Barry Rosenfeld,[1] Ph.D. and Kerri Ritchie,[2] B.A.

Competence to Stand Trial: Clinician Reliability and the Role of Offense Severity*

ABSTRACT: As research strategies for studying competence to stand trial evolve, interest in the decision-making process of clinicians and the factors that may influence the clinical process has grown. This study assesses the reliability of clinical opinions regarding the basis for competence decisions and the influence that severity of the offense has upon the clinical determinations of competence. One-hundred eighty-eight criminal defendants in an outpatient forensic clinic were evaluated by pairs of clinicians (psychiatrists and clinical psychologists). Each clinician was asked to complete a questionnaire regarding the defendant's overall degree of competence on a 0 to 10 rating scale, ability to understand charges and proceedings and ability to assist in one's defense, and the likelihood that the defendant was malingering. Results demonstrated a high degree of reliability not only in clinical determinations of competence (agreement in 187 of 188 cases), but in estimates of overall degree of competence and the basis for findings of competence or incompetence (e.g., the Dusky criteria). Misdemeanor defendants were more likely to be found incompetent to stand trial. However, degree of competence was only associated with offense severity for defendants found incompetent to stand trial. There was no relationship between these variables for competent defendants. These results suggest that clinicians may require a higher degree of competence for defendants charged with more serious offenses, although this process may not apply equally to competent and incompetent defendants. Furthermore, conducting competence interviews jointly appears to increase the concordance of competence opinions.

KEYWORDS: forensic science, forensic psychiatry competence to stand trial, reliability, criminal offense

Research regarding the construct of competence to stand trial has grown considerably over the past two decades. Although much of the past research has focused on methods for evaluating competence and identifying the correlates of competent and incompetent defendants, more recently studies have begun to explore the decision-making process of clinicians and the role of clinical, situational, and systemic factors that may influence the clinical process. This study addresses two such issues regarding competence to stand trial: the reliability of clinical opinions regarding the basis

for competence decisions and the influence of offense severity on competence determinations.

Reliability of Clinical Opinions

The first of these topics, the reliability of clinical determinations, has been the subject of little empirical research, and the research that has been offered has typically focused only on the determination of competence/incompetence. For example, Poythress and Stock described a study of criminal defendants evaluated by pairs of clinical psychologists (1). The clinicians agreed in each of the 44 competence assessments, indicating a high degree of reliability in clinical judgments. Goldstein and Stone also reported a high degree of concordance in opinions of competence to stand trial (CST) in an outpatient forensic clinic, with disagreements in only 35 of 1400 cases evaluated during the study year (2). This report, however, noted that a large proportion of defendants were referred for inpatient evaluation when their competence was questionable (24% of all referrals), likely inflating the degree of concordance found since disagreements are more likely among defendants whose competence is marginal.

Golding and colleagues, in their validation study of the Interdisciplinary Fitness Interview (IFI), reported a high degree of interrater reliability in competence decisions made by trained raters (mental health clinicians and attorneys) (3). Two raters agreed in 75 of 77 competence decisions using this structured competence interview, Kappa = .93. These authors also assessed the reliability of 5 subdomains of competence: appreciation of charges, relationship with one's attorney, attitude towards attorneys in general, anticipated courtroom demeanor and appreciation of legal options. However, the reliability of these judgments was quite modest (Kappa range: .42 to .58) even after omitting the 2 cases in which raters disagreed as to competence. Furthermore, this study, like several other studies that have addressed aspects of competence to stand trial, focused on establishing the reliability of trained raters in utilizing competence assessment tools rather than focusing on whether clinicians' opinions of competence were based on similar rationale (4). More recently, Skeem and Golding described the reliability of competence opinions based on 50 defendants evaluated by pairs of clinicians, finding that two evaluators agreed in 41 of 50 cases (Kappa = .64) (5). These authors assessed reliability along several dimensions of competence (appreciation of charges, relationship with attorney, capacity to testify, etc). However, their judgments of concordance were based on the information contained in the clinicians' reports (i.e., whether a particular aspect of competence was mentioned in each of 2 reports) rather than direct examination of clinical opinions.

[1]Forensic Psychiatry Clinic, New York City Criminal and Supreme Courts and Department of Psychology, John Jay College of Criminal Justice, New York, NY.

[2]Department of Psychology, John Jay College of Criminal Justice, New York, NY.

Competence and Offense Severity

In addition to the clinical/legal basis for competence determinations, many authors have suggested that situational or systemic factors may impact on competence determinations as well. One such factor that has generated somewhat confusing findings is the relationship between alleged offense and clinical findings of competence/incompetence. Clinicians have often suggested that higher levels of competence are required for more complex decisions or serious situations (e.g., 4,6). For example, the American Bar Association's Criminal Justice Mental Health Standards state that a competence "evaluation should consider a defendant's mental ability in relation to the severity of the charge and the complexity of the case (7)." Hence, one might expect findings of incompetence more often among defendants charged with more serious offenses or facing lengthy jail sentences if convicted. Indeed, Steadman found that homicide and other serious offenses were over-represented among incompetent defendants relative to the rate of such charges overall (8). However several other researchers have often found the reverse; that findings of incompetence are more common among defendants charged with minor or "nuisance" crimes (e.g., trespassing, criminal mischief, disorderly conduct). These findings have occasionally been interpreted as indication that the legal system may be used as a mechanism for obtaining mental health treatment for indigent or otherwise unwilling individuals (i.e., the "criminalization" of the mentally ill). Thus, while evidence has been offered in support of both of these seemingly contradictory hypotheses, no research has attempted to resolve these competing influences.

The existing research on the relationship between criminal charge and competence has fallen into two categories: studies that focus on offense severity (e.g., misdemeanor versus felony charge, violent versus non-violent offense) and studies that divide index offense (most serious charge) into discrete categories (e.g., homicide, property crimes, nuisance offenses). For example, Bittman and Convit reported that among a sample of 354 defendants assessed at an outpatient forensic clinic, defendants charged with misdemeanor offenses were nearly twice as likely to be found incompetent as defendants charged with felonies (no statistical tests were included in this report) (9). Johnson, Nicholson and Service, on the other hand, classified offense severity along a 19-point rating scale and found no relationship between competence and offense, although the authors did not describe their method for classifying offense severity (10). Studies addressing the relationship between violence (violent versus non-violent offense) and competence, however, have been more consistent, with several studies reporting no relationship between these two variables (11–14).

Research focusing on type of offense and competence determinations has been less clear. Four studies conducted in inpatient forensic settings have reported no relationship between the type of offense alleged and determinations of competence (15–18). Two of these studies, however, failed to describe the method for classifying offense and indicated their null findings in an incidental manner (16–17). Conversely, several large studies conducted in outpatient settings have found significant difference in the proportion of defendants found incompetent for various offense categories (19–21). In general, these studies have supported the conclusion that defendants charged with more serious offenses (e.g., homicide, sex offenses, robbery) were more likely to be found competent to stand trial while defendants charged with nuisance/public order offenses were more likely to be found incompetent. Warren and colleagues recently described a comparison study of forensic evaluations conducted in three different states (Ohio, Michigan, Virginia) (21). They found that defendants charged with homicide or sex offenses were twice as likely as other defendants to be found competent to stand trial, whereas defendants charged with "public order" offenses were nearly 3 times more likely to be judged incompetent.

The present study seeks to resolve these seemingly contradictory hypotheses and ascertain whether clinicians apply higher standards of competence to defendants charged with more serious offenses. In addition, the reliability of clinical opinions is examined by analyzing not only findings of competence/incompetence, but the basis for clinical determinations of competence based on the prongs outlined in Dusky v. U.S. (1960) (22).

Method

Subjects

Pre-trial criminal defendants evaluated with regard to competence to stand trial in the Forensic Psychiatry Clinic of the New York City Criminal and Supreme Courts between 1/1/96 and 6/30/96 were used in these analyses. The Forensic Psychiatry Clinic (FPC) is a division of Bellevue Hospital located within the Criminal Courts building and is the primary facility responsible for conducting court ordered forensic evaluations for New York County (Manhattan), New York. The clinic evaluates approximately 1200 criminal defendants annually, roughly 40% of whom are referred for evaluation of competence to stand trial (CST). In addition to the evaluations conducted in the FPC, a small number of Manhattan defendants (approximately 20–30 per year) are evaluated on an inpatient basis at Bellevue Hospital, typically after an unsuccessful attempt to complete the evaluation on an outpatient basis.

According to New York State law, an evaluation of a defendant's competence to stand trial (termed "fitness to stand trial" under New York State law) should be obtained whenever questions arise as to the defendant's ability to understand trial proceedings or participate meaningfully in their defense (23). The specific criteria for competence in New York State is modeled after the Dusky criteria used by all states and the Federal Government (22, 24).[3]

Unlike many states and jurisdictions, competence evaluations in New York State are primarily ordered by the trial judge, although defense attorneys may request a court-ordered evaluation or obtain an independent evaluation at their own or their defendant's expense (see 21, 24 for a description of interstate differences in CST evaluations). Once an evaluation has been ordered, the defendant must be evaluated by two independent clinicians (psychiatrists or clinical psychologists), although it is common practice for these evaluations to be conducted jointly.

Procedures

Upon completion of the CST evaluation, each clinician was asked to complete a brief questionnaire regarding the defendant's competence and psychiatric condition (see below). These questionnaires were completed independently, although they were typically

---

[3]Although a slightly expanded criteria for competence to stand trial is detailed in People v. Valentino (25), the New York Criminal Procedure Law follows the Dusky criteria (24). In addition, because all 50 states and the federal court system base determinations of competence on Dusky, these criteria were considered as the primary components of competence for this study.

based upon a single, joint interview. All of the 8 clinicians employed in the FPC (6 psychiatrists and 2 clinical psychologists) participated in the study. However, rates of completion of the study questionnaire varied (range: 42% to 100%, median: 60%). These clinicians had been employed in the FPC for an average of 11.8 years (range: 1 to 23 years, median: 12.5 years) and had each conducted hundreds of CST evaluations. There was no difference, on any variables, between cases in which clinicians did or did not complete the study questionnaire ($p < .05$), nor was there any difference among clinicians in the proportion of defendants found CST/IST, Chi-Square $= 7.96$, df $= 7$, $p =$ NS.

Additional data (e.g., demographic, offense data, legal history) was collected from the defendant's clinic record after the evaluation had been completed. Because the data collection process required no participation or modification in the evaluation procedure, and had no impact on the outcome of their evaluation, informed consent was not obtained. The study was approved by the Forensic Psychiatry Clinic's Research Review Committee.

Study Measures

Competence to stand trial was assessed in two ways: as a dichotomous variable, competent (CST)/incompetent (IST), and as a continuous variable based on clinician's ratings using a 0–10 numerical rating scale. In addition, each defendant's ability to understand charges/proceedings (understand) and assist in his/her defense (assist) was rated as a dichotomous variable (impaired/not impaired). For defendants who were considered IST, clinicians were asked to estimate the likelihood that the defendant could be restored to competence (likely, unsure, unlikely). Finally, evaluators rated the likelihood that defendants were malingering incompetence (unlikely, possible, probable), along with offering diagnostic judgments and psychiatric symptom ratings (reported elsewhere).

Data regarding the alleged instant offense was classified in several different ways including whether the index offense (most serious charge) was a felony or a misdemeanor and violent or nonviolent, based on the New York State Penal Code (26). Alleged offense was also divided into one of several discrete categories based on the index offense (most serious charge): homicide, sex offense, assault/crimes against others, robbery, drug offenses, nuisance offenses (e.g., subway fare evasion, trespassing, criminal mischief) and other offenses (e.g., possession of weapon). Finally, offense severity was further refined into two related variables. The first of these indices was based on the ordinal classification system used in New York State, in which each charge is classified as a Violation, a Misdemeanor (level A or B) or a Felony (levels A through E) (26). A second method of classifying offense severity was based on the median possible sentence according to state judicial sentencing guidelines (26). These sentencing guidelines incorporate the ordinal classification system described above, the defendant's past criminal history (number of prior felony convictions) and whether or not the alleged offense was violent or non-violent.

Several felony defendants had not been indicted prior to referral for evaluation. Thus, charges listed in the arrest record are not necessarily an accurate reflection of offense severity. For example, arrest charges might indicate attempted felony assault while the criminal complaint suggests that the defendant was merely waving his arms and gesturing at a passerby on the street. In cases where a felony defendant had not been indicted, arrest records were reviewed to determine whether less serious charges more accurately reflected offense severity (revisions were made in 7 of 37 cases).

Statistical Analyses

Data were analyzed using three related but somewhat different samples. One set of analyses utilized each of the 188 cases in which a completed evaluation was conducted in the clinic during the study period (the total sample of defendants evaluated in the clinic during the 6-month period). This sample was used for descriptive purposes, as well as to assess demographic and offense characteristics potentially associated with competence findings (e.g., using frequency analyses, Chi-square and Fisher's Exact Test statistics). A second sample was comprised only of cases for which both clinicians completed the study data collection forms. These paired ratings were used to assess the inter-rater reliability of clinicians' determinations of competence (e.g., using intraclass correlation coefficients and Kappa statistics). Finally, a sample was generated that included one study data collection form for each subject evaluated that had such data available (drawn at random for subjects in which 2 forms were completed). These data were used to evaluate the relationships between demographic/offense data and the different aspects of competence and clinical opinions rendered. This sample was used (rather than averaging ratings from the paired ratings) because of the possibility of discrepancies between the ratings of the 2 evaluators with regard to aspects of competence impaired, level of competence, etc. Although one data collection form was selected from each pair at random, an attempt was made to balance the proportion of ratings made by each clinician in order to avoid potential bias by over-representation of any one evaluator.

Results

Sample Characteristics

Of the 200 defendants referred for evaluation during the 6-month study period, clinicians were able to complete the evaluations for 188 defendants (12 were referred for inpatient evaluation). The following data are based on the 188 cases in which a clinical conclusion with regard to CST was offered. Of these 188 completed evaluations, the majority of defendants were male ($N = 168$, 89.4) and of African American descent ($N = 105$, 56.5%). Thirty-two defendants were Caucasian (17.2%), 46 were Hispanic (24.7%) and 3 were Asian (1.6%). Most defendants had never been married ($N = 131$, 72.4%), and spoke English fluently ($N = 167$, 88.8%), although 17 defendants were evaluated with the assistance of a Spanish-speaking interpreter (9.1%) and 4 required other interpreters (French, Chinese, 2.1%).

A copy of the criminal complaint, indicating the charges and a description of the alleged offense, was available prior to each evaluation. Cases in which the defendant was indicted with a felony charge typically included a copy of the NYSID report ($N = 92$, 48.9% of all cases, 94.7% of felony defendants), a computerized "rap" sheet indicating past arrests, convictions and sentences. When official records of legal history were not available, evaluators were forced to rely on self-report data. Among those defendants who either provided this information or were accompanied by official records, nearly half had prior felony convictions ($N = 71$, 43.8%).

Eighty-four defendants were charged with violent crimes (44.7%) and 104 with non-violent crimes (55.3%). The majority of defendants were charged with felony offenses ($N = 114$, 60.6%) while 74 (39.4%) were charged with misdemeanors. Only 77 of these 114 felony defendants, however, had been indicted on a felony charge and 7 were subsequently reclassified as misdemeanor

cases for the purposes of data analysis after a review of the criminal complaint (described above). Fifty-one of 188 defendants (27.3%) had been previously evaluated in the Forensic Psychiatry Clinic and 39 (20.9%) had been previously evaluated specifically with regard to competence to stand trial in the clinic. The number of prior evaluations ranged from 1 to 8 prior CST evaluations and 1 to 11 prior evaluations of any kind. Of the 188 evaluations completed during the study period, 2 study questionnaires were completed in 90 cases (47.9%, although some of these questionnaires contained missing data) and at least one questionnaire was completed in 138 cases (73.9%).

Concordance of Clinical Opinion

Of the 188 cases completed during the study period, the two clinicians' opinions agreed with regard to competence in 187 cases (99.5%), Fisher's Exact Test = 184.02, $p < .0001$, Kappa = .99, 95%CI = .97 − 1.0.[4] More detailed information with regard to the basis for competence findings (e.g., degree of competence, ability to understand charges or assist in one's defense) was not available for all cases. Therefore, subsequent analyses were based on only those cases in which two completed study questionnaires were available ($N = 90$).

With regard to the estimates of the defendant's degree of competence, there was a high degree of concordance between the two clinicians' ratings using the 0–10 point numerical rating scale, Spearman $r(N = 88) = .92$, $p < .0001$. The two ratings were within one point of one another in 68 of 88 cases, and differed by 3 points in only 1 case (in 19 cases the two ratings differed by 2 points). A more conservative measure of reliability, an intraclass correlation coefficient (one-way, random-effects model), yielded similarly positive results regarding the reliability of clinicians' competence ratings, ICC = .91 (27).

Evaluators also demonstrated relatively high rates of agreement as to whether defendants were able to understand the proceedings against them and/or assist in their defense. Of the 89 cases in which two opinions were available as to whether or not the defendant understood the charges and proceedings against him/her (one case was missing this data for one evaluator), the two clinicians agreed in 82 cases (92.1%), Fisher's Exact Test (1,89) = 62.4, $p < .0001$, Kappa = .83, 95% CI = .71 − .95. Clinicians agreed even more often with regard to the defendant's ability to assist in his defense, with agreement in 87 of 90 cases (96.7%), Fisher's Exact Test (1,90) = 78.4, $p < .0001$, Kappa = .93, 95% CI = .86 − 1.0.

The level of interrater agreement was somewhat more modest with regard to the likelihood of malingered incompetence. Of the 79 cases in which two ratings of the probability of malingering were available (data were missing in 21 cases), the two clinicians agreed in 71 cases (89.9%), Fisher's Exact Test (1,79) = 15.9, $p < .001$ Kappa = .45, 95% CI = .12 − .77 (defendants who were rated as "possible" or "probable" regarding malingering incompetence were collapsed into a single group because of the infrequent occurrence of suspected malingering). Forty-three defendants were considered incompetent of the 90 cases in which two evaluators had completed the study questionnaire, however only 31 cases (72.1%) included two ratings of the likelihood that the defendant could be restored to competence. The level of agreement in these opinions was also modest, with agreement between the

two clinicians in only 25 of 31 cases (80.6%), Fisher's Exact Test (1,31) = 7.52, $p < .02$; Kappa = .46, 95% CI = .11 − .81.

Variables Associated with Competence Determinations

Univariate analyses revealed no relationship between findings of competence to stand trial and any demographic variables studied (e.g., gender, race; see Table 1). Likewise, there was no relationship observed between competence and either specific offense categories (homicide, property offenses, drug, etc.) or violent/non-violent charges. Misdemeanor defendants, however, were significantly more likely to be found IST than felony defendants (55% versus 40%), Fisher's Exact Test (1,188) = 4.36, $p < .04$, although no relationship was observed between other measures of offense severity and clinical findings (see Table 2).

Clinician estimates of the defendants' degree of competence approximated a bimodal distribution, with roughly normal distribution of scores for both competent and incompetent defendants yet few ratings in the mid-range (4,5,6 on the 0–10 rating scale, see Fig. 1). These degree of competence ratings were significantly correlated with the measures of offense severity including the ordinal ranking based on the state Penal Code, Spearman $r(N = 137) = .25$, $p < .004$, and the median expected sentence based on judicial sentencing guidelines, $r(N = 137) = .19$, $p < .03$. The dichotomous categorization of felony versus misdemeanor defendants also revealed significant differences in competence ratings with felony defendants more likely to be considered to have a

---

4State statute requires that a third clinician evaluate any defendant whenever a disagreement exists with regard to competence. Although this process ultimately generated 3 opinions, only the first two were considered in the assessment of reliability.

TABLE 1—*Variables associated with CST/IST.*

Variable	CST $N = 101$ (53.7%)	IST $N = 87$ (46.3%)	Chi-Square (d f.)	p
Gender:			0.13(1)	n.s.
Male	91 (54.2.1%)	77 (45.8%)		
Female	10 (50.0%)	10 (50.0%)		
Race:			5.21(3)	n.s.
Caucasian	12 (37.5%)	20 (62.5%)		
African-American	57 (54.3%)	48 (45.7%)		
Hispanic	29 (63.0%)	17 (37.0%)		
Asian	2 (66.7%)	1 (33.3%)		
Marital status:			3.63(2)	n.s.
Single	70 (53.4%)	61 (46.6%)		
Married	14 (73.7%)	5 (26.3%)		
Separated/divorce/ widow	15 (46.9%)	17 (53.1%)		
Language:			1.59 (1)	n.s.
English	87 (52.1%)	80 (47.9%)		
Interpreter	14 (66.7%)	7 (33.3%)		
Prior CST evaluation:			0.81 (1)	n.s.
Yes	81 (80.1%)	20 (19.9%)		
No	68 (78.2%)	19 (21.8%)		
Charge:			4.36 (1)	.04
Felony	66 (60.0%)	44 (40.0%)		
Misdemeanor	35 (44.8%)	43 (55.2%)		
Violence:			0.01 (1)	n.s.
Violent	45 (53.6%)	48 (46.4%)		
Non-violent	56 (53.9%)	39 (46.1%)		
Offense category:			5.38 (7)	n.s.
Homicide	3 (60.0%)	2 (40.0%)		
Sex offense	4 (80.0%)	1 (20.0%)		
Assault	23 (48.9%)	24 (51.1%)		
Robbery	17 (56.7%)	13 (43.3%)		
Larceny	19 (59.4%)	13 (40.6%)		
Drug	14 (56.0%)	11 (44.0%)		
Nuisance	14 (45.2%)	17 (54.8%)		
Other	7 (53.8%)	6 (46.2%)		

TABLE 2—*Correlation between level of competence and offense severity.*

| Variable	Misdemeanor/ Felony	Offense Severity	Median Sentence
CST/IST finding (N = 188):	−0.14*	−0.10	−0.10
Overall sample (N = 139):			
Level of competence	0.26‡	0.24‡	0.21*
Understand proceedings	0.22†	0.19*	0.21†
Assist in defense	0.19*	0.15	0.17*
CST only (N = 75):			
Level of competence	0.09	0.14	−0.01
Understand proceedings	N/A	N/A	N/A
Assist in defense	N/A	N/A	N/A
IST only (N = 64):			
Level of competence	0.34†	0.37‡	0.33†
Understand proceedings	0.22	0.22	0.30*
Assist in defense	0.13	0.11	0.06

*$p < .05$; †$p < .01$; ‡$p < .005$.
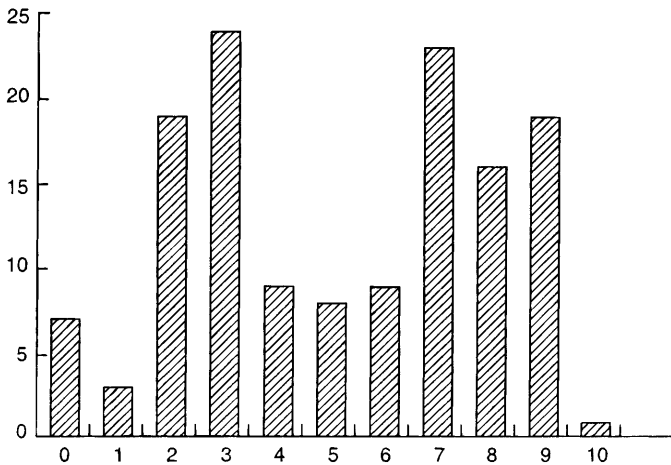
FIG. 1—*Clinician ratings of defendant's degree of competence.*

higher degree of competence (5.80 versus 4.44), $t(df = 137) = 2.90$, $p < .005$. There was no relationship between violent/nonviolent charges and estimated level of competence, $t(df = 133) = 0.26$, $p =$ NS.

Interestingly, the association between offense severity and clinician estimates of the defendant's degree of competence was considerably stronger among defendants found incompetent than for competent defendants. Among defendants found IST, there was a significant correlation between clinician's competence ratings and offense severity rankings based on the state Penal Code, $r(N = 62) = .37$, $p < .004$, median expected sentence based on state sentencing guidelines, $r(N = 62) = .32$, $p < .02$, and misdemeanor/felony charge, $r(N = 62) = .34$, $p < .006$. Among defendants considered CST, on the other hand, there was no significant association between level of competence and Penal Code classification, $r(N = 75) = .15$, $p =$ n.s., median expected sentence, $r(N = 75) = .04$, $p =$ n.s., or misdemeanor/felony charge, $r(N = 75) = .10$, $p =$ n.s.

Discussion

These data offer several insights regarding the process of competence to stand trial evaluations. First, we are encouraged by the high level of concordance in clinical opinions regarding competence to stand trial. The extremely high concordance observed in CST/IST decisions (99.5% agreement), however, may be misleading given the clinic practice of conducting joint evaluations and possible social pressure to agree on clinical findings (e.g., because disagreements regarding competence must be resolved by a third evaluator, these cases have a much greater likelihood of necessitating a court hearing). That is, evaluators may agree as to a finding of competence but have very different reasons for their decisions. For example, one evaluator may consider the defendant as competent as any defendant he/she has seen ("10" on a 0–10 rating scale), while the second considers the defendant only marginally competent ("6" on the 0–10 scale), yet both agree that the defendant is competent. These data, however, revealed no such hidden disagreements regarding competence decisions, either in estimating the defendant's degree of competence or whether they possessed the requisite abilities to proceed (e.g., ability to understand charges and proceedings, ability to assist in one's defense). Clinician ratings on the 0 to 10 numerical rating scale were highly consistent (reliability > .90), and the two ratings differed by 3 points in only 1 case. This high level of reliability is more noteworthy in light of the fact that clinicians, although quite experienced in conducting competence evaluations, were not "trained" in completing the study questionnaires or numerical rating scale, and therefore relied solely on their judgment and past experience.

Several factors may have bolstered the high degree of concordance in competence opinions observed in this study, including the practice of conducting interviews jointly. When two clinicians base their opinions on the same evaluation (i.e., observing the defendant in the same mental condition, answering the same interview questions, and behaving in the same manner), one would expect a greater degree of concordance of opinions than when the competence evaluations are conducted separately and differently, possibly separated by several weeks in which treatment or additional stressors may have occurred. It is also likely that clinicians continued their evaluation whenever an aspect of competence was unclear, thus increasing the information available to both clinicians and increasing the likelihood of reaching an apparent consensus. Alternatively, a clinician who is unsure which decision to render may be swayed by the apparent confidence of the second evaluator, particularly if that clinician has extensive experience in CST evaluations, or may even agree with the more confident clinician in an effort to avoid conflicts, embarrassment or disagreements. Nevertheless, the high level of concordance in opinions regarding CST/IST criteria (the Dusky criteria, degree of competence) suggests that clinicians generally agreed as to which, if any, aspects of competence were impaired and to what extent.

In addition to a high degree of inter-rater reliability in competence opinions, these data also support the hypothesis that a greater degree of competence is required of defendants charged with more serious offenses. Among defendants found incompetent to stand trial, there was a significant association between severity of charge (based on several different indices) and degree of competence estimated. This finding could be explained by clinicians' willingness to allow marginally fit misdemeanor defendants to proceed to trial whereas marginal felony defendants, facing lengthy jail sentences if incarcerated, would be found incompetent. Thus, one would expect "more competent" (a higher degree of competence on the 0–10 rating scale) incompetent felony defendants whereas only the most grossly impaired misdemeanor defendants to be found incompetent.

Surprisingly, no association between degree of competence and offense severity was found for defendants found competent to stand

trial since the above hypothesis would be expected to hold for both competent and incompetent defendants. One possible explanation for this lack of association between degree of competence and offense severity for defendants found competent to stand trial is the impact of study methodology. Clinicians rated each defendant's competence after they had concluded their clinical evaluation and therefore had already determined whether they believed the defendant was competent or incompetent. As a result, these competence ratings were likely "anchored" by the clinical determination already made (e.g., "The defendant is competent, therefore his competence must fall between 5 and 10"). Such a process is consistent with the roughly bimodal distribution of competence ratings observed in this study. This anchoring process may also be facilitated by the demands of judges and attorneys, who request absolute opinions regarding competence (competent or incompetent) rather than ambiguous findings. Thus, even if competence abilities were roughly normally distributed, the requirement that clinical opinions be framed in polar terms is likely to influence how clinician's view defendants' competence abilities.

Although one might expect an anchoring process to apply equally for both competent and incompetent defendants, such an assumption may not be accurate. Blashfield and colleagues, based on the results of their analogue study of competence decisions, suggested that clinicians may err on the side of considering defendants competent and this bias may alter the anchoring process apparent in this study (28). For example, clinicians may unconsciously adjust their estimates of competence for marginal misdemeanor defendants because of a belief that a finding of competence requires a certain minimum level of abilities while the same logic does not necessarily hold for marginal felony defendants found unfit. Regardless of the basis for these findings, the fact that the discrepancy in the competence/offense severity association was consistent across all 3 different measures of offense severity supports the validity of these findings.

The post-hoc nature of clinicians' competence ratings used in this study preclude a disentanglement of the possible impact of anchoring biases from other factors that may have influenced these results. Consequently, these data are more accurately characterized as an indication of how clinicians conceptualized the competence evaluation, rather than as a predictive model. Hence, while these data offer some insights into the clinical decision-making process in competence evaluations, further research is needed to better understand this complex process.

This study also revealed several apparent differences in the clinical evaluation process compared to data from other settings and forensic systems. Unlike most published studies, in which a relatively small proportion of defendants are found incompetent, nearly half of the defendants evaluated in New York City, were so judged (14). Although it is possible that this relatively high rate of incompetence reflects the idiosyncratic practice of one large clinic, rather than the impact of the referral practice in New York State or the prevalence of severe mental illness in the New York City criminal justice system, such an explanation appears unlikely. A more likely explanation is that this high rate of incompetence is related to the referral process, in which defendants are primarily referred by the court rather than by defense attorneys and the results of evaluations are made available to both the prosecution as well as the defense. Thus, defense attorneys have less motivation for requesting referrals merely for exploratory purposes. In addition, the volume of cases processed by the New York City courts may necessitate ap-

plication of more stringent criteria for referral (e.g., only referring those defendants obviously in need of evaluation). In fact, a study conducted in this clinic more than 20 years earlier reported having conducted approximately 1400 competence evaluations during the study ear compared to approximately 400 such cases annually during the previous times several years (2). Regardless of the explanation for this referral practice, application of more stringent criteria for referral minimizes the likelihood that needless competence evaluations would be requested, but may increase the possibility that incompetent defendants could go unidentified.

The evaluation of competence to stand trial is a complicated endeavor and the factors that impact on clinical decisions are numerous. Research identifying factors that increase, as well as decrease the reliability of clinical decisions can yield important information for understanding discrepancies in clinical opinions, as well as for improving the accuracy of opinions offered. Understanding how clinicians integrate case-specific information into their competence decisions may help clarify possible biases as well as allow more attention to be paid to previously overlooked factors. Given the importance of clinical determinations of competence and incompetence, such research continues to be an important and necessary endeavor.

References

1. Poythress NG, Stock HV. Competency to stand trial: A historical review and some new data. J Psychiatry Law 1980;8:131–46.
2. Goldstein RL, Stone M. When doctors disagree: Differing views on competency. Bull Am Acad Psychiatry Law 1977;5:90–7.
3. Golding SL, Roesch R, Schreiber J. Assessment and conceptualization of competency to stand trial. Law Human Behav 1984; 8(3/4):321–34.
4. Roesch R, Golding SL. Defining and assessing competency to stand trial. In: Geiner IB, Hess A, editors. Handbook of forensic psychology. New York: John Wiley, 1987;378–94.
5. Skeem JL, Golding SL. The logic and reliability of evaluations of competence to stand trial. Paper presented at the meeting of the American Psychological Association; 1995 August 11–15; New York, NY.
6. Grisso T. Evaluating competencies: Forensic assessments and instruments. New York: Plenum, 1986.
7. American Bar Association. Criminal Justice Mental Health Standards. Chicago: American Bar Association, 1984.
8. Steadman HJ. Beating a rap? Defendants found incompetent to stand trial. Chicago: University of Chicago, 1979.
9. Bittman BJ, Convit A. Competency, civil commitment, and the dangerousness of the mentally ill. J Forensic Sci 1993; 38(6):1460–6.
10. Johnson WG, Nicholson RA, Service NM. The relationship of competency to stand trial and criminal responsibility. Crim Justice Behav 1990;17(2):169–85.
11. Hart SD, Hare RD. Predicting fitness to stand trial: The relative power of demographic, criminal, and clinical variables. Forensic Reports 1992;5(2):53–65.
12. Heller MS, Traylor WJ, Ehrlich SM, Lester D. Is the issue of competence raised more frequently in crimes of violence? Am J Forensic Psychol 1986;4:22–6.
13. Nicholson RA, Johnson WG. Prediction of competency to stand

trial. Contribution of demographics, type of offense, clinical charac-teristics, and psycholegal ability. Int J Law Psychiatry 1991;14(3): 287–97.

14. Nicholson RA, Kugler KE. Competent and incompetent criminal defendants: A quantitative review of comparative research. Psycho-logical Bull 1991;109(3):355–70.

15. Daniel AE, Beck NC, Herath A, Schmitz M, Menninger K. Factors correlated with psychiatric recommendations of incompetence and insanity. J Psychiatry Law 1984;12(4):527–44.

16. Reich J, Wells J. Psychiatric diagnosis and competency to stand trial. Comprehensive Psychiatry 1985;26(5):421–32.

17. Roger R, Gillis JR, McMain S, Dickens SE. Fitness evaluation: A retrospective study of clinical, criminal, and sociodemographic characteristics. Canadian J Behav Sci 1988;20(2):192–200.

18. Roesch R, Eaves D, Stollner R, Normandin M, Glackman W. Evalu-ating fitness to stand trial: A comparative analysis of fit and unfit defendants. Int J Law Psychiatry 1981;4:145–57.

19. Hartman HL. Competency to stand trial. Am J Forensic Psychiatry 1984;5(2):63–75.

20. Warren J, Fitch L, Dietz PE, Rosenfeld B. Criminal offense, psychiatric diagnosis, and psycholegal opinion: An analysis of 894 Pretrial referrals. Bull Am Acad Psychiatry Law 1991;22(3): 379–88.

21. Warren J, Rosenfeld B, Fitch WL, Hawk G. Forensic mental health clinical evaluation: An analysis of interstate and intersystemic dif-ferences. Law Human Behavior. 1997;21(5):377–90.

22. Dusky v. United States, 362 U.S. 402, 1960.

23. New York Criminal Procedure Law 730.10.

24. Weiner BA: Mental disability and the criminal law. In Brakel SJ, Parry J, Weiner BA, editors. The mentally disabled and the law. Chicago: American Bar Foundation, 1985;693–801.

25. People v. Valentino, 356 N.Y.S.2d 962, 1974

26. New York State Penal Code, Article 55.05, 55.10

27. Shrout PE, Fleiss JL: Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979;86(2):420–8.

28. Blashfield RK, Robbins L, Barnard GW. An analogue study of the factors influencing competency decisions. Bull Am Acad Psychia-try Law 1994;22(4):587–94.