). *Behavioral Sciences & the Law*, *10*, 419–435. doi:10.1002/bsl.2370100311
- Barnard, G. W., Thompson, J. W., Freeman, W. C., & Robbins, L. (1991). Competency to stand trial: Description and initial evaluation of a new computer-assisted assessment tool (CADCOMP). *Bulletin of the American Academy of Psychiatry & the Law*, *19*(4), 367–381.
- Beckham, J. C., Annis, L. V., & Bein, M. F. (1986). Don't pass go: Predicting who returns from court as remaining incompetent for trial. *Criminal Justice and Behavior*, *13*, 99–109. doi:10.1177/0093854886013001007
- ± Bennett, A. D. (2006). *The measurement of adjudicative competence: A comparison of three types of competencies in a sample of individuals with mental retardation*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2006-99014-226)
- Bennett, G. T., & Kish, G. R. (1990). Incompetency to stand trial: Treatment unaffected by demographic variables. *Journal of Forensic Sciences*, *35*(2), 403–412.
- Bendt, R. H., Balcanoff, E. J., & Tragellis, G. S. (1973). Incompetency to stand trial: Is psychiatry necessary? *The American Journal of Psychiatry*, *130*(11), 1288–1289.
- Berman, L. M., & Osborne, Y. H. (1987). Attorneys' referrals for competency to stand trial evaluations: Comparisons of referred and nonreferred clients. *Behavioral Sciences & the Law*, *5*, 373–380. doi:10.1002/bsl.2370050310
- ± Bertman, L. J. (2000). *Effect of an individualized treatment protocol on competency restoration in pretrial forensic inpatients*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2000-95010-190)
- Bertman, L. J., Thompson, J. W., Jr., Waters, W. F., Estupinan-Kane, L., Martin, J. A., & Russell, L. (2003). Effect of an Individualized Treatment Protocol on Restoration of Competency in Pretrial Forensic Inpatients. *Journal of the American Academy of Psychiatry and the Law*, *31*(1), 27–35.

(Appendices continue)



- Bertsch, J. M., Younglove, J. A., & Kerr, M. G. (2002). A pilot study of the Porterville Developmental Center's court competency training program. *Criminal Justice Policy Review*, 13(1), 65–77.
- Buigas, R. A. (1996). *Development of scales for the Computer-Assisted Determination of Competency*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1996-95009-060)
- Carbonell, J. L., Heilbrun, K., & Friedman, F. L. (1992). Predicting who will regain trial competency: Initial promise unfulfilled. *Forensic Reports*, 5(1), 67–76.
- ◆ Chellsen, J. A. (1986). Trial competency among mentally retarded offenders: Assessment techniques and related considerations. *Journal of Psychiatry and Law*, 14(1–2), 177–185.
- Cooke, G., Johnston, N., & Pogany, E. (1973). Factors affecting referral to determine competency to stand trial. *American Journal of Psychiatry*, 130(8), 870–875.
- Cuneo, D. J., & Brelje, T. B. (1984). Predicting probability of attaining fitness to stand trial. *Psychological Reports*, 55(1), 35–39.
- Daniel, A. E., Beck, N. C., Herath, A., Schmitz, M., & Menninger, K. (1984). Factors correlated with psychiatric recommendations of incompetency and insanity. *Journal of Psychiatry & Law*, 12(4), 527–544.
- Edens, J. F., Poythress, N. G., Nicholson, R. A., & Otto, R. K. (1999). Effects of state organizational structure and forensic examiner training on pretrial competence assessments. *Journal of Behavioral Health Services & Research*, 26, 140–150. doi: 10.1007/BF02287487
- Everington, C. (1989). Demographic variables associated with competence to stand trial referral and evaluation of criminal defendants with mental retardation. *Journal of Psychiatry & Law*, 17(4), 627–640.
- Everington, C. T. (1990). The Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST-MR): A validation study. *Criminal Justice and Behavior*, 17, 147–168. doi:10.1177/0093854890017002001
- Everington, C., & Dunn, C. (1995). A second validation study of the Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST-MR). *Criminal Justice and Behavior*, 22, 44–59. doi:10.1177/0093854895022001004
- Everington, C., Notario-Smull, H., & Horton, M. L. (2007). Can defendants with mental retardation successfully fake their performance on a test of competence to stand trial? *Behavioral Sciences & the Law*, 25, 545–560. doi:10.1002/bsl.735
- Exner, J. E., Jr. (1980). Diagnosis versus description in competency issues. In F. Wright, C. Bahn, & R. W. Rieber (Eds.), *Forensic psychology and psychiatry* (pp. 20–26). New York: New York Academy of Sciences (vol. 347).
- ± Felchli, M. A. (1992). *Construct validity of the Competency Screening Test*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1993-71809-001)
- ± Ficken, C. T. (2004). *Using demographic and clinical variables to predict the length of stay of "Incompetent to Stand Trial" patients*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2004-99002-010)
- Fontaine, P. A. (1985). The case dispositions of temporarily trial-incompetent criminal defendants. *Journal of Psychiatry & Law*, 13(3–4), 435–448.
- ± Gabel, J. E. (2007). *Consistency of screening measures of malingering among pre-trial adjudicative competence defendants*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2007-99006-167)

(Appendices continue)

- ±Gannon, J. L. (1990). *Validation of the Competency Assessment Instrument and elements of competency to stand trial*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1990-55942-001)
- Gold, L. H. (1969). A statistical review of court-appointed psychiatric examinations. *Journal of Forensic Sciences, 14*(3), 294–301.
- Gold, L. H. (1973). Discovery of mental illness and mental defect among offenders. *Journal of Forensic Sciences, 18*, 125–129.
- Greenland, C., & Rosenblatt, E. (1972). Remands for psychiatric examination in Ontario, 1969-1970. *The Canadian Psychiatric Association Journal, 17*(5), 397–401.
- Heilbrun, K., Bennett, W. S., Evans, J. H., & Offutt, R. A. (1992). Assessing treatability in mentally disordered offenders: Strategies for improving reliability. *Forensic Reports, 5*(1), 85–96.
- Herbel, B. L., & Stelmach, H. (2007). Involuntary medication treatment for competency restoration of 22 defendants with delusional disorder. *Journal of the American Academy of Psychiatry and the Law, 35*(1), 47–59.
- Ho, T. (1999). Examination of racial disparity in competency to stand trial between White and African American retarded defendants. *Journal of Black Studies, 29*, 771–789. doi:10.1177/002193479902900602
- Ho, T., & Henderson, B. B. (1998). Relationship of psychological, demographic, and legal variables to court decisions of competency to stand trial among mentally retarded criminal defendants. *Journal of Criminal Justice, 26*, 307–320. doi:10.1016/S0047-2352(98)00013-0
- Hochstedler, E. (1987). Twice-cursed? The mentally disordered criminal defendant. *Criminal Justice and Behavior, 14*, 251–267. doi:10.1177/0093854887014003001
- Hoge, S. K., Poythress, N., Bonnie, R. J., Monahan, J., Eisenberg, M., & Feucht-Haviar, T. (1997b). The MacArthur adjudicative competence study: Diagnosis, psychopathology, and competence-related abilities. *Behavioral Sciences & the Law, 15*, 329–345. doi:10.1002/(SICI)1099-0798(199722/06)15:3<329::AID-BSL276>3.0.CO;2-Z
- ±Holmes, D. A. (1991). *Validation of the IFI with hospital staff judgments of competency to proceed to trial*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1992-75092-001)
- Hubbard, K. L., & Zapf, P. A. (2003). The role of demographic, criminal, and psychiatric variables in examiners' predictions of restorability to competency to stand trial. *International Journal of Forensic Mental Health, 2*(2), 145–155.
- Jacobs, M. S., Ryba, N. L., & Zapf, P. A. (2008). Competence-related abilities and psychiatric symptoms: An analysis of the underlying structure and correlates of the MacCAT-CA and the BPRS. *Law and Human Behavior, 32*, 64–77. doi:10.1007/s10979-007-9086-8
- ±Jones, B. L. (1995). *Factors related to the competency to proceed to trial of defendants who are mentally retarded*. (Doctoral Dissertation.) Retrieved from Dissertation Abstracts International. (1995-95003-133)
- ±Kennedy, J. D. (2008). *Diagnosis, instant offense, substance history, and trial competency: An explorative study*. (Doctoral Dissertation.) Retrieved from Dissertation Abstracts International. (2008-99080-361)
- Ladds, B. (1995). Involuntary electro-convulsive therapy to restore competency to stand trial: A five year study in New York State. *Journal of Forensic Sciences, 40*(2), 183–187.

(Appendices continue)

- Ladds, B., Convit, A., Zito, J., & Vitrai, J. (1993). The disposition of criminal charges after involuntary medication to restore competency to stand trial. *Journal of Forensic Sciences*, 38(6), 1442–1459.
- Lamb, H. R. (1987). Incompetency to stand trial: Appropriateness and outcome. *Archives of General Psychiatry*, 44(8), 754–758.
- Lewis, C. F., Fields, C., & Rainey, E. (2006). A Study of geriatric forensic evaluatees: Who are the violent elderly? *Journal of the American Academy of Psychiatry and the Law*, 34(3), 324–332.
- Manguno-Mire, G. M., Thompson, J. W., Shore, J. H., Croy, C. D., Artecona, J. F., & Pickering, J. W. (2007). The use of telemedicine to evaluate competency to stand trial: A preliminary randomized controlled study. *Journal of the American Academy of Psychiatry and the Law*, 35(4), 481–489.
- McGarry, A. L. (1965). Competency for trial and due process via the state hospital. *American Journal of Psychiatry*, 122, 623–631.
- McKee, G. R. (1998). Competency to stand trial in preadjudicatory juveniles and adults. *Journal of the American Academy of Psychiatry and the Law*, 26(1), 89–99.
- ◆ Menzies, R. J., Jackson, M. A., & Glasberg, R. E. (1982). The nature and consequences of forensic psychiatric decision-making. *Canadian Journal of Psychiatry*, 27(6), 463–470.
- ◆ Menzies, R. J., Webster, C. D., Roesch, R., Jensen, F. A. S., & Eaves, D. (1983). The Fitness Interview Test: A semi-structured instrument for assessing competency to stand trial, with a proposal for its implementation. *Medicine & Law*, 3, 151–162.
- Miller, H. A. (2004). Examining the use of the M-FAST with criminal defendants incompetent to stand trial. *International Journal of Offender Therapy and Comparative Criminology*, 48, 268–280. doi:10.1177/0306624X03259167
- Mossman, D. (2007). Predicting restorability of incompetent criminal defendants. *Journal of the American Academy of Psychiatry and the Law*, 35(1), 34–43.
- Mowbray, C. T. (1979). A study of patients treated as incompetent to stand trial. *Social Psychiatry*, 14, 31–39. doi:10.1007/BF00583571
- Nicholson, R. A., Barnard, G. W., Robbins, L., & Hankins, G. (1994). Predicting treatment outcome for incompetent defendants. *Bulletin of the American Academy of Psychiatry & the Law*, 22(3), 367–377.
- Nicholson, R. A., & Johnson, W. G. (1991). Prediction of competency to stand trial: Contribution of demographics, type of offense, clinical characteristics, and psycho-legal ability. *International Journal of Law and Psychiatry*, 14, 287–297. doi:10.1016/0160-2527(91)90009-C
- Nicholson, R. A., & McNulty, J. L. (1992). Outcome of hospitalization for defendants found incompetent to stand trial. *Behavioral Sciences & the Law*, 10, 371–383. doi:10.1002/bsl.2370100307
- Nussbaum, D., & Amaral, R. (2001). Internal reliability and item analysis of the MET-FORS Fitness Questionnaire. *American Journal of Forensic Psychology*, 19(2), 65–74.
- Paramesh, C. R. (1987). The Competency Screening Test: Application to a state maximum security hospital population. *American Journal of Forensic Psychology*, 5(1), 11–15.
- ± Patterson, H. R. (2005). *Length of stay in forensic inpatients adjudicated as incompetent to stand trial*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2005-99012-012)

(Appendices continue)

- ±Peacock, M. A. (2005). *Mental illness as a mediator of competent behavior in persons with mental retardation*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2005-99016-329)
- ±Pierrel, A. S. (1986). *Competency to stand trial and the mentally retarded defendant*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1987-53458-001)
- Quinsey, V. L., Pruesse, M., & Fernley, R. (1975). A follow-up of patients found unfit to stand trial or not guilty because of insanity. *The Canadian Psychiatric Association Journal*, *20*(6), 461–467.
- Redding, R. E. (1997). Depression in jailed women defendants and its relationship to their adjudicative competence. *Journal of the American Academy of Psychiatry and the Law*, *25*(1), 105–119.
- Redlich, A. D., Silverman, M., & Steiner, H. (2003). Pre-adjudicative and adjudicative competence in juveniles and young adults. *Behavioral Sciences & the Law*, *21*, 393–410. doi:10.1002/bsl.543
- Reich, J., & Wells, J. (1986a). Defendants with repeated competency evaluations. *Journal of Nervous and Mental Disease*, *174*, 120–122. doi:10.1097/00005053-198602000-00009
- Reich, J., & Wells, J. (1986b). The association between legal charge and psychiatric diagnosis in a competency-to-stand-trial population. *Social Psychiatry*, *21*, 206–209. doi:10.1007/BF00584002
- ±Richardson, H. A. (2005). *A community competency restoration program for defendants who do not meet commitment criteria*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2005-99020-025)
- ±Roach, R. L. (1994). *Assessment of trial competency in hospitalized criminal defendants*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1996-74338-001)
- Robbins, E., Waters, J., & Herbert, P. (1997). Competency to stand trial evaluations: A study of actual practice in two states. *Journal of the American Academy of Psychiatry and the Law*, *25*(4), 469–483.
- ◆Roesch, R., & Golding, S. L. (1977). *A systems analysis of competency to stand trial procedures: Implications for forensic services in North Carolina* (Pre-final draft). Urbana, IL: University of Illinois Press.
- Rogers, R., Jackson, R. L., Sewell, K. W., & Harrison, K. S. (2004). An Examination of the ECST-R as a Screen for Feigned Incompetency to Stand Trial. *Psychological Assessment*, *16*, 139–145. doi:10.1037/1040-3590.16.2.139
- Rogers, R., Jackson, R. L., Sewell, K. W., Tillbrook, C. E., & Martin, M. A. (2003). Assessing dimensions of competency to stand trial: Construct validation of the ECST-R. *Assessment*, *10*, 344–351. doi:10.1177/1073191103259007
- Rogers, R., Sewell, K. W., Grandjean, N. R., & Vitacco, M. (2002). The detection of feigned mental disorders on specific competency measures. *Psychological Assessment*, *14*, 177–183. doi:10.1037/1040-3590.14.2.177
- Rogers, R., Ustad, K. L., Sewell, K. W., & Reinhardt, V. (1996). Dimensions of incompetency: A factor analytic study of the Georgia court competency test. *Behavioral Sciences & the Law*, *14*, 323–330. doi:10.1002/(SICI)1099-0798(199622)14:3<323::AID-BSL243>3.0.CO;2-Z

(Appendices continue)

- ±Ryba, N. L. (2005). *An evaluation of the influence of psychiatric symptoms and cognitive abilities on competence-related abilities*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2005-99012-013)
- Shatin, L. (1979). Brief form of the Competence Screening Test for mental competence to stand trial. *Journal of Clinical Psychology, 35*, 464–467. doi:10.1002/1097-4679(197904)35:2<464::AID-JCLP2270350249>3.0.CO;2-I
- Siegel, A. M., & Elwork, A. (1990). Treating incompetence to stand trial. *Law and Human Behavior, 14*, 57–65. doi:10.1007/BF01055789
- ±Simon, A. (1999). *The restoration of competency to stand trial: An analysis of intervention efficacy*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1999-95018-281)
- Smith, S. A. (1996). Comparison of two screening tests of competency to stand trial for defendants with mental retardation. *Psychological Reports, 78*(1), 234.
- Smith, S. A., & Broughton, S. F. (1994). Competency to stand trial and criminal responsibility: An analysis in South Carolina. *Mental Retardation, 32*(4), 281–287.
- Smith, S. A., & Hudson, R. L. (1995). A quick screening test of competency to stand trial for defendants with mental retardation. *Psychological Reports, 76*(1), 91–97.
- Stoops, R., Hess, J., Scott, T., Tolan, J., Gallucci, G., & Brierley-Bowers, P. (2007). Training competency to stand trial in an individual with intellectual disability and behavioral health concerns. *Mental Health Aspects of Developmental Disabilities, 10*(2), 47–52.
- Ustad, K. L., Rogers, R., Sewell, K. W., & Guarnaccia, C. A. (1996). Restoration of competency to stand trial: Assessment with the Georgia Court Competency Test and the Competency Screening Test. *Law and Human Behavior, 20*, 131–146. doi:10.1007/BF01499351
- Vann, C. R., & Morganroth, F. (1965). The psychiatrist as judge: A second look at the competence to stand trial. *University of Detroit Law Journal, 43*(1), 1–12.
- Vernon, M., Steinberg, A. G., & Montoya, L. A. (1999). Deaf murderers: Clinical and forensic issues. *Behavioral Sciences & the Law, 17*, 495–516. doi:10.1002/(SICI)1099-0798(199910/12)17:4<495::AID-BSL361>3.0.CO;2-6
- Viljoen, J. L., Roesch, R., & Zapf, P. A. (2002). An examination of the relationship between competency to stand trial, competency to waive interrogation rights, and psychopathology. *Law and Human Behavior, 26*, 481–506. doi:10.1023/A:1020299804821
- Viljoen, J. L., & Zapf, P. A. (2002). Fitness to stand trial evaluations: A comparison of referred and non-referred defendants. *International Journal of Forensic Mental Health, 1*(2), 127–138.
- Vitacco, M. J., Rogers, R., Gabel, J., & Munizza, J. (2007). An evaluation of malingering screens with competency to stand Trial patients: A known-groups comparison. *Law and Human Behavior, 31*, 249–260. doi:10.1007/s10979-006-9062-8
- Wall, B. W., Krupp, B. H., & Guilmette, T. (2003). Restoration of competency to stand trial: A training program for persons with mental retardation. *Journal of the American Academy of Psychiatry and the Law, 31*(2), 189–201.
- ±Whittemore, K. E. (2001). *Releasing the mentally disordered offender: Disposition decisions for individuals found unfit to stand trial and not criminally responsible*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (2001-95002-139)

(Appendices continue)

- ◆ Wildman, R. W., II, Batchelor, E. S., Thompson, L., Nelson, F. R., Moore, J. T., Patterson, M. E., & de Laosa, M. (1978). *The Georgia Court Competency Test: An attempt to develop a rapid, quantitative measure of fitness for trial*. Unpublished manuscript, Forensic Services Division, Central State Hospital, Milledgeville, GA.
- Wygant, D. B., Sellbom, M., Ben-Porath, Y. S., Stafford, K. P., Freeman, D. B., & Heilbronner, R. L. (2007). The relation between symptom validity testing and MMPI-2 scores as a function of forensic evaluation context. *Archives of Clinical Neuropsychology*, 22, 489–499. doi:10.1016/j.acn.2007.01.027
- ± Zapf, P. A. (1999). *An investigation of the construct of competence in a criminal and civil context: A comparison of the fit, the MacCAT-CA, and the MacCAT-T. (adjudicative competence, treatment competence, Macarthur competence assessment tool)*. (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (1999-95024-036)
- Zapf, P. A., & Roesch, R. (2001). A comparison of MacCAT-CA and the FIT for making determinations of competency to stand trial. *International Journal of Law and Psychiatry*, 24, 81–92. doi:10.1016/S0160-2527(00)00073-X
- Zapf, P. A., & Roesch, R. (2005). An Investigation of the Construct of Competence: A Comparison of the FIT, the MacCAT-CA, and the MacCAT-T. *Law and Human Behavior*, 29, 229–252. doi:10.1007/s10979-005-2194-4

Received July 14, 2010

Revision received September 24, 2010

Accepted September 24, 2010 ■