

# Response to Disclosure as an Indicator of Competence to Stand Trial

Douglas Mossman, MD, Rebecca M. Brown, DO, Bridget A. Casey-Leavell, DO, Christopher P. Marett, MD, MPH, and Elliot R. Lee, PhD, MD

Ethics guidelines recommend that forensic mental health professionals begin in-person assessments by explaining the nature and purpose of the examination. To learn whether evaluatees have understood and can give consent, forensic practitioners may ask evaluatees to paraphrase the explanation. This article explores how a forensic evaluatee's disclosure response (DR) reveals substantive information relevant to the purposes of a forensic examination. We examined archival data from 255 reports on competence to stand trial (CST) that a Midwest public sector hospital had previously submitted to courts. We classified each evaluatee's DR at one of three levels: DR = yes (accurate paraphrasing), DR = no (inability to paraphrase or provide a relevant response), or DR = other (an intermediate level implying a less-than-accurate response). None of the 28 DR = no evaluatees was CST, and only 7 (17%) of the 48 DR = other evaluatees were CST. Thus, a CST evaluatee who cannot paraphrase an examiner's explanation is likely to be incompetent to stand trial, and an examiner would need to adduce a strong argument to support any opinion to the contrary.

**J Am Acad Psychiatry Law 46:195–203, 2018. DOI:10.29158/JAAPL.003747-18**

In its *Ethics Guidelines for the Practice of Forensic Psychiatry*, the American Academy of Psychiatry and the Law (AAPL) tells psychiatrists, “At the outset of a face-to-face evaluation, notice should be given to the evaluatee of the nature and purpose of the evaluation and the limits of its confidentiality.”<sup>1</sup> Similar instructions appear in specialty guidelines for forensic psychologists, who should “strive to inform service recipients about the nature and parameters of the services to be provided” (Ref. 2, p 12). The obligation to provide notice flows from a commitment to the doctrine of informed consent, “one of the core values of the ethical practice of medicine and psychiatry. It reflects respect for the person, a fundamental principle in the practices of psychiatry and

forensic psychiatry.”<sup>1</sup> In medical practice generally, requiring consent prevents clinicians from imposing their views on patients and gives patients “the privacy to make decisions in line with their own values” (Ref. 3, p 356).

To obtain a forensic evaluatee's genuine assent or consent, an examiner needs to know that the evaluatee has understood the examiner's explanation of the examination's purpose and appreciates its significance. For this reason, forensic practitioners may invite evaluatees to ask questions after describing the nature and purpose of an examination.<sup>4</sup> Examiners may also ask evaluatees to paraphrase or answer questions about the information disclosed (as is done in obtaining informed consent for medical care)<sup>5,6</sup> and then correct any misconceptions that evaluatees hold.

The disclosure process for an evaluation of adjudicative competence, more often termed competence to stand trial (CST), usually includes an explanation of the reason for the evaluation, the nontreating nature of the examination, who has appointed or retained the examiner, the lack of confidentiality, who will receive the examiner's report, and the right of the evaluatee not to answer particular questions, with a warning that the examiner may have to report such noncooperation or refusal.<sup>7</sup> To paraphrase this information satisfactorily, the evaluatee must hear it, appre-

---

The coauthors are eternally grateful for their mentor and friend, Douglas Mossman, M.D., who passed away on January 4, 2018. Dr. Mossman was Professor of Psychiatry, Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati College of Medicine, Cincinnati, OH. Drs. Brown, Casey-Leavell, and Marett are psychiatrists at Summit Behavioral Healthcare, Cincinnati, OH. Dr. Marett is also Assistant Professor of Psychiatry, Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati College of Medicine, Cincinnati, OH. Dr. Lee is a Clinical Adjunct Assistant Professor, Department of Psychiatry, University of Wisconsin School of Medicine and Public Health, Madison, WI. Address correspondence to: Christopher Marett, M.D., M.P.H., Department of Psychiatry and Behavioral Neuroscience, U.C. College of Medicine, 260 Stetson Street, Suite 3200, Cincinnati, OH 45219. E-mail: marettcr@ucmail.uc.edu.

Disclosures of financial or other potential conflicts of interest: None.

ciate its bearing on the situation, and explain it rationally to the examiner; that is, the evaluatee must exercise the very sorts of mental faculties required to be CST. How much information about the defendant's competence, therefore, is revealed by the quality of a defendant's responses during the CST examination's disclosure process? Is the disclosure response (DR) itself an indicator or "test" (as it were) of a defendant's CST?

In an effort to answer these questions, this article describes a study of archival material developed under actual CST examination conditions and taken from actual CST reports submitted to criminal courts. Before undertaking the study, we anticipated that these reports would contain data that would validate or fail to prove the hypothesis that a defendant's responses during a CST examination's disclosure process provide substantive information about whether that defendant is competent or not.

## Method

### Setting

This study used archival data gathered from CST reports prepared at Summit Behavioral Healthcare, an Ohio public sector hospital operated by the Ohio Department of Mental Health and Addiction Services (OhioMHAS). Among the services provided by Summit Behavioral Healthcare are assessments of criminal defendants sent by courts for evaluations of their adjudicative competence<sup>8</sup> or for treatment to restore defendants found incompetent to stand trial (IST).<sup>9</sup> Pursuant to Ohio statutory requirements,<sup>10</sup> the hospital sends reports of the competence assessments to the region's criminal courts and retains copies of the reports in patients' records.

### Sample

Study data came from 255 reports that Summit Behavioral Healthcare had sent to courts during the years 2010–2015. These reports represented all the CST evaluations conducted by four psychiatrists who, when the evaluations took place, were forensic psychiatry fellows at an academic medical center. During their training, the fellows had examined hospitalized patients and prepared reports that served as the hospital's opinion on CST.

The study sample included 54 women, a proportion (21% of the full sample) fairly similar to recent

statistics about arrest rates across the United States<sup>11</sup> and in Ohio.<sup>12</sup> All sample members were adults whose mean (SD) age was 39.4 (12.9) years. The majority ( $n = 141$ ; 55%) were African American, 106 (42%) were white, and 8 (3%) were of other backgrounds (Latino, Asian, or African ancestry). These proportions are consistent with the ethnic make-up of the hospital's catchment area<sup>13,14</sup> and with the arrest ratio of African-American to white defendants for Ohio generally.<sup>12</sup> The sample's mean education level was  $10.9 \pm 2.4$  years; 143 (56%) of the evaluatees had either completed high school or had passed a General Educational Development (GED) test.

The evaluatees' most serious charges ranged from second-degree misdemeanors (e.g., obstructing official business and resisting arrest), to higher level felonies (e.g., felonious assault, rape), to homicides. Nearly one-fourth ( $n = 62$ ; 24%) of the sample's most serious charges were assaults, a proportion that rose to nearly one-third ( $n = 83$ ; 33%) if one includes domestic violence charges. More than half ( $n = 137$ ; 54%) the evaluatees faced misdemeanor charges only. In Ohio, incompetent misdemeanor defendants may spend no longer than 60 days undergoing competence restoration; for felony defendants, restoration may last up to one year.<sup>15</sup> Also, 23 (9%) of the evaluatees had undergone hospitalization for purposes of a 20-day inpatient CST evaluation.<sup>16</sup> These factors account for the relatively short median length of stay (42 days; intraquartile range, 28–118 days) at the time of evaluation.

The four psychiatric conditions that most commonly underlay evaluatees' possible or adjudicated incompetence to stand trial were schizoaffective disorder-bipolar type ( $n = 77$ ; 30%), schizophrenia ( $n = 72$ , 28%), bipolar disorders ( $n = 50$ ; 20%), and other psychoses ( $n = 17$ ; 7%). Substance use disorders was in fifth place ( $n = 8$ ; 3%), but 49 evaluatees (19%) had substance use disorders in addition to the principal psychiatric condition related to possible incompetence. Only three (1%) of the evaluatees were thought to be malingering, a rate much lower than is reported for initial CST evaluations.<sup>17,18</sup> In our sample, however, all the evaluatees had undergone at least one prehospitalization CST evaluation by other clinicians at local forensic centers, which likely had filtered out many court-referred defendants who were malingering.

## Measurement

In response to our inquiries, the four original psychiatric examiners confirmed that at the outset of every CST evaluation, they had tried conscientiously to complete a disclosure process that included a description of the examination's purpose and nonconfidentiality, the parties that would receive the CST report, the potential use of the report in determining case disposition, and the nontreatment purpose of the examination; and asking the evaluatee to paraphrase this explanation to assess understanding of the description.

The original examiners confirmed that they had accurately summarized evaluatees' DRs in all their CST reports. These reports also included narrative summaries of evaluatees' personal background, educational attainment, employment history, legal history, psychiatric and other medical history, hospital course, mental status findings, and responses to inquiries specific to CST. Further, examiners documented their findings from administrations of assessment instruments (e.g., the Georgia Court Competency Test (GCCT)<sup>19</sup> or MacArthur Competence Assessment Tool-Criminal Adjudication (MacCAT-CA)<sup>20</sup>) that they had used. The reports concluded with the examiners' opinions on psychiatric diagnosis and CST; the latter was assessed against the jurisdiction's statutory and case law criteria for competence to stand trial.<sup>21,22</sup> In Ohio, a defendant is presumed competent to stand trial and only found incompetent to stand trial if "the defendant is incapable of understanding the nature and objective of the proceedings against the defendant or of assisting in the defendant's defense."

## Data Collection Procedure

We undertook this archival study after receiving approvals from the Institutional Review Boards of the University of Cincinnati and Summit Behavioral Healthcare. From each CST report prepared by the original four examiners, three of this article's authors [R.B., B.C.L., C.M.] extracted data that included: evaluatees' demographic features (e.g., age, sex, race/ethnicity, educational attainment); reason for hospitalization (i.e., for CST evaluation or restoration); length of hospital stay when the examination occurred; offense type for the most serious charge; the examiner's diagnoses; scores on assessment instruments (e.g., the GCCT); the examiner's opinion on

CST; and a classification (described in the next subsection) of the evaluatee's DR.

Data were recorded on an Excel spreadsheet with identifiers removed such that no information could be linked to any individual evaluatee. To verify that the extractors' classifications were reasonable, the first author [D.M.] reviewed (but in no case changed) the original data on which extractors' had based "DR = no" and "DR = other" ratings. (The DR = yes ratings came from reports with highly consistent descriptions of evaluatees' responses.)

## Data Analysis

We compared the evaluatees whom examiners had opined CST and IST along several demographic, treatment, and assessment variables. We also compared evaluatees who fell into each of the three DR classifications described above. For ordered values, we used Mann-Whitney U tests; for categorical values, we used chi-square tests or (when expected categorical values were below 10 subjects) Fisher's exact test. In evaluating significance, we corrected for multiple comparisons using the Holm-Bonferroni method.<sup>23</sup>

To quantify what the evaluatees' DRs said about their CST status, we needed to examine the DR mathematically as though it were a diagnostic test of CST. As a first step, we assigned each evaluatee's DR to one of three categories:

DR = yes. This category designated evaluatees whom the examiners described as having paraphrased the disclosure statement satisfactorily before agreeing to participate. Typically, the examiners reported this by writing something very close to the following: "To see whether Mr. A had understood, I asked him to paraphrase what I had told him about the nature and purpose of the interview. Aided by some prompting questions, he did so accurately, and he agreed to speak with me." Here, "prompting questions" referred to portions of the disclosure that the evaluatee had not paraphrased spontaneously. If, for example, the evaluatee had not mentioned who would receive the report, the examiner might then ask, "Did I say anything about who gets my report?" and await the evaluatee's response. An evaluatee who (for example) responded, "Oh yeah, you said the court, my lawyer, and the prosecutor" would then be described in the report as having paraphrased accurately.

DR = no. This category contained evaluatees who, examiners reported, offered little or no response, who gave responses that were irrelevant or too disorganized or chaotic to understand or who could not remember or paraphrase more than an element or two of the disclosure information. Examples of examiner descriptions that led to DR = no categorizations included:

“[The evaluatee] did not acknowledge my presence. I interpreted his ignoring me as an indication that he would not participate in an interview.”

“[The evaluatee] asked, ‘What am I going to court for?’ . . . I could not elicit [further] responses.”

“[The evaluatee] responded, ‘I need to meet and then go home.’”

“[The evaluatee] laughed briefly, but he did not talk, gesture, or make any other response.”

DR = other. This category contained evaluatees for whom we could not neatly categorize responses as clearly DR = yes or DR = no. Responses in this category tended to indicate problems with grasping information disclosed or the presence of symptoms (e.g., paranoia) that affected the evaluatee’s participation in an interview. DR = other responses did not indicate complete inability to grasp the interview’s purpose, however. To illustrate, we provide a few examples taken from the examiners’ reports:

“He could not restate the information, but the answers he provided to my questions showed some grasp of what I had explained.”

“He could not restate the information, but his answers to my questions showed adequate understanding of this information.”

“She asked me what court and where the court was located. She asked me to repeat my answers a few times. When she restated the disclosure, she did it accurately and agreed to speak with me.”

“She said, ‘It’s for competent to stand trial; it’s between me and you.’ I again explained that the evaluation was not confidential, and she agreed to participate.”

“He began to speak rapidly and tangentially about his case, and said, ‘If they don’t show up, throw the case out of court.’ After a few minutes, I asked him again if he could tell me why we were meeting. Mr. A said that it was because he was found incompetent to stand trial, and he agreed to participate in the evaluation.”

We assigned the numbers 1 to DR = no, 2 to DR = other, and 3 to DR = yes to reflect our assumption that this order represented decreasing likelihood that the evaluatee was IST.

We wrote OpenBUGS<sup>24</sup> code to generate Bayesian estimates of accuracy parameters, examining how the DR performed as a diagnostic test of CST from three perspectives:

*Opinion as truth*, which treated the examiners’ CST opinions, formulated based on all data available to the examiner, as the criterion for whether the evaluatee was indeed CST or IST.

*Opinion as an imperfect gold standard*, which treated the examiners’ CST opinions as a very accurate but not perfect criterion (or gold standard) for the true CST-IST status of the evaluatees.

*Agnostic*, which adopted latent class analysis techniques use in other recent studies of forensic data<sup>25,26</sup> and are applicable when data include test results from subjects who have undergone evaluation for a condition with more than one diagnostic method. Here, four findings potentially differentiated between IST and CST evaluatees: the evaluatees’ DRs, the examiners’ CST opinions, whether the evaluatee had a diagnosis of schizophrenia or schizoaffective disorder (designated  $\Sigma+$ ) or not ( $\Sigma-$ ),<sup>27</sup> and evaluatees’ scores on the GCCT.<sup>19,28</sup>

We calculated accuracy indices for the DR categorizations and, for comparison purposes, the  $\Sigma+/\Sigma-$  distinction and GCCT scores. These indices included the area under the receiver operating characteristic (ROC) curve (AUC), which is a global measure of classificatory performance. In this context, the AUC equals the probability that, if an IST defendant and a CST defendant were chosen at random from all IST and CST defendants, the diagnostic method would correctly classify the defendants. (This interpretation of the AUC assumes that if the two randomly chosen evaluatees had tied scores, the tie would be resolved by flipping a coin to decide which way to classify the evaluatees.)

We also calculated another accuracy index: the stratum-specific likelihood ratio (SSLR). One formulation of Bayes’ theorem expresses the posterior odds of being IST as the product of the pretesting or prior odds of being IST and the SSLR associated with  $T_k$ , a test result in category  $k$ , or  $ODDS_{\text{post}} = ODDS_{\text{prior}} \times SSLR_k$ . The SSLR equals the slope of the portion of the ROC curve that corresponds to a particular test result category. On an ROC graph, test result categories are demarcated by points that



**Table 1** Comparisons of Defendants Opined CST or IST

	CST ( <i>n</i> = 126)	IST ( <i>n</i> = 129)	<i>p</i>
Age (mean ± SD)	36.4 ± 12.5	42.3 ± 12.6	0.00024*
Sex			
Men	94	107	0.14 <sup>†</sup>
Women	32	22	
Ethnicity			
African American	70	71	0.43 <sup>‡</sup>
White	54	52	
Other	2	6	
Length of stay (mean ± SD)	68.4 ± 69.5	96.9 ± 185.7	0.42*
Education			
≥12 Years	54	55	0.81 <sup>§</sup>
GED	17	21	
<12 Years, no GED	55	53	
Most serious charge			
Misdemeanor	60	77	0.071 <sup>†</sup>
Felony	66	52	
GCCT score (mean ± SD) ( <i>n</i> = 177)	88.7 ± 6.3	57.2 ± 17.5	~10 <sup>-26</sup> *
MacCAT-CA scores (mean ± SD)			
Understanding ( <i>n</i> = 24)	13.7 ± 3.6	8.3 ± 5.4	0.021*
Reasoning ( <i>n</i> = 23)	14.2 ± 3.7	8.4 ± 5.5	0.012*
Appreciation ( <i>n</i> = 23)	11.4 ± 2.7	6.0 ± 4.8	0.025*
Diagnosis			
Schizophrenia	27	45	~10 <sup>-5</sup> §
Schizoaffective disorder	28	49	
Others	71	35	
Medications when evaluated			
Antipsychotics	104	108	0.92 <sup>†</sup>
Mood stabilizers	51	56	0.73 <sup>†</sup>
Antidepressants	34	19	0.024 <sup>†</sup>
Others	68	70	0.92 <sup>†</sup>
None	12	14	0.89 <sup>†</sup>
Disclosure response			
DR = yes	119	60	~10 <sup>-19</sup> †
DR = other	7	41	
DR = no	0	28	

\* Mann-Whitney U test.

<sup>†</sup>  $\chi^2$  test, *df* = 1.

<sup>‡</sup> Fisher's exact test.

<sup>§</sup>  $\chi^2$  test, *df* = 2.

represent (*fpr*, *tpr*) pairs of false-positive and true-positive rates. One calculates  $SSLR_k$  from the (*fpr*, *tpr*) pairs as follows:

$$SSLR_k = \frac{tpr_{c-1} - tpr_c}{fpr_{c-1} - fpr_c}$$

where  $c = \{1, \dots, K\}$ ,  $c = \{1, \dots, K-1\}$  correspond to the  $K-1$  (*fpr*, *tpr*) pairs that delimit the  $K$  result categories,  $fpr_c = 0$  and  $tpr_c = 0$  when  $c = K$ , and  $fpr_{c-1} = 1$  and  $tpr_{c-1} = 1$  when  $c = 1$ .

The rationale for undertaking the imperfect gold standard and agnostic analyses was to evaluate explicitly the accuracy of the examiners' opinions on CST and IST. These two analyses showed, however, that the examiners were almost perfectly accurate. Further, all three methods for quantifying diagnostic

performances of the DR, the  $\Sigma+/\Sigma-$  distinction, and GCCT scores produced similar results. The Results section that follows therefore presents just the findings from the "opinion as truth" analyses.

## Results

Table 1 presents univariate comparisons of the evaluatees whom the examiners thought were CST and IST. IST evaluatees tended to be older, to have lower scores on forensic assessment instruments, to be more likely to have diagnoses of schizophrenia and schizoaffective disorder, and to have had trouble paraphrasing the initial disclosure.

Although taking an antidepressant medication appeared to be an indicator of being CST, a statistical

## Response Disclosure as an Indicator of Trial Competence

**Table 2** Examination Findings Construed as Diagnostic Tests of Incompetence to Stand Trial

Type of Finding	Category	SSLR	AUC
Diagnosis	Σ+	1.6 (1.3–2.1)	0.64 (0.58–0.70)
	Σ−	0.49 (0.35–0.67)	
GCCT score	92–100	0.045 (0.0076–0.15)	0.95 (0.91–0.97)
	82–90	0.13 (0.065–0.23)	
	72–80	1.5 (0.83–2.8)	
	62–70	55 (9.5–4300)	
	≤60	5000 (130–10 <sup>9</sup> )	
Disclosure response	DR = no	10 <sup>6</sup> (41–10 <sup>14</sup> )	0.73 (0.68–0.78)
	DR = other	5.2 (2.5–12)	
	DR = yes	0.52 (0.42–0.63)	

Data are median stratum-specific likelihood ratio (SSLR) and area under the ROC curve (AUC), with 95% confidence intervals. SSLRs and AUCs are rounded to two significant digits. Σ+, diagnosis of schizophrenia or schizoaffective disorder; Σ−, other diagnoses.

adjustment for multiple comparisons suggested that this association was coincidental.

Table 2 describes the diagnostic performance of three types of examination findings in differentiating between competent and incompetent evaluatees: having schizophrenia or schizoaffective disorder (Σ+) or not (Σ−), GCCT scores (which, as Table 1 shows, were available for almost 70% of evaluatees), and DR. We present median values with 95 percent Bayesian credible intervals in Table 2, rather than means ± SDs, because the Bayesian posterior distributions for the *SSLRs* were highly skewed, which made medians better estimators of central tendency. Table 2 shows that around 95 percent of the time, a randomly chosen IST evaluatee would have a lower GCCT score than a randomly chosen CST evaluatee. Similarly, about three-fourths of the time, a randomly chosen defendant who turns out to be incompetent will have a poorer DR than a randomly chosen defendant who turns out to be competent.

We illustrate use of the *SSLRs* in Table 2 with some numerical examples. Suppose that Mr. Jones, a CST evaluatee, gave a response to the disclosure that would be classified DR = other. Recent reports estimate that about one-fourth of defendants referred for CST evaluations are actually incompetent.<sup>29</sup> If this estimate appeared applicable to the population from which Mr. Jones came, then the prior odds of being IST would be 1:3. Table 2 shows that for DR = other, the *SSLR* was 5.2. Thus, the postdisclosure odds of being IST would be  $5.2 \times 1/3 = 1.73$ , or 1.73:1, and the postdisclosure probability of being CST =  $\frac{1.73}{1.73 + 1} \approx 0.63$ , or 63 percent. Similar calculations show that a rating of DR = no would

make it highly likely that the evaluatee was IST. Finally, a rating of DR = yes would reduce the odds of being IST by about one-half, so that the posttest odds of being IST would be 1:5.77, and the posttest probability of being IST =  $\frac{1}{5.77 + 1} \approx 0.15$ , or 15 percent.

Finally, Table 3 compares the members of the three DR groups. Although the DR = yes group appeared slightly younger than the other two groups, this difference probably is not significant when corrected for multiple comparisons. The only significant differences are those that involve the diagnoses rendered by the examiners and the evaluatee's GCCT score. That DR = yes was associated with having bipolar disorder makes sense, in that persons with affective illnesses tend to have less cognitive impairment and better rates of competence restoration than do persons with schizophrenia-spectrum disorders.<sup>21,27,30</sup> Similarly, one would expect that individuals who could do relatively well on a forensic assessment instrument such as the GCCT would assimilate the DR and paraphrase it satisfactorily.

## Discussion

This study showed that the disclosure process that begins a CST examination foreshadows later findings from the examination and the evaluatee's CST status itself. In particular, an evaluatee's inability to paraphrase the examiner's disclosure is a strong indicator of incompetence to stand trial. Thus, an examiner learns much about the evaluatee's competence, even before the evaluatee has agreed to disclose information relevant to CST.

**Table 3** Comparisons of Defendants According to Their DR

	DR = no ( <i>n</i> = 28)	DR = other ( <i>n</i> = 48)	DR = yes ( <i>n</i> = 179)	<i>p</i>
Median age (IQR)	43.0 (31.3–55.6)	41.2 (31.6–53.3)	36.1 (26.8–48.2)	0.029*
Sex				
Men	23	42	136	0.30 <sup>†</sup>
Women	5	6	43	
Ethnicity				
African American	20	26	95	0.14 <sup>‡</sup>
White	6	19	81	
Other	2	3	3	
Median LOS (IQR)	42 (21–51)	40.5 (28–145)	42 (29–117)	0.50*
Education				
≥12 Years	11	16	82	0.58 <sup>§</sup>
GED	3	7	28	
<12 years, no GED	14	25	69	
Most serious charge				
Misdemeanor	18	26	93	0.60 <sup>†</sup>
Felony	10	22	86	
Median GCCT (IQR)	54 (49–59)	60 (42–69)	86 (72–92)	~10 <sup>-8*</sup>
Examiner's diagnosis				
Schizophrenia	9	15	48	0.0039 <sup>§</sup>
Schizoaffective disorder	12	16	49	
Other psychoses	3	7	7	
Bipolar disorder	2	2	46	0.0023 <sup>  </sup>
Others	2	8	29	

(*n* = 177). IQR, interquartile range; LOS, length of stay in days.

\* Kruskal-Wallis test, 2 *df*.

<sup>†</sup>  $\chi^2$  test, *df* = 2.

<sup>‡</sup>  $\chi^2$  test, *df* = 2 (comparing African Americans and Caucasians).

<sup>§</sup>  $\chi^2$  test, *df* = 4.

<sup>||</sup>  $\chi^2$  test, *df* = 1 (comparing bipolar disorder to all other diagnostic groups).

Based on our findings, the DR appears to be a fairly specific indicator of being IST, though not a sensitive one. That is, less-than-satisfactory paraphrasing of the disclosure greatly heightened the odds that the evaluatee was IST, but (contrary to what we had expected) many evaluatees who could paraphrase a disclosure satisfactorily still were not CST. Of course, a poor disclosure response should not be the sole basis for an examiner's opinion on adjudicative competence. Yet our findings suggest that examiners should provide substantial data to support and documentation to justify any opinion that an evaluatee who paraphrased poorly was nonetheless competent to stand trial.

Our findings come with several important qualifications. Our efforts at quantification were limited by our study's retrospective design and our use of archival data that were not collected with research goals in mind. Although fellowship supervisors had instructed the original four examiners to describe evaluatees' DRs faithfully in their reports, the examiners did not use any prespecified scoring system or description formats for the DR that may be evaluated for interrater reliability. Nor did the examiners use

standardized wording for their disclosures; consistent with good forensic practice, examiners individualized how they conveyed information to optimize individual evaluatees' comprehension. Although the article's first author reviewed his coauthors' DR = no and DR = other ratings to confirm that their classifications were reasonable, we did not conduct formal tests of interrater reliability among the data gatherers.

Despite these limitations, we believe our findings constitute a useful view of what a poor DR signifies: likely incompetence. Supporting this conclusion are findings from recent research showing that the Montreal Cognitive Assessment,<sup>32</sup> the Repeatable Battery for the Assessment of Neuropsychological Status,<sup>33</sup> and the Bender Gestalt Test<sup>34</sup> identify individuals whose cognitive impairments limit their competence and capacity to be restored. Moreover, our findings should not surprise individuals who are familiar with the disclosure process for a forensic CST examination, although the low sensitivity and high specificity of the DR may. After all, navigating the disclosure process successfully requires that a CST evaluatee exercise mental faculties that also are relevant to adjudicative competence, including atten-

tion, orientation, verbal comprehension, memory, reasoning, executive functioning, and aspects of social cognition.<sup>35</sup> Reports by CST examiners may therefore include a poor DR as one of several findings that support an opinion that a defendant-evaluee is IST.

A not-yet-mentioned limitation relates to our using the original four examiners' CST opinions as the criterion for whether an evaluee actually was CST or IST. Yet, this choice makes sense in light of reasons articulated in a previous study of CST restoration<sup>27</sup>: in all but a handful of cases, the criminal courts accepted the hospital's opinion based on the examiner's report; using the examiner's opinions provided a truth criterion that was more uniform across evaluees than would be the opinions from dozens of judges; judges (not examiners) could be wrong; and judges' opinions may differ because of posthospitalization changes in the evaluees' mental status. Further, our imperfect gold standard and agnostic analyses showed that treating the examiners' opinions as the truth did not distort estimates of accuracy.

Readers may wonder whether evaluees' DRs led examiners to reach premature closure about evaluees' competence. If this were the case, then hearing evaluees' DRs at the outset of the examination would have caused rather than merely presaged the examiners' final opinions; the examiners would have succumbed to what cognitive psychologists call confirmatory bias.<sup>36</sup> Our findings tend to dispel this notion, however. As Table 3 shows, obvious features of the evaluees known to the examiners before the interview predicted neither how examiners viewed the DRs nor their ultimate points of view on CST: their age, sex, ethnicity, length of hospital stay, and charges. The GCCT score, obtained *after* the disclosure, tended to confirm what the DR suggested. The GCCT also was a much better gauge of CST than was the disclosure response, as implied by its much higher AUC (see Table 2).

If an evaluee is too impaired to understand why a CST examination is taking place or otherwise cannot give consent to the examination, a court order or the express permission of the evaluee's defense attorney provides acceptable legal authorization for a CST examination to continue.<sup>7</sup> In all the CST examinations that generated our study data, examiners had acted pursuant to court authorizations grounded in Ohio's statutory provisions.

Nonetheless, our findings show that seeking evaluees' consent undermined the ethical rationale for disclosure (fostering and respecting self-determination regarding use of one's personal information) by eliciting clinically and legally significant findings relevant to the evaluees' mental functioning before they had actually agreed to have this information exposed. For evaluees who preferred not to convey such information, the examiners' disclosures and efforts to obtain consent contravened the evaluees' wishes. Statutory provisions usually provide legal justification for CST evaluations of defendants who cannot give valid consent to participate, and by helping courts to identify defendants who are too impaired to assist counsel or understand their legal proceedings, forensic examiners provide an ethically valuable service to those defendants themselves. But when performing other types of forensic evaluations, undertaking a careful consent process may help examiners identify defendants who should not undergo further assessment absent an express authorization from the court or the evaluee's attorney.

## References

1. American Academy of Psychiatry and the Law: Ethics Guidelines for the Practice of Forensic Psychiatry, May 2005. Available at: <http://www.aapl.org/ethics.htm/>. Accessed November 30, 2016.
2. American Psychological Association: Specialty Guidelines for Forensic Psychology. *Am Psychol* 68:7–19, 2013
3. Wilson J: Is respect for autonomy defensible? *J Med Ethics* 33: 353–6, 2007
4. American Academy of Psychiatry and the Law: AAPL Practice guideline for forensic psychiatric evaluation of defendants raising the insanity defense. *J Am Acad Psychiatry Law* 42(4 suppl):S3–76, 2014
5. Janofsky J, McCarthy R, Folstein M: The Hopkins Competency Assessment Test: a brief method for evaluating patients' capacity to give informed consent. *Hosp Community Psychiatry* 43: 132–6, 1992
6. Appelbaum PS: Assessment of patients' competence to consent to treatment. *N Engl J Med* 357:1834–40, 2007
7. Mossman D, Noffsinger S, Ash P, et al: AAPL Practice guideline for the forensic psychiatric evaluation of competence to stand trial. *J Am Acad Psychiatry Law* 35:S3–72, 2007
8. Ohio Rev. Code § 2945.371(D)
9. Ohio Rev. Code § 2945.38(B)(1)
10. Ohio Rev. Code § 2945.38(F-G)
11. Federal Bureau of Investigation: Uniform Crime Report, Crime in the United States, 2012. Available at: <https://ucr.fbi.gov/crime-in-the-u.s/2012/crime-in-the-u.s.-2012/>. Accessed December 7, 2016
12. Office of Criminal Justice Services: Ohio Criminal Justice Statistics. Columbus, OH: Ohio Department of Public Safety, 2011. Available at: [http://www.publicsafety.ohio.gov/links/ocjs\\_Statistics.pdf](http://www.publicsafety.ohio.gov/links/ocjs_Statistics.pdf). Accessed December 9, 2016
13. United States Census Bureau: Quick facts: Ohio. Statistical Atlas. Washington DC: U.S. Department of Commerce, Bureau of Sta-



- tistics, 2017. Available at <https://www.census.gov/quickfacts/table/PST045215/39/>. Accessed December 7, 2016
14. Race and ethnicity in Cincinnati, Ohio. Statistical Atlas. San Francisco, CA: Cedar Lake Ventures, April 2015. Available at <http://statisticalatlas.com/place/Ohio/Cincinnati/Race-and-Ethnicity/>. Accessed December 7, 2016
  15. Ohio Rev. Code § 2945.38(C)
  16. Ohio Rev. Code § 2945.371(C–D)
  17. Mittenberg W, Patton C, Canyock EM, *et al*: Base rates of malingering and symptom exaggeration. *J Clin Exp Neuropsychol* 24: 1094–102, 2002
  18. Rogers R, Salekin RT, Sewell KW, *et al*: A comparison of forensic and nonforensic malingerers: a prototypical analysis of explanatory models. *Law & Hum Behav* 22:353–67, 1998
  19. Johnson WG, Mullet, N. Georgia Court Competency Test–R, in *Dictionary of Behavioral Assessment Techniques*. Edited by Hersen M, Bellack AS. New York: Pergamon Press, 1988, p 234
  20. Poythress NG, Nicholson R, Otto RK, *et al*: The MacArthur Competence Assessment Tool–Criminal Adjudication: Professional Manual. Odessa, FL: Psychological Assessment Resources, Inc., 1999
  21. Ohio Rev. Code § 2945.37(G)
  22. *In re Williams*, 687 N.E.2d 507, 510–11 (Ohio Ct. App. 1997)
  23. Abdi H: Holm’s sequential Bonferroni procedure, in *Encyclopedia of Research Design*. Edited by Salkind NJ, Dougherty DM, Frey B. Thousand Oaks, CA: Sage, 2010, pp 573–7
  24. Lunn D, Spiegelhalter D, Thomas A, *et al*: The BUGS project: evolution, critique and future directions. *Stat Med* 28:3049–67, 2009
  25. Mossman D, Bowen MD, Vanness DJ, *et al*: Quantifying the accuracy of forensic examiners in the absence of a “gold standard.” *Law & Hum Behav* 34:402–17, 2010
  26. Mossman D, Wygant DB, Gervais RO: Estimating the accuracy of neurocognitive effort measures in the absence of a “gold standard.” *Psychol Assess* 24:815–22, 2012
  27. Mossman D: Predicting restorability of incompetent criminal defendants. *J Am Acad Psychiatry Law* 35:34–43, 2007
  28. Wildman RW II, White PA, Brandenburg CE: The Georgia Court Competency Test: the base-rate problem: Percept Motor Skills 70:1055–8, 1990
  29. Pirelli G, Gottdiener WH, Zapf PA: A meta-analytic review of competency to stand trial research. *Psychol Pub Pol’y & Law* 17:1–53, 2011
  30. Bortolato B, Miskowiak KW, Köhler CA, *et al*: Cognitive dysfunction in bipolar disorder and schizophrenia: a systematic review of meta-analyses. *Neuropsychiatr Dis Treat* 11:3111–25, 2015
  31. Vöhringer PA, Barroilhet SA, Amerio A, *et al*: Cognitive impairment in bipolar disorder and schizophrenia: a systematic review. *Front Psychiatry* 4:87, 2013
  32. Najolia GM, Dunning C, Smith L, *et al*: Use of the Montreal Cognitive Assessment in evaluations of adjudicative competency. *J Forensic Psychol Pract* 14:342–54, 2014
  33. Ross PT, Padula CB, Nitch SR, *et al*: Cognition and competency restoration: using the RBANS to predict length of stay for patients deemed incompetent to stand trial. *Clinical Neuropsychol* 29: 150–65, 2015
  34. Mosotho NL, Timile I, Joubert G: The use of computed tomography scans and the Bender Gestalt Test in the assessment of competency to stand trial and criminal responsibility in the field of mental health and law. *Int’l J Law & Psychiatry* 50:68–75, 2017
  35. Schwable E, Medalia A: Cognitive dysfunction and competency restoration: using cognitive remediation to help restore the un-restorable. *J Am Acad Psychiatry Law* 35:518–25, 2007
  36. Erard RE: If it walks like a duck: a case of confirmatory bias. *Psychol Inj & L* 9:275–7, 2016