

---

---

## Opinion Formation in Evaluating Sanity at the Time of the Offense: An Examination of 5175 Pre-Trial Evaluations

Janet I. Warren, D.S.W\*, Daniel C. Murrie,  
Ph.D., Preeti Chauhan, B.S., Park E. Dietz,  
M.D., Ph.D., M.P.H. and James Morris, Ph.D.

---

**Sanity evaluations are high-stake undertakings that explicitly examine the defendant's culpability for a crime and implicitly explore clinical information that might inform a plea agreement. Despite the gravity of such evaluations, relatively little research has investigated the process by which evaluators form their psycholegal opinions. In the current study, we explore this process by examining 5175 sanity evaluations conducted by a cohort of forensic evaluators in Virginia over a ten-year period. Our analyses focus on (i) the clinical, criminal, and demographic attributes of the defendant correlated with opinions indicative of insanity; (ii) the clinical content of the evaluations and the legal criteria referenced as the basis for the psycholegal opinion; (iii) the process and outcome differences in the sanity evaluations conducted by psychologists versus psychiatrists; and (iv) the consistency in these opinions over a ten year period. Analyses predicting an opinion of insanity indicate a positive relationship with psychotic, organic, and affective diagnoses and previous psychiatric treatment. Analyses also indicate a negative relationship with prior criminal history, drug charges, personality disorder diagnosis, and intoxication at the time of the offense. Modest racial disparities were observed with evaluators offering opinions that the defendant was insane more often for white than for minority defendants despite comparable psychiatric and criminal characteristics. Copyright © 2003 John Wiley & Sons, Ltd.**

---

\*Correspondence to: Janet I. Warren, Institute for Law, Psychiatry, and Public Policy, University of Virginia Health System, P.O. Box 800660, Charlottesville, VA 22908-0660. U.S.A.

E-mail: [jiw@virginia.edu](mailto:jiw@virginia.edu)

This research was funded under contract by the Virginia Department of Mental Health, Mental Retardation and Substance Abuse Services and the Office of the Attorney General of the Commonwealth of Virginia. The points of view expressed in this document are those of the authors and do not necessarily represent the official position of the Virginia Department of Mental Health, Mental Retardation and Substance Abuse Services or the Office of the Attorney General.

## INTRODUCTION

Although the field of forensic psychology has demonstrated substantial expansion and maturation during its brief 50 year history, practice guidelines remain ill defined for many types of forensic evaluation (Otto & Heilbrun, 2002). Indeed, a review of the research reveals such variability in clinical–forensic practice that some scholars have concluded that the level of forensic practice “falls far short of professional aspirations for the field” (Nicholson & Norwood, 2000, p. 9). Such a status warrants concern, given the trend for both federal and state courts to require an increasingly empirical standard of practice (*Daubert v. Merrell Dow Pharmaceuticals*, 1993; *Kumho Tire Co. v. Carmichael*, 1999) and growing societal concerns about the relationship between psychiatric disorders and violent crime.

How, then, is forensic practice conducted? What exactly do evaluators *do* to arrive at their psycholegal opinion? The research addressing the process or outcome of forensic mental health evaluations has tended to focus on trial competence. Traditionally, these studies have examined the factors associated with a finding of incompetence (Cooper & Grisso, 1997; Grisso, 1992; Hart & Hare, 1992; Nicholson & Kugler, 1991; Rosenfeld & Wall, 1998). More recently, studies addressing competence evaluation have begun to examine the reliability among evaluating clinicians (Rosenfeld & Ritchie, 1998) as well as the factors considered and the process documented (Nicholson & Norwood, 2000; Skeem, Golding, Cohn, & Berge, 1998). Other emergent research explores the roles that psychological testing (Borum & Grisso, 1995; Heilbrun, 1990), third party information (Heilbrun & Collins, 1995; Heilbrun, Rosenfeld, Warren, & Collins, 1994; Heilbrun, Warren, & Picarello, 2003), and different methods of quantification (Slovic, Monahan, & MacGregor, 2000) can play in forensic evaluation and opinion formation.

Less often available in forensic research has been the empirical study of sanity at the time of the offense, even though these pre-trial evaluations are central not only in determining a defendant’s culpability for often serious felony crimes, but also in discerning clinical information that might be relevant to plea bargaining and sentencing. This relative lack of empirical research may reflect the smaller number of sanity evaluations conducted each year and the low base rate for an insanity finding by evaluators or the courts (Warren, Fitch, Dietz, & Rosenfeld, 1991). Undoubtedly, however, it also reflects the complex nature of the sanity evaluation and the difficulties intrinsic to studying it empirically.<sup>1</sup> Unlike the relatively consistent battery of questions used in standard evaluations of trial competence, sanity evaluations require a retrospective examination of mental state, which demands a detailed scrutiny of intrinsically unique crimes, as well as integration of clinical and collateral information. To the extent that subjective clinical impressions guide the inquiry and determine the progression from clinical observation to psycholegal opinion formation, the process may be too covert to permit direct observation of the sanity evaluation and its conclusions.

Given the limits of the available research, we attempt to describe these dynamics by examining the process used and the conclusions reached in 5175<sup>2</sup> sanity

<sup>1</sup>For a more on conducting sanity evaluations and empirical research using a standardized measure designed to address legal sanity, see Rogers and Shuman (2000).

<sup>2</sup>Although the overall  $N = 5175$ ,  $ns$  vary by analysis. In some cases, certain variables were missing, while in other cases we created a smaller subset ( $n = 3174$ ) to control for the effects of a few prolific evaluators.

evaluations conducted in Virginia over a 10 year period. In particular, we focus on (i) the clinical, criminal, and demographic attributes of the defendant that correlate with opinions indicative of insanity; (ii) the forensic process and legal criteria used by the evaluating clinician(s) in reaching their psycholegal opinions; (iii) the process and outcome differences in the sanity evaluations conducted by psychologists versus psychiatrists; and (iv) the consistency in these opinions reached by a large cohort of evaluators over a 10 year period. These analyses are intended to further describe the reliability and validity of these sanity evaluations as well as to inform debate over the relative ability of different mental health professionals to perform these evaluations.

### **Offenses, Diagnoses, and Psycholegal Opinion**

A few studies of large samples have investigated the relationship between criminal offenses or psychiatric diagnoses and subsequent psycholegal opinion regarding sanity at the time of the offense. Warren and colleagues (1991) examined 894 pretrial referrals in Virginia, 617 of which featured sanity as a referral issue either alone or in addition to competency to stand trial (CST). In that study, evaluators reached an opinion supporting an insanity claim in only 47 (8%) of the cases. Defendants' diagnoses were significantly related to the legal opinion. Not surprisingly, schizophrenia (13 defendants, 28%) was the most frequently cited diagnostic category among the 47 defendants opined to be insane. Affective disorder (15% of those opined insane) and mental retardation (11% of those opined insane) were also relatively common diagnoses given in evaluations that supported a finding of insanity. In contrast, diagnoses of personality disorders and substance abuse disorders were least likely to be associated with a finding of insanity. The above findings are consistent with the conclusions that Melton, Petrila, Poythress, and Slobogin (1997) drew after reviewing six smaller studies conducted across four states at different points between 1967 and 1987. Namely, the authors reported, "data suggests that the presence of a major psychosis is required for the insanity defense to succeed," particularly in more recent years (p. 216). Similarly, Callahan, Steadman, McGreevy, and Clark-Robins (1991) reported that 55% of those who pled not guilty by reason of insanity and 84% of those acquitted as such carried a diagnosis of "schizophrenia or another major mental illness (other psychotic or affective disorder)" (p. 336).<sup>3</sup>

In the Warren et al. (1991) study, the category of offense was also significantly related to the opinion supporting insanity. For example, between 0 and 4% of defendants charged with robbery, sex offending, or murder were opined to be insane, as compared with 10% of defendants charged with property crimes and public order offenses.

A decade later, Cochrane, Grisso, and Frederick (2001) conducted a similar study, examining competence and sanity opinions among 1710 federal pretrial defendants, and found similar results. Their data reflected a strong relationship between diagnoses and psycholegal opinion: 40% of defendants with psychotic

<sup>3</sup>The stated goal of the Callahan et al. (1991) study was to present a systemic description of "the volume, rates, and composition of insanity pleas and acquittals across states" (p. 332). As such, it differs from the other studies reviewed, and the present one, which report the evaluator's clinical or psycholegal opinion rather than a legal strategy (entering an insanity plea) or judicial decision (acquittal on such a plea).

disorders were opined to be insane as compared to only 6% of those with personality disorders. Regarding criminal charges, insanity rates varied greatly based on the type of offense. For example, those charged with threatening a government official or assault were most likely to be found insane by forensic examiners (36 and 31%, respectively), as compared with none of the defendants charged with sex crimes or kidnapping.

Unlike previous research, Cochrane and colleagues investigated the relationship between diagnoses and charges. The authors proposed that the high rates of insanity for certain crime categories were best explained by the high rates of psychotic diagnoses for defendants within these crime categories. Consistent with their hypothesis, logistic regression revealed that there were no significant relationships between charges and psycholegal opinion once diagnoses were also considered. Rather, diagnoses were related to the types of offense the defendant committed and diagnostic presentation was the main variable to affect psycholegal opinion. This conclusion about the primacy of diagnostic presentation is particularly noteworthy when compared with a study of trial competence (e.g. Rosenfeld & Ritchie, 1998) that demonstrated a relationship between charges and psycholegal opinion and concluded that forensic evaluators may be biased towards particular conclusions due to the severity of the offense.

### **Other Factors Involved in Psycholegal Opinion Formation**

Beyond the defendant's diagnostic presentation, what other factors do examiners consider in sanity evaluations? Interestingly, the forensic literature appears to offer more guidance regarding what evaluators *should* consider rather than what they *do* consider in practice. For example, several authors have provided well reasoned guidelines about the ways in which psychological testing (see, e.g., Borum & Grisso, 1996; Heilbrun, 1992) and third party information (Heilbrun *et al.*, 1994, 2003) should inform forensic evaluations. However, normative data regarding the degree to which evaluators actually use such sources of clinical information is more sparse. Among other benefits, such normative data could be useful in evaluating some of the critiques of forensic evaluations (Heilbrun & Collins, 1995). For example, Grisso's (1986) summary of the criticisms levied against forensic practice included the claim that evaluators failed to collect sufficient or legally relevant information on which to base their opinion.

An early study by Petrella and Poythress (1983) yielded data regarding the sources of information forensic evaluators consider when performing sanity evaluations. For the 80 sanity cases examined, 33% of reports documented consultation with the defendant's attorney, 26% documented a review of previous medical or psychiatric records, and 14% documented some type of consultation with other professionals.

Heilbrun and Collins (1995) also examined forensic evaluation reports ( $n = 277$ ) including evaluations performed in the community ( $n = 110$ ) as well as those from an inpatient state hospital ( $n = 167$ ) to investigate differences between settings. The legal issues addressed in this sample of reports included competence, sanity, both, or other. For cases in which sanity was at issue (either alone or in addition to competence), the authors presented information only on procedures used by the

community sample. A clinical interview was used in 98–100% of cases, a mental status exam was performed in 67–69% of cases, an arrest report was reviewed in 65–67% of cases, and prior mental health records were reviewed in only 31–33% of the cases.

Regarding the use of psychological testing in criminal forensic evaluation, only 16% of the reports in the Heilbrun and Collins (1995) sample (which included hospital and community evaluation of competence, sanity, and other issues) mentioned testing, with the MMPI and WAIS-R cited most frequently. Testing was used more often in hospital-based reports, but this is likely a function of the evaluators' discipline; hospital-based evaluators were psychologists, whereas most community-based evaluators were psychiatrists. A full breakdown of test usage by legal issue (allowing for examination of test administration in all cases where sanity was raised as an issue) was not available. Finally, Borum and Grisso (1995) surveyed 53 psychologists and 43 psychiatrists who conducted Criminal Responsibility (sanity) evaluations. Among forensic psychologists, 68% reported that they used testing "frequently" or "always," whereas 32% reported that they used testing only "sometimes" or "rarely." Rates were slightly lower among forensic psychiatrists, of whom 42% reported that they "frequently" or "always" use psychological testing.

### Interdisciplinary Differences

Another focus in the research has examined whether forensic evaluators of different disciplines differ in the degree to which they use various sources of information. Although the Heilbrun and Collins (1995) study included variables reflecting the discipline to which an evaluator belonged, the authors did not perform analyses contrasting the procedures employed by psychiatrists versus psychologists. However, the Petrella and Poythress (1983) study was designed primarily to compare the quality of forensic evaluations between disciplines. The authors examined both the final written product and the sources of information consulted during the evaluation process. For sanity evaluations, psychologists sought outside sources of information more frequently than did psychiatrists (0.88 versus 0.58 average consultations per case), and produced a greater quantity of "clinical notes" (172 versus 136 average lines of clinical notes per case), both of which were interpreted as suggesting a more thorough evaluation. When two legal experts, a trial judge and a prosecuting attorney, rated the quality of the final reports, those by psychologists tended to be scored more favorably, with differences reaching statistical significance on two criteria. Results from this study "fail(ed) to support the conventional wisdom that psychiatrists perform forensic evaluations that are superior to those conducted by non-medically trained clinicians" (Petrella & Poythress, 1983, p. 76).

Despite two decades of evolving forensic practice and the findings from the Petrella and Poythress (1983) study, the same "conventional wisdom" (p. 76) those authors described apparently prevails today. A recent study (Redding, Floyd, & Hawk, 2001) surveyed Virginia trial judges ( $n=59$ ), prosecutors ( $n=46$ ), and defense attorneys ( $n=26$ ) regarding their preferences for various aspects of mental health testimony. When asked to rank which mental health professional they would prefer to have complete a forensic psychological evaluation for the court, the respondents favored a psychiatrist, followed by a PhD psychologist, followed by

other clinicians. Sixty-eight percent of respondents ranked a psychiatrist first in preference, while only 32% ranked a PhD psychologist first, indicating a significant preference for psychiatrists ( $\chi^2(1) = 24, p < 0.001$ ). Given these preferences, it would be worthwhile to investigate in a large sample what differences, if any, actually exist between psychologists and psychiatrists in the practice of forensic evaluation.

## **Current Study Objectives**

In the current study, we investigate these various considerations by examining the content and process of 5175 evaluations conducted in Virginia by a cohort of psychiatrists and clinical psychologists over a 10 year period. Analyses examine factors that were related to a psycholegal opinion regarding insanity, the process used by psychologists and psychiatrists in reaching their conclusions, disciplinary differences in opinion formation, and the consistency and change in these opinions over a 10 year period. We also report the cognitive and volitional prongs of the insanity defense the evaluators identify in their opinions supportive of insanity.

## **METHOD**

### **Participants**

We reviewed the characteristics of 5175 sanity at the time of offense evaluations conducted by evaluators trained by the Institute of Law, Psychiatry and Public Policy at the University of Virginia over a 10 year period. Each of the evaluators had completed a 7 or 5 day training program that addressed both the clinical and legal parameters of evaluating legal sanity and the unique demands of this particular type of evaluative setting. Participants in the training program were required to submit an opinion segment of a report for review and to pass an examination at the end of the training program.

### **Measures**

Data were compiled using the Forensic Evaluation Information Form, a 2 page instrument that includes information concerning (i) the defendant's prior criminal and psychiatric histories, compliance at the time of the offense with prescribed psychotropic medication, use of alcohol or other non-prescribed substances at the time of the offense, and current and past psychiatric condition and criminal history; (ii) the nature of the examination, including the number and discipline of the evaluators, the time spent on different components of the evaluation, the use of psychological tests, and the sources of information both requested and obtained; and (iii) the nature of the psycholegal opinion in terms of both a penultimate opinion and the prongs of the insanity defense that are identified as suggesting impairment or not. The Forensic Evaluation Information Form is submitted voluntarily by evaluators practicing privately as well as those based in hospitals, community service

boards, and other institutions. Earlier research suggests that 50–80% of court-ordered forensic evaluations conducted in Virginia each year are entered into the Forensic Evaluation Information System (FEIS; Warren, 1992).

### **Evaluators**

A total of 222 evaluators (164 psychologists and 58 psychiatrists) evaluated 5073 cases. In 102 additional cases the individual evaluator was unknown; however the discipline of the evaluator was known, allowing the cases to remain in the dataset for analyses. Two psychologists evaluated 32% of all cases evaluated by their discipline. One evaluated 856 cases and found 34 (4%) of these to be insane; another evaluated 603 cases and found 26 (4%) of these to be insane. Similarly, two psychiatrists evaluated 50% of the cases conducted by psychiatrists. One evaluated 118 cases and found 11 (9%) to be insane and the other evaluated 210 cases and found 22 (10%) to be insane. To avoid having these four evaluators unnecessarily skew the results of the comparisons between the two disciplines, only the mean number of evaluations from each discipline ( $n=27$  for psychologists,  $n=11$  for psychiatrists) were randomly selected and allowed to remain in the dataset for these four evaluators. However, this smaller subset ( $n=3471$ ) was only used for analyses examining interdisciplinary differences and not for those analyses examining opinion variables overall.

### **Age**

Subtracting the date of birth from the date of the evaluation yielded age. Only defendants who were 18 or older at the time of the evaluation were used in age related data analyses. The mean age (36) was used to divide the variable into older or younger defendants.

### **Offense**

The instant offense was coded using the seven-category classification system developed by Policy Research Associates, Inc. for use in cross state studies. These include: violent, potentially violent, other crimes against persons, sex, property, drugs, and minor offenses. For a defendant with multiple charges only the most serious offense was considered in data analyses.

### **Diagnoses**

Diagnosis was coded using either the DSM-III-R or DSM-IV classification, depending on the year that the evaluation was completed. The diagnoses were divided into ten categories: psychotic disorder; organic disorder; mental retardation/learning disorder; affective disorder; anxiety/somatoform/conversion disorder; substance abuse disorder; personality disorder; paraphilia; and other. For those

defendants with multiple diagnoses, the most severe disorder was used in data analyses. For example, if a defendant were diagnosed with both schizophrenia and a personality disorder, the former would be used in the data analyses.

### **Psychological Testing**

Psychological tests were arranged into eight categories: personality/mood—objective; personality/mood—projective; personality/mood—projective and objective; cognitive; competence; malingering; sexual and other. Neuropsychological and neurological tests were entered as present or absent but were not further categorized.

### **Substance and Medication**

Substance intake at the time of offense was categorized as alcohol, marijuana, cocaine/amphetamines, heroine/opiates, prescription medications, alcohol combined with another drug, combination of other drugs excluding alcohol, other/unknown, and none. Prescription medication was divided into seven categories: anti-psychotics, lithium, anti-depressants, anti-convulsants, anti-anxiety and combination of others, and others. If more than one category was prescribed, the most potent medication was coded. For example, if both Prozac and Haldol were prescribed to the defendant, Haldol, the anti-psychotic, would be used in data analyses.

### **Process Variables**

Combining the time spent interviewing, collecting information, and report writing yielded the total time spent on the evaluation. All time variables were collapsed into categorical variables of either above or below the mean. We dichotomized these time variables both to facilitate interpretation and because our additional analyses examining these variables on a continuous level yielded similarly insignificant results (for a defense of dichotomization, see Farrington & Loeber, 2000).

### **Insanity Opinion**

Opinion based on mental status at the time of the offense was derived using the three prongs of the insanity standard: (a) the ability to understand the nature, character, and consequences of the act; (b) the ability to distinguish right from wrong; and (c) the ability to resist the impulse of the act. This study does not include those evaluations in which the prongs—and hence the opinion—were not identified.

## **RESULTS**

Table 1 summarizes information regarding the demographic and crime characteristics of the defendants opined to be either sane or insane by the evaluating

Table 1. Demographic and crime characteristics of defendant by opinion regarding sanity at the time of the offense

Variable	Sane	Insane	$\chi^2(1)$	CC
Age			21.24***	0.11
<36	852 (91%)	88 (9%)		
>36	609 (83%)	124 (17%)		
Gender			5.94*	0.04
Male	2867 (90%)	304 (10%)		
Female	568 (87%)	83 (13%)		
Minority status			6.88**	0.04
Yes	1489 (91%)	143 (9%)		
No	2008 (89%)	257 (11%)		
Type of crime				
Violent	1702 (89%)	208 (11%)	0.00	0.00
Potentially violent	717 (89%)	85 (11%)	0.01	0.00
Other crimes against persons	381 (86%)	61 (14%)	4.25*	0.03
Sex	365 (93%)	29 (7%)	5.45*	0.03
Property	733 (88%)	104 (12%)	2.46	0.02
Drugs	155 (96%)	8 (5%)	6.19*	0.03
Minor	195 (86%)	32 (14%)	2.54	0.02
Defendant's criminal history			7.76**	0.04
Prior convictions	2381 (90%)	256 (10%)		
No prior convictions	1568 (88%)	221 (12%)		

CC = contingency coefficient, a measure of effect size.

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.0001$ .

ns vary by analysis due to some missing data.

clinicians. As summarized in Table 1, significant differences were found on seven demographic and crime variables. However, even the largest effect size (CC = 0.11) is considered small (Cohen, 1992). Because 66% of the data was missing for age, it was not entered into subsequent analyses. Violent, potentially violent, property crimes, and minor crimes did not reflect significant differences and were not subsequently entered into the logistic regressions.

Table 2 summarizes the psychiatric characteristics of the defendants opined to be sane or insane by the evaluating clinician. The presence of a psychotic disorder emerged as the most relevant psychiatric characteristic, yielding an effect size (CC = 0.21) in the small to medium range (Cohen, 1992). Organic diagnoses, mental retardation/learning disorders, paraphilias, and anxiety/somatiform/conversion disorders did not reveal significant differences and were not subsequently entered into the logistic regression.

Table 3 summarizes certain process related characteristics of the evaluations conducted by psychologists versus psychiatrists. Although statistically significant, the effect sizes for the interdisciplinary differences reported in Table 3 tended to be small. *Post hoc* analysis of the differences in opinion by the two disciplines revealed that these differences actually reflected degree of experience ( $\chi^2(1, N = 3471) = 8.90, p = 0.003$ ) by the various evaluators rather than actual disciplinary differences. Evaluators who had conducted less than the mean number ( $n = 21$ ) of evaluations were more likely to opine that the defendant met the standard for insanity. In other words, less experienced evaluators were more likely to "find" legal insanity. Since a smaller number of experienced psychiatrists were present in this sample, the trend tended to show that psychiatrists were more likely to opine that a defendant met the standard for legal insanity.

Table 2. Psychiatric characteristics of the defendant at the time of the offense

Variable	Sane	Insane	$\chi^2(1)$	CC
Diagnosis received			21.75***	0.08
Yes	2635 (88%)	364 (12%)		
No	494 (95%)	27 (5%)		
Psychiatric diagnoses				
Psychotic disorders	823 (76%)	255 (24%)	229.22***	0.21
Organic disorders	106 (80%)	26 (20%)	10.86**	0.05
Mental retardation/learning disorders	314 (89%)	37 (11%)	0.04	0.00
Affective disorders	658 (87%)	99 (13%)	4.42*	0.03
Substance abuse	907 (97%)	26 (3%)	76.88***	0.12
Anxiety/somatiform/conversion	20 (100%)	—	2.45	0.02
Dissociative disorders	9 (69%)	4 (31%)	5.32*	0.03
Personality disorder	498 (96%)	20 (4%)	29.24***	0.08
Paraphilias	30 (100%)	—	3.68	0.03
Other	88 (97%)	3 (3%)	5.49*	0.03
Prior psychiatric history			72.36***	0.12
Prior hospitalization	2372 (87%)	398 (14%)		
No prior hospitalization	1732 (94%)	117 (6%)		
Prescribed medication at the time of the offense			83.12***	0.14
Yes	1324 (83%)	270 (17%)		
No	2617 (92%)	226 (8%)		
Taking medication at the time of the offense			6.80*	0.06
Yes	635 (88%)	85 (12%)		
No	1124 (84%)	215 (16%)		
Using substance at the time of the offense			64.68***	0.12
Yes	1601 (94%)	112 (7%)		
No	2164 (86%)	367 (15%)		

CC = contingency coefficient, a measure of effect size.

\* $p < 0.05$ ; \*\* $p < 0.001$ ; \*\*\* $p < 0.0001$ .

ns vary by analysis due to some missing data.

Table 4 summarizes logistic regressions that were conducted to assess the cumulative value of the various significant variables in explaining differences in opinions regarding sanity and insanity. Three logistic regressions were conducted based upon the demographic, criminal, and psychiatric characteristics of the defendant. All significant variables and the experience variable were introduced into an overall model predicting an opinion supportive of insanity. The final outcome model was highly significant ( $\chi^2(12, N = 1680) = 241.88, p < 0.000$ ) and contained nine variables. These include non-minority status; no drug offense charges; having a prior conviction; diagnosis of psychotic, organic, or affective disorder without a personality disorder diagnosis; previous psychiatric hospitalization; and not being under the influence of substances at the time of the offense. It is worth noting that the final model in Table 4 used the data subset that did not include the most prolific evaluators.

In an attempt to replicate findings from Cochrane *et al.* (2001), we ran a logistic regression including seven offense variables and ten diagnostic variables. Results were consistent with findings by Cochrane and colleagues in that there were no significant relationships between charges and psycholegal opinion once

Table 3. Process characteristics of sanity evaluations by clinical psychologists and psychiatrists

Variable	Clinical psychologist ( <i>N</i> = 3121)	Psychiatrist ( <i>N</i> = 350)	$\chi^2(1)$	CC
Opinion			11.84***	0.06
Insane	407 (13%)	69 (20%)		
Sane	2714 (87%)	281 (80%)		
Number of Evaluators			53.67***	0.12
Only one	2260 (74%)	182 (55%)		
More than one	801 (26%)	150 (45%)		
Hours spent				
Interviewing defendant	3.04 (2.18)	3.00 (3.36)	3.03	0.03
Collecting information	1.75 (1.89)	1.78 (2.23)	0.19	0.01
Report writing	3.43 (3.62)	3.39 (3.75)	0.07	0.00
Total time spent	8.28 (6.04)	8.21 (8.29)	2.55	0.03
Sources of information obtained				
Copy of warrant	2792 (91%)	288 (85%)	12.40***	0.06
Reasons for evaluation	2682 (87%)	313 (92%)	7.59**	0.05
Statements by the defendant	1207 (40%)	121 (37%)	0.97	0.02
Defendant's criminal history	1347 (45%)	156 (48%)	0.58	0.01
Information about alleged offense	2632 (86%)	264 (78%)	12.57***	0.06
Psychiatric/medical records	1906 (63%)	242 (72%)	10.52**	0.06
Witnesses statements	1322 (44%)	105 (32%)	15.23***	0.07
Use of psychological or neuropsychological tests				
Psychological	400 (22%)	11 (6%)	27.15***	0.12
Neuropsychological	185 (10%)	3 (2%)	15.03***	0.09
Neurological	21 (1%)	—	2.25	0.03

CC = contingency coefficient, a measure of effect size.

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.0001$ .

$n = 3471$ , a sample reduced from the overall dataset to control for the influence of a few prolific evaluators.

diagnoses were also considered. The model was highly significant ( $\chi^2(17, N = 5175) = 332.23, p < 0.0001$ ) and contained eight predictive variables. Unlike the previous study, however, our analysis did indicate one offense type—i.e. drug offense ( $\beta = -0.42$  (0.20),  $p = 0.037$ , OR = 0.43)—as being significantly, although negatively, related to an insanity opinion. Diagnoses significant in rendering an opinion of insanity included psychotic diagnoses ( $\beta = 0.68$  (0.07),  $p < 0.001$ , OR = 3.86), organic diagnoses ( $\beta = 0.57$  (0.12),  $p < 0.001$ , OR = 3.12), affective diagnoses ( $\beta = 0.31$  (0.08),  $p \leq 0.000$ , OR = 1.87), mental retardation/learning disorders diagnoses ( $\beta = 0.22$  (0.10),  $p = 0.031$ , OR = 0.43), dissociative diagnoses ( $\beta = 0.82$  (0.31),  $p = 0.007$ , OR = 5.18), and not having a substance abuse ( $\beta = -0.52$  (0.11),  $p < 0.001$ , OR = 0.56) or a personality disorder diagnosis ( $\beta = -0.35$  (0.13),  $p = 0.005$ , OR = 0.49).

An additional question involved whether the proportion of defendants opined to be insane remained consistent over the 10 year period during which these evaluations were conducted. As there were no changes during this period in the laws governing the sanity evaluations or the sanity standard itself, the degree of consistency was explored as a crude measure of the temporal stability of the opinions offered to the courts by this large cohort of primarily community based evaluators. An ANOVA of the percentages of defendants opined to be insane over each year was non-significant ( $\chi^2(1, N = 3638) = 2.44, p = 0.118$ ) suggesting no meaningful

Table 4. Logistic regression summary for demographic, offense characteristics, psychiatric diagnoses and overall model for predicting an insanity opinion

Variable	$\beta$	SE	<i>p</i>	OR
Demographic variables ( <i>N</i> = 3524)				
Female	0.15	0.07	0.026	1.36
Minority	-0.13	0.06	0.032	0.78
Offense variables ( <i>N</i> = 4426)				
Drug offense	-0.47	0.20	0.016	0.39
Prior conviction	-0.14	0.05	0.004	0.75
Psychiatric variables ( <i>N</i> = 2813)				
Diagnosis received	0.31	0.14	0.032	1.86
Psychotic diagnosis	0.51	0.11	0.000	2.76
Organic diagnosis	0.46	0.18	0.009	2.52
Affective diagnosis	0.24	0.12	0.041	1.62
Substance abuse diagnosis	-0.47	0.16	0.003	0.39
Personality disorder	-0.45	0.17	0.007	0.40
Prior hospitalization	0.27	0.08	0.001	1.73
Under the influence at the time of the offense	-0.33	0.07	0.000	0.52
Overall model* ( <i>N</i> = 1680)				
Minority	-0.23	0.09	0.010	0.63
Drug offense	-1.13	0.51	0.029	0.11
Prior conviction	-0.20	0.08	0.016	0.67
Psychotic diagnosis	0.87	0.14	0.000	5.66
Organic diagnosis	0.48	0.22	0.029	2.61
Affective diagnosis	0.43	0.15	0.004	2.39
Personality disorder	-0.47	0.22	0.031	0.39
Prior hospitalization	0.32	0.10	0.001	1.90
Under the influence at the time of the offense	-0.38	0.09	0.000	0.47

\*This model includes all the previous variables and the variable representing evaluator experience.

changes in the frequency of and insanity opinion offered the courts throughout the 10 years of data collection and the large number of evaluations/evaluators.

Finally, Table 5 summarizes the various prongs of the insanity defense coded by the evaluators as supporting their psycholegal opinion of insanity. As summarized, of the 563 individuals opined to be insane by the evaluators, 91% of these involved an opinion that encompassed the two cognitive prongs of the insanity standard either singularly, in combination, or in conjunction with recognition of the defendant's inability to resist the impulse to act. In only 51(9%) instances did the evaluating clinician identify irresistible impulse as the only prong of the insanity defense relevant to the legal question.

## DISCUSSION

These findings are highly consistent in many respects with clinical wisdom concerning the evaluation of sanity at the time of the offense. The overall model that best predicted an opinion supportive of insanity included various serious Axis I diagnoses, the absence of an Axis II disorder as the primary diagnosis, prior psychiatric hospitalizations, not being under the influence of substances at the time of the offense, and not being charged with a drug offense. Clearly, and not surprisingly, a history of serious mental illness was the most influential clinical

Table 5. Prongs cited in evaluations supportive of insanity

Type of impairment found	
Impairment on all three prongs, A, B, C* (both cognitive and volitional)	223 (40%)
Impairment on both A and B (solely cognitive)	119 (21%)
Impairment on both A and C (both cognitive and volitional)	22 (4%)
Impairment on both B and C (both cognitive and volitional)	20 (4%)
Impairment on A only (solely cognitive)	58 (10%)
Impairment on B only (solely cognitive)	70 (12%)
Impairment on C only (solely volitional)	51 (9%)

\*The three prongs of the insanity standard: (A) the ability to understand the nature, character, and consequences of the act; (B) the ability to distinguish right from wrong; and (C) the ability to resist the impulse of the act.

factor, with individuals with a psychotic diagnosis being over five times more likely to be found insane than those without this type of diagnosis. As in the Cochrane et al. (2001) study, clinical characteristics of the defendant were found to override offense characteristics in predicting an opinion of insanity, except—in our sample—for a negative association with drug offenses. The professional discipline of the evaluators did not remain significant in this final model, suggesting that well trained forensic psychiatrists and psychologists, while using a slightly different methodology, are fairly consistent in terms of the proportion of defendants they opine to meet the insanity standard.

Evaluators often offered their opinions on the basis of incomplete data. As shown in Table 3, in more than half the cases, evaluators of both professional disciplines were offering sanity opinions without having seen statements by the defendant, the defendant's criminal history, and/or the statements of witnesses. In a smaller proportion of cases, evaluators had not obtained information about the alleged offense. There are many cases in which a screening evaluation can be completed without these data, and some hospital-based evaluators may feel that the inpatient setting affords enough first-hand data to make collateral data less necessary. However, the forensic evaluator who reaches an opinion about criminal responsibility without looking at these sources of information is at greater risk of reaching the wrong conclusion and would have great difficulty defending this omission on vigorous cross-examination. He or she may appear even more negligent given the practice standards implied by survey research, which indicates that most forensic professionals view such collateral information as necessary for conducting a sanity evaluation (Borum & Grisso, 1996). More broadly, when many forensic evaluators fail to consider relevant data, they leave the field vulnerable to the types of criticism that Grisso (1986) summarized nearly two decades ago: specifically, that evaluators fail to consider clinically and legally relevant information. In fairness to these examiners, it should be noted that we are aware of some jurisdictions within the state wherein prosecutors consistently decline requests to share non-statutorily-mandated offense-related information with forensic evaluators, presumably out of concern that this information will be shared with defense counsel. Although such practice does not violate the rules of discovery, it may make it more difficult for evaluators to reach an accurate psycholegal opinion in certain cases.

One finding of broad societal concern is that minority status had a statistically significant negative association with an opinion of insanity, even when all the

relevant variables were controlled within the final outcome model. Although the effect size was small ( $CC = 0.04$ , reflecting 8.5% of minorities versus 11.4% of whites opined to be insane), the discrepancy is important in light of the longstanding racial disparity within the American criminal justice system. Racial bias was not observed in the percentage of defendants referred for evaluation: 43% minority and 57% white defendants were referred for evaluation as compared to 42% minority and 59% white suspects arrested in Virginia (Virginia Department of State Police, 2001), suggesting that racial dynamics are influential in the outcome of the forensic evaluation rather than in the process of representation and referral by counsel. It was not possible to determine whether the racial disparity in our sample occurred only across evaluator/defendant racial lines, as data were not available on the race of the referring attorney or evaluator, although it is clear that the majority of evaluators were of non-minority status.

All models reflected a negative relationship between the use of substances at the time of the offense and a psycholegal opinion supportive of insanity. While this finding likely reflects the legal responsibility that derives from voluntary intoxication, it may also reflect the connotation of wrong-doing and, therefore, culpability in instances wherein substances are paired with criminal behavior. Apparently, there are cases in which evaluators of psychiatrically impaired individuals who offended while under the influence of a substance have the opinion that the influence of the substance overrode the psychotic thinking that might otherwise have been relevant to the criminal behavior, whether or not it is the pivotal factor in the motivation for the crime. The methods used here do not permit an analysis of the possible explanations for this finding, which could include a reasoned opinion that the underlying illness did not cause the defendant to meet the legal test of insanity, a moral bias against substance abusers, and/or the biasing effect of anticipating that juries are unlikely to negate culpability in instances in which substance abuse is involved.

The analyses that were conducted across the 10 years of data collection suggest no significant changes in the proportion of defendants opined to be insane from one year to the next. These numbers are noteworthy given the lengthy time period and the numerous evaluators, which make for a sample size that should be sensitive even to modest year-to-year changes in opinion rates. These results suggest that a community-based forensic mental health system is able to offer the courts a reliable cohort of forensic evaluators. The consistent base rate of insanity opinions across years in Virginia, which varied from 7 to 15% with a mean of 12%, is also consistent with national trends indicating that findings of insanity are unusual regardless of locale (Warren, Rosenfeld, Fitch, & Hawk, 1997). It is likely that the consistency observed within the Commonwealth of Virginia derives in part from the rigorous nature of the training required of forensic evaluators prior to their involvement in the community-based system. This multi-disciplinary training, which is mandated by statute and funded by the state Department of Mental Health, is within a university environment, integrates legal knowledge with forensic principles, and does so in an applied manner that is designed to build upon the preexisting clinical experiences of the trainees.

The analysis of the insanity “prongs” upon which the forensic evaluators based their findings of insanity documents the relative importance of the cognitive, as contrasted with the volitional, test of insanity. In 91% of the opinions supporting

insanity that were formulated by this cadre of evaluators, examiners identified at least one of the two cognitive prongs of the insanity standard as relevant to the opinions they offered to the courts. This pattern appears to reflect the longstanding difficulty that clinicians have in differentiating between an irresistible impulse and an impulse not resisted. Psychotic delusions are often central to a determination that the defendant was not able to differentiate right from wrong at the time of an offense and reflect a more clear-cut set of symptoms than are available to address the behavioral inhibitions implied by the volitional test. Bonnie suggested in 1983 that “whatever the precise terms of the volitional test, the question is unanswerable, or it can be answered only by ‘moral guesses’” (p. 196).

From a different perspective, however, one could argue that, in the majority (57%) of cases, the evaluating clinician identified the volitional prong as relevant to their opinion of legal insanity. It is our impression that, in those cases in which the volitional prong was combined with the cognitive prong (48% of all cases), the evaluators perceived the cognitive prong to be the determinative factor for reaching an opinion supportive of insanity, with the volitional prong reflecting more the emotional nature of the crime versus an actual compelling state of volitional dyscontrol. The current data set, however, does not allow us to examine this hypothesis or to otherwise compare the two different interpretations of this data.

The findings reported here may be more representative than previous studies because they reflect the work of over 200 forensic examiners. Attempts were made to control for the biases of a few prolific examiners, but the lack of standardization and reliability in the actual evaluations are one substantial limitation of our data. Nevertheless, these findings speak to questions of public policy, by supporting the viability and reliability of community-based forensic training, and to questions of forensic practice, reflecting the factors that forensic psychologists and psychiatrists consider when offering opinions to the court.

## REFERENCES

- Bonnie, R. (1983). The moral basis of the insanity defense. *American Bar Association Journal*, 69, 194–204.
- Borum, R., & Grisso, T. (1995). Psychological test use in criminal forensic evaluations. *Professional Psychology: Research and Practice*, 26, 465–473.
- Borum, R., & Grisso, T. (1996). Establishing standards for criminal forensic reports: An empirical analysis. *Bulletin of the American Academy of Psychiatry and the Law*, 24, 297–317.
- Callahan, L. A., Steadman, H. J., McGreevy, M. A., & Clark-Robins, P. (1991). The volume and characteristics of insanity defense pleas: An eight-state study. *Bulletin of the American Academy of Psychiatry and Law*, 19, 331–338.
- Cochrane, R. E., Grisso, T., & Frederick, R. I. (2001). The relationship between criminal charges, diagnoses, and psycholegal opinions among federal pretrial defendants. *Behavioral Sciences and the Law*, 19, 565–582.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Cooper, D. K., & Grisso, T. (1997). Five year research update (1991–1995): Evaluations for competence to stand trial. *Behavioral Sciences and the Law*, 15, 347–364.
- Daubert, V. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).
- Farrington, D. P., & Loeber, R. (2000). Some benefits of dichotomization in psychiatric and criminological research. *Criminal Behaviour and Mental Health*, 10, 100–122.
- Grisso, T. (1986). *Evaluating competencies*. New York: Plenum.
- Grisso, T. (1992). Five year research update (1986–1990): Evaluations for competence to stand trial. *Behavioral Sciences and the Law*, 10, 353–369.
- Hart, S. D., & Hare, R. D. (1992). Predicting fitness to stand trial: The relative power of demographic, criminal, and clinical variables. *Forensic Reports*, 5, 53–65.

- Heilbrun, K. (1990). Response style, situation, third-party information and competency to stand trial: Research issues in practice. *Law and Human Behavior*, *14*, 193–196.
- Heilbrun, K. (1992). The role of psychological testing in forensic assessment. *Law and Human Behavior*, *16*, 257–272.
- Heilbrun, K., & Collins, S. (1995). Evaluations of trial competency and mental state at time of offense: Report characteristics. *Professional Psychology: Research and Practice*, *26*, 61–67.
- Heilbrun, K., Rosenfeld, B., Warren, J., & Collins, S. (1994). The use of third party information in forensic assessments: A two state comparison. *Bulletin of the American Academy of Psychiatry and Law*, *22*, 399–406.
- Heilbrun, K., Warren, J., & Picarello, K. (2003). The use of third party information in forensic assessment. In I. B. Weiner, & A. M. Goldstein (Eds.), *Comprehensive handbook of psychology*, Vol. 11, *Forensic Psychology* (pp. 69–86). New York: John Wiley & Sons Ltd.
- Kumho Tire Company v. Carmichael, No. 97–1079, 1999 U.S.LEXIS 2189 (March 23, 1999).
- Melton, G., Petrilla, J., Poythress, N., & Slobogin, C. (1997). *Psychological evaluations for the courts* (2nd ed.). New York: Guilford.
- Nicholson, R. A., & Kugler, K. E. (1991). Competent and incompetent criminal defendants: A quantitative review of comparative research. *Psychological Bulletin*, *3*, 355–370.
- 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.
- Otto, R. K., & Heilbrun, K. (2002). The practice of forensic psychology: A look toward the future in light of the past. *American Psychologist*, *57*, 5–18.
- Quinnell, F. A., & Bow, J. N. (2001). Psychological tests used in child custody evaluations. *Behavioral Sciences and the Law*, *19*, 491–501.
- Petrella, R. C., & Poythress, N. G. (1983). The quality of forensic evaluations: An interdisciplinary study. *Journal of Consulting and Clinical Psychology*, *51*, 76–85.
- Redding, R. E., Floyd, M. Y., & Hawk, G. L. (2001). What judges and lawyers think about the testimony of mental health experts: A survey of the courts and the bar. *Behavioral Sciences and the Law*, *19*, 583–594.
- Rogers, R., & Shuman, D. W. (2000). *Conducting insanity evaluations*. New York: Guilford.
- Rosenfeld, B., & Ritchie, K. (1998). Competence to stand trial: Clinician reliability and the role of offense severity. *Journal of Forensic Science*, *43*, 151–157.
- Rosenfeld, B., & Wall, A. (1998). Psychopathology and competence to stand trial. *Criminal Justice and Behavior*, *25*, 443–462.
- Skeem, J. L., Golding, S. L., Cohn, N. B., & Berge, G. (1998). Logic and reliability of evaluations of competence to stand trial. *Law and Human Behavior*, *22*, 519–547.
- Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior*, *24*(4), 271–296.
- U.S. Department of Justice. (1997). *Correctional population in the United States*. Washington, DC: Bureau of Justice Statistics.
- U.S. Department of Justice. (2000). *Crime in the United States: Uniform crime reports*. Washington, DC: Federal Bureau of Investigation.
- Virginia Department of State Police. (2001). *Crime in Virginia: Virginia uniform crime reporting program*. Richmond, VA: Author.
- Warren, J. I. (1992). *Forensic evaluation reimbursement study: The Virginia Supreme Court*. Confidential report compiled for the Department of Mental Health, Mental Retardation and Substance Abuse.
- Warren, J. I., Fitch, W. L., Dietz, P. E., & Rosenfeld, B. D. (1991). Criminal offense, psychiatric diagnoses, and psycholegal opinion: An analysis of 894 pretrial referrals. *Bulletin of the American Academy of Psychiatry and the Law*, *19*, 63–69.
- Warren, J. I., Rosenfeld, B., Fitch, W. L., & Hawk, G. (1997). Forensic mental health clinical evaluation: An analysis of interstate and intersystemic differences. *Law and Human Behavior*, *21*, 377–390.