

Structured instruments for insanity defense evaluations: Opportunities and limitations

Giovanna Parmigiani¹ | Gabriele Mandarelli² | Gerben Meynen^{3,4} | Stefano Ferracuti¹

¹Department of Human Neurosciences, "Sapienza" University of Rome, Rome, Italy

²Interdisciplinary Department of Medicine, Section of Criminology and Forensic Psychiatry, University of Bari, Bari, Italy

³Willem Pompe Institute for Criminal Law and Criminology, Utrecht University, Utrecht, The Netherlands

⁴Faculty of Humanities, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Correspondence

Giovanna Parmigiani, Department of Human Neurosciences, "Sapienza" University of Rome, Viale dell'Università 30, Rome 00185, Italy.
Email: giovanna.parmigiani@uniroma1.it

Abstract

Insanity evaluations are often criticized for their—alleged—lack of objectivity, reliability and transparency. Structured tools to guide and support forensic evaluators during these evaluations have been developed—but they are rarely employed in forensic practice. In the present article, we consider the value of these tools for forensic practice in terms of opportunities and limitations. First, we briefly describe different insanity criteria used in Western countries. Next, we will review five structured instruments to guide insanity assessment together with their performance measures. Finally, we draw conclusions on the value of such instruments for forensic practice.

KEYWORDS

insanity criteria, insanity evaluation, structured instruments

1 | INTRODUCTION

The insanity defense dates back to ancient times and is an element of almost all Western legal systems. Even though, ultimately, it is up to the court or the jury to decide about a defendant's sanity; in practice, the evaluation by a forensic psychiatrist or psychologist is required. This evaluation represents one of the most challenging and debated evaluations at the intersection of psychiatry and the law—whose objectivity, reliability and transparency have frequently been questioned. These qualms may be based on several reasons, among which:

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. Behavioral Sciences & The Law published by John Wiley & Sons Ltd.

- The retrospective nature of the assessment: the forensic evaluator must infer the defendant's state of mind at the time of the crime (often weeks or months before the assessment) on the basis of the subject's examination and history, psychiatric assessment, and collateral sources of information (e.g., police and health records, witnesses);
- The risk of faking good or faking bad (malingering) (Arin & Mengchuay, 2022; Knoll & Resnick, 2008). Given the considerable reliance of psychiatric diagnosis on the person's own words (history taking), the outcome of the examination may be influenced by a defendant's (possibly) untruthful answers to the psychiatrist's questions.
- Whether or not legal experts should give explicit advice about the defendant's sanity/criminal responsibility (Meynen, 2016). Legal insanity is a legal matter, and therefore, in some jurisdictions, experts are not allowed to make explicit statements about the defendant's sanity; they should restrict themselves to describing the presence, nature and (behavioral) impact of a mental condition. However, in others, such as The Netherlands and Italy, psychiatrists and psychologists are allowed—and even asked—to make statements about the defendant's sanity.
- Research shows that forensic experts often disagree about the final judgement on criminal responsibility regarding the same case (Gowensmith et al., 2013; Guarnera & Murrie, 2017);
- Standardized procedures on how to perform an insanity defense evaluation, as well as biological markers, are lacking (Beckham et al., 1989; Meyer & Valença, 2021).
- The forensic evaluators' decisional processes may be influenced by the presence of unintentional or cognitive biases, money, prestige, and the amount of public attention attracted by the case (Beckham et al., 1989; Commons et al., 2004; Dattilio et al., 2006; Homant & Kennedy, 1986; Murrie et al., 2013).

These qualms deserve to be addressed because of the significant forensic and procedural consequences of these evaluations and the far-reaching ramifications for defendants, victims, and the community. When a defendant, whose severe psychiatric disorder had a decisive influence on their criminal behavior, is erroneously considered to be responsible for a crime, they will be punished for a crime for which they should not be held accountable. In addition, they could have fewer possibilities to be treated for their disease (this, however, depends on the legal system). On the other hand, when a sane defendant is mistakenly considered insane at the moment of the crime, they will not be punished for a crime they should have been held responsible for. Finally, by entering a forensic psychiatric system, they will use treatment resources that are usually limited. In both cases, no justice is served.

Several authors have developed theoretical models, guidelines and instruments to support and structure insanity evaluations in order to render them more objective, reliable and transparent (American Academy of Psychiatry and the Law, 2014; Cai et al., 2014; Kalis & Meynen, 2014; Meyer et al., 2020; Meynen, 2010a, 2010b, 2011, 2012, 2013, 2016, Parmigiani, Mandarelli, Meynen, Carabellese, & Ferracuti, 2019; Parmigiani et al., 2022; Rogers et al., 1981; Rogers et al., 1984; Slobogin et al., 1984). The purpose of the present article is to analyze the opportunities and limitations of five structured assessment tools for insanity defense evaluations. Firstly, we will briefly discuss some differences between legal insanity criteria used in Western countries. Next, we will review five assessment tools designed to guide the insanity defense evaluation and their performance measures. Finally, we will consider the advantages and limitations of these tools and draw conclusions.

2 | INSANITY DEFENSE CRITERIA

Insanity criteria tend to vary considerably across jurisdictions. Firstly, the burden of proof may differ with some countries requiring insanity to be proven, while others requiring sanity to be proven (Simon & Ahn-Redding, 2006). Secondly, the threshold for proof may vary with some legal systems demanding a "preponderance of the evidence," while others "clear and convincing evidence". Thirdly, the degrees of responsibility may be different, ranging from two (the dichotomy of sanity or insanity) to three or even more (Simon & Ahn-Redding, 2006). Fourthly, legal systems may be inquisitorial or adversarial, which may have consequences for the behavioral experts: they may be experts for the defense or prosecution (in adversarial systems) and for the court (in inquisitorial, continental systems). Finally, and

importantly, the legal criteria for the defense itself may vary considerably as we will discuss below. Specifically, we will just focus on some of those standards, which provide contrasting examples on the variety of insanity criteria used, with an emphasis on Anglo-American standards (for a more comprehensive description of standards across various jurisdictions, see Simon & Ahn-Redding, 2006).

In the United States, there are several insanity standards, among which the M'Naghten Rule (1843) and the Model Penal Code test (American Law Institute, 1962). The first standard relies specifically on the assessment of the defendant's knowledge, also known as the "cognitive component", as it states that a defendant is not to be found responsible if, due to a mental disorder, he did "not know the nature and quality of the act he was doing; or if he did know it, that he did not know what he was doing was wrong". However, the Model Penal Code's test (American Law Institute, 1962), used in many other states in the United States, adds a control element to the criteria as it states "a person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity either to appreciate the criminality (wrongfulness) of his conduct or to conform his conduct to the requirements of the law". So, even if a defendant did appreciate the wrongfulness of his behavior, he can still be considered insane, if he was unable to control his behavior. These two components—knowledge/appreciation on the one hand and control on the other—are reflected by the criteria of many Western legal systems. Besides these two insanity criteria, there are others, less commonly used, such as "The Product Test or Durham Rule",¹ which states: "No man shall be held accountable, criminally, for an act which was the offspring and product of mental disease", and "The Irresistible Impulse Test",² which focuses on "whether the mental disease or defect has prevented the person from controlling his behavior at the time of the offense".

In Italy, Articles 88 and 89 of the Italian Penal Code establish the criteria for determining whether an offender can be considered non-responsible or with substantially diminished responsibility. Specifically, Article 88 states that "a person who, at the moment in which he/she committed the crime, was, because of infirmity, in such a state of mind as to exclude the capacity of appreciating or willing, is non-responsible" (translated by Parmigiani et al., 2022),³ while Article 89 states that "a person who, at the moment in which he/she committed the crime, was, because of infirmity, in such a state of mind as to greatly diminish, without excluding, the capacity of appreciating and willing, is responsible for the crime committed, but the penalty is reduced" (translated by Parmigiani et al., 2022).⁴

Another example is Brazil, where Article 26 of the Brazilian Penal Code reads that:

It is exempt from punishment the agent who, on account of mental illness or incomplete or delayed mental development, was at the time of the action or omission completely incapable of understanding the illicit nature of the fact or of self-determination according to this understanding.

Single paragraph. The sentence may be reduced by one-to two thirds, if the agent, in virtue of mental disorder or incomplete or delayed mental development, was not completely capable of understanding the illicit nature of the fact or of self-determine according to this understanding.

(translated by Meyer et al., 2020).⁵

Overall, to our knowledge, most of the jurisdictions require the disorder to have an impact on the defendant's knowledge and/or behavioral control concerning the relevant legal act.

Norway, however, is an exception. Its new rule, which came into force in 2020, reads:

A person that at the time of committing the act is under 15 years old is not accountable. The same applies to a person who at the time of the act is not accountable due to a:

- (a) severe divergent state of mind,
- (b) severe impairment of consciousness, or
- (c) a high degree of intellectual disability.

When deciding whether a person is not accountable according to the first section, a–c, emphasis should be placed on the degree of failure of understanding reality and functional ability.

(translated by Gröning, 2021).⁶

Notably, this criterion of insanity does not require a relationship between the disorder and the criminal act. It is merely about the *presence* of a (severe) mental condition at the time of the act—not about a specific impact either on knowledge or on behavioral control as related to the crime (Gröning et al., 2020). This is different from both M'Naghten and the Model Penal Code test and the criteria used in Italy and Brazil.

Finally, in The Netherlands, there is no explicit legal standard for legal insanity, as Section 39 of the Dutch Criminal Code states that “A person who commits an offense for which he cannot be held responsible by reason of the mental disorder, psychogeriatric condition or intellectual disability, is not criminally liable.” (translated by Meynen, 2022).⁷

Notably, it is not specified under which condition the defendant cannot be held accountable. This is called an “open” criterion, and it was meant to be like this to provide the behavioral expert and the judge ample room to decide about legal insanity on an individual basis (Bijlsma, 2016; Gröning, 2021). Interestingly, in The Netherlands, the legal expert is asked to give explicit advice about the defendant's (degree of) criminal responsibility in the absence of a criterion.

Apart from the different insanity criteria present across jurisdiction, another concept that has been considered as central to legal insanity is the notion of irrationality or lack of rationality (Felthous, 2022; Meynen, 2016; Morse, 1998; Simon & Ahn-Redding, 2006). According to Morse (1998), “nonculpable irrationality is the underlying basis for the insanity defense” and “Irrationality is the genuine excusing condition that is operative.” In this case, is the subject's incapacity for rationality, a condition more frequent in defendants affected by psychotic disorder that is relevant for the insanity defense (Morse, 1998). In forensic and legal practice, cases in which a defendant is considered insane often concern psychotic disorders (Meynen, 2016). Clearly, the fact that there are different legal criteria for insanity is relevant for (the use and value of) the structured instruments as we will address in the Discussion section.

3 | STRUCTURED INSTRUMENTS TO ASSESS LEGAL INSANITY

In recent years, several tools (Cai et al., 2014; Meyer et al., 2020; Parmigiani et al., 2019, 2022; Rogers et al., 1981, 1984; Slobogin et al., 1984) have been developed to assist forensic evaluators in the insanity assessment, and theoretical models (Meynen, 2010a, 2010b, 2012, 2013, 2016; Kalis & Meynen, 2014; Parmigiani et al., 2017) and some practice guidelines (American Academy of Psychiatry and the Law, 2014) have been published. Despite these efforts, well-validated instruments for insanity evaluations remain sparse. At present, we are aware of five specialized instruments that are available:

- The Mental State at the Time of the Offense Screening Evaluation (MSE) (Slobogin et al., 1984)
- The Rogers Criminal Responsibility Assessment Scales (R-CRAS) (Rogers et al., 1981, 1984)
- The Rating scale of criminal responsibility for mentally disordered offenders (RSCRs) (Cai et al., 2014)
- The Criminal Responsibility Scale (CRS) (Meyer et al., 2020)
- The Defendant's Insanity Assessment Support Scale (DIASS) (Parmigiani et al., 2022)

These tools vary greatly in their design, validation and forensic applications, and consequently, will be described separately in this section.

The MSE is a semi-structured interview developed to be used as a screening device to guide forensic experts during the evaluation of defendants' psychopathological symptoms and psychological functioning at the time of the crime. It is composed of three parts: (a) defendant's psychological history (to identify the presence of previous and

enduring mental disease); (b) offense information (to detect the presence and the influence on the criminal act of the mental disorder) and (c) defendant's mental status (whose evaluation can inform the forensic expert about the validity of the defendant's report about his or her past mental state). The validation procedure involved 24 mental health professionals, grouped into 12 teams of two persons each, one of whom was either a psychiatrist or Ph.D. clinical psychologist. They were trained on the use of the MSE and then asked to express an opinion about the presence of significant dysfunction due to mental illness at the time of the offense of three defendants, randomly selected from the forensic unit at the Central State Hospital in Petersburg, Virginia. Their interviews were then evaluated by one of the two inpatient forensic evaluation teams at the Central State Hospital (each composed by one psychiatrist, one Ph.D. psychologist, and one social worker). Their assessments consisted of expressing an opinion about the presence of signs or symptoms to support an insanity defense and their judgments compared with those of the trainees. The instrument was developed as a screening tool to screen out those defendants whose criminal conduct was clearly not linked to a mental disorder (Slobogin et al., 1984).

The R-CRAS is a 25-item psychometric scale, whose variables are grouped in the following dimensions: (a) reliability (including malingering), (b) organic factors, (c) psychopathology, (d) cognitive control and (e) behavioral control. Each item score ranges from 0 (no information) to 6 (extreme severity). To complete the R-CRAS, the examiner must administer two parts: Part I, which deals with the quantification of individual variables, and Part II, where a specific decision model leads to the final judgment about sanity or insanity. The R-CRAS validation underwent a three-phase evaluation. It was initially validated through 10 vignettes based on real cases, which were independently evaluated by three forensic psychiatrists and four forensic psychologists; in a second phase, it was employed for the evaluation of 25 actual forensic cases by two forensic psychiatrists and 1 forensic psychologist (Rogers et al., 1981). In the third phase, 48 forensic insanity cases from the Court Diagnostic and Treatment Center in Toledo and the Nebraska Psychiatric Institute were added to the previous 25 actual forensic cases and evaluated by the forensic psychiatrists and psychologists of the related Centers (Rogers et al., 1984).

The RSCR is an 18-item scale grouped into 2 dimensions: capacity of cognition (which investigates, e.g., the presence of "realistic motivation", "inducement of crime", "understanding the nature of the behavior") and capacity of control (which assesses, among others aspects, "time selectivity of crime", "place selectivity of crime", "shirking responsibility after offense" and "concealing the truth during the inquest") (Cai et al., 2014). The score of each item ranges from 0 to 4. It was developed on the base of the R-CRAS. For its validation, the authors enrolled 1187 defendants recruited from the Institute of Forensic Science, Ministry of Justice, the Second Xiangya Hospital of Central South University, the Mental Health Center of Shanghai, the Psychosis Hospital of Guangzhou and Nanjing Brain Hospital. These defendants were then evaluated independently by raters (their qualifications were not reported) and forensic psychiatrists (Cai et al., 2014).

The CRS is a tool structured in vignettes that portray a brief history of an offense (Meyer et al., 2020). It is composed of 12 items with 19 questions that investigate the defendant's "capacity for understanding" (CU) and "capacity for self-determination" (CD). Each question can be scored: "0" (i.e., no acknowledgment of the cited psychopathological elements, or answered with pathological elements), "1" (i.e., partial acknowledgment of at least one of the cited psychopathological elements) or "2" (i.e., assertive acknowledgment of at least one of the cited psychopathological elements). For the validation of the instrument, the criminal responsibility of 88 defendants was evaluated through a hypothetical scenario, a clinical vignette portraying the brief history of an offense:

Two people, John and Paul, are playing cards in a room. They are sitting facing each other. Suddenly they begin to argue, and John pushes Paul. Paul falls off the chair, hits his head on the floor, and dies.

On the basis of this hypothetical scenario, the defendants' CU and CD were evaluated through questions about, for example, their "notion of legal goods and illegality" ("Do you think that there is something strange in the story I told you?"—CU), and "perception of social and legal disapproval" ("What do you think John felt after he pushed Paul and realized the consequence of the accident?"—CD). Consequently, the insanity assessment was not performed in

relation to the alleged crime, but only on the effective decisional capacity of the subject at the time of the forensic evaluation related to a hypothetical crime (Meyer et al., 2020).

The DIASS is a semi-structured interview composed of 8 binary items (present/absent) grouped into 4 dimensions: "Knowledge/understanding of the crime" (2 items), "Appreciating of the crime" (1 item), "Reasoning" (3 items) and "Control of voluntary motor activity" (2 items). One item of the first dimension and the second and third dimensions refer to the "Epistemic component", while 1 item of the first dimension and the all the items of the fourth dimension refer to the "Control component". A box at the end of the scale refers to the final judgments on the Epistemic and Control components, which are scored on a 3-point scale (Intact, partially compromised, and compromised). The validation procedure involved 40 forensic experts for the evaluation of 10 hypothetical forensic cases based on real cases modified to make them unrecognizable (Parmigiani et al., 2022). The items of the DIASS were designed to assess the defendants' knowledge/understanding and appreciation of the criminal act, the ability to reason about it, and their inhibitory control at the time of the crime. However, despite the DIASS having been validated on hypothetical forensic cases, it has been designed to be used in forensic practice to constitute a guide and support for forensic evaluators during the criminal responsibility assessment. Currently, the instrument is undergoing further validation with real forensic cases.

The characteristics, validation procedures and performance measures of the instruments are shown in Table 1 and Table 2.

Regarding the evaluation of criminal responsibility, which may be guided by the abovementioned instruments, an issue that forensic evaluators should bear in mind is the possibility of malingering or the intentional fabrication or aggravation of mental symptoms with the intention to achieve secondary benefits. For the purpose of assessing

TABLE 1 Instruments to assess criminal responsibility.

Instrument	Country	Assessed domains	Cognitive/ Volitional Prong	Scenario	Standard	Format
R-CRAS Rogers et al. (1981)	USA	Patient's reliability Organicity Psychopathology Cognitive control Behavioral control	C, V	Real (alleged crime)	ALI standard	25 items scale
MSE Slobogin et al. (1984)	USA	Historical information Offense information Present mental state examination		Real (alleged crime)		Semi-structured Interview (screening)
DIASS Parmigiani et al. (2019)	Italy	Knowledge/understanding Appreciating Reasoning Inhibitory control	C, V	Real (alleged crime)	ALI standard	8-items scale
RSCRs Cai et al. (2014)	China	Capacity of cognition Capacity of control	C, V	Real (alleged crime)	ALI standard	18-items scale
CRS Meyer et al. (2020)	Brazil	Understanding Self-determination	C, V	Hypothetical (vignettes)	ALI standard	Clinical vignettes

Abbreviations: CRS, Criminal Responsibility Scale; DIASS, Defendant's Insanity Assessment Support Scale; MSE, Mental State at the Time of the Offense Screening Evaluation; R-CRAS, Rogers Criminal Responsibility Assessment Scales; RSCRs, Rating scale of criminal responsibility for mentally disordered offenders.

TABLE 2 Psychometric properties of the instruments to assess criminal responsibility.

Instrument	Evaluators	Sample	Validation	Interrater-reliability on final judgement	Internal consistency	Concurrent validity	Results
R-CRAS	I phase 3 forensic psychiatrists and 4 forensic psychologists II phase 2 forensic psychiatrists and 1 forensic psychologist III phase Forensic psychiatrists and psychologists	I phase 10 vignettes on the R-CRAS II phase 25 forensic cases from the Isaac Ray Center III phase 48 forensic cases from the Court Diagnostic and Treatment Center in Toledo and the Nebraska Psychiatric Institute	I and II phases Evaluation by forensic evaluators at the Isaac Ray Center III phase Evaluations by forensic psychiatrists and psychologists	Cohen's kappa = 0.77	----	----	The R-CRAS differentiates between sane and insane groups; insane groups exhibited more severe psychopathology
MSE	24 mental health professionals (4 psychiatrists, 13 clinical psychologists, 7 psychiatric social workers, and 1 psychiatric nurse)	36 defendants (18 involved in minor offenses, 18 involved in major offenses)	Evaluation by inpatient forensic evaluation teams at Central state Hospital (one psychiatrist, one PhD psychologist, and one social worker)	----	----	----	There was no significant difference in the rate of agreement between the trainees and Central state staff in evaluation of defendants charged with major offenses versus evaluation of defendants charged with minor offenses.

TABLE 2 (Continued)

Instrument	Evaluators	Sample	Validation	Interrater-reliability on final judgement	Internal consistency	Concurrent validity	Results
DIASS	30 forensic psychiatrists 5 forensic psychologists 3 medico-legal experts 2 neuropsychiatrists	10 hypothetical forensic cases based on real cases	Evaluation of the 10 hypothetical vignettes by 40 forensic experts	Cohen's kappa = 0.72	Cronbach's alpha = 0.86	rho = 0.674; p < 0.001	The capacities analyzed through the DIASS were mainly affected by mania/excitement and psychotic dimensions in nonresponsible and with substantially diminished responsibility defendants, while by hostility and negative symptoms in responsible defendants.
RSCRS	6 forensic psychiatrists	10 forensic cases	Independent evaluation by raters and forensic psychiatrists	Cohen's kappa = 0.79	Cronbach's alpha = 0.96	-----	By establishing discrimination analysis among three levels (whole responsibility, diminished responsibility and irresponsibility), classification results suggested that 88.90% of the original grouped cases were correctly classified, and the discriminant value had high conformity with the experts' opinions
CRS	3 researchers with experience in forensic psychiatric assessments	88 defendants in a forensic medical institute	Independent evaluation by a forensic expert was compared with the evaluation of the researchers	----	Cronbach's alpha = 0.72	----	The CRS contributed to the forensic expert's distinction between criminally responsible versus not responsible individuals. Special attention should be paid to the group with partial CR, which did not show statistical reaffirmation in the present study

symptoms validity (or the possibility of malingering), several tools have been developed, which may be useful in forensic evaluations. Examples are: The Structured Interview of Reported Symptoms (SIRS) (Rogers, 1992), the Miller Forensic Assessment of Symptom Test (M-FAST) (Miller, 2001), the Structured Inventory of Malingered Symptomatology (SIMS) (Smith & Burger, 1997), and multiscale self-report inventories, such as the Minnesota Multiphasic Personality Inventory, Second Edition (MMPI-2) (Butcher et al., 1989) and the Personality Assessment Inventory (PAI) (Morey, 1991). Finally, even though this paper considers instruments specifically developed for insanity evaluations, we would like to mention that other psychological instruments may also be valuable in the assessment of defendants. We briefly refer to just some of them. Among the multiscale inventories are the widely used MMPI-2 (already mentioned in the section on malingering) and the Millon Clinical Multiaxial Inventory (MCMI; Millon, 1983, 1987, 1994). It is a 175-item multiscale inventory designed to be a theory-driven diagnostic measure coordinated with DSM (Millon, 1983). The use of both these instruments alone, however, is not enough to evaluate insanity (Rogers & Shuman, 2000). Among the Intellectual and Neuropsychological Tests, the Wechsler Adult Intelligence Scale—Third Edition (WAIS-II Wechsler, 1997) is a widely acknowledged standard for intellectual assessment and it is composed of two domains: verbal and performance abilities. It is a well validated instrument to detect intellectual disability, which can provide very valuable information for the assessment of criminal responsibility.

4 | DISCUSSION

In the previous section, we reviewed the structured assessment tools for insanity evaluations. Interrater reliabilities of the R-CRAS, the RSCRs and the DIASS were good, ranging from 0.72 to 0.79 (Cai et al., 2014; Parmigiani et al., 2022; Rogers et al., 1984), as well as internal consistency of the RSCRs, CRS and the DIASS ranging from 0.72 to 0.96 (Cai et al., 2014; Meyer et al., 2020; Parmigiani et al., 2022). Only one article provided concurrent validity (DIASS; $\rho = 0.674$; $p < 0.001$) (Parmigiani et al., 2022). However, we must be careful when trying to quantitatively compare and derive general conclusions from the validity and reliability of such instruments as the validation process involved different procedures and scenarios.

For example, the MSE is an interesting assessment tool (see Slobogin et al., 1984—Appendix for how to perform the evaluation), whose aim however is just to screen out those defendants whose criminal conduct was not clearly linked to a mental disorder. Despite being accurate and useful, also for those professionals with limited experience in the forensic field, it is not informative on how to evaluate the (level of) criminal responsibility, but just to detect whether there are grounds to support an insanity defense. In addition, reservations have been expressed about its development as it relies on unstandardized questions and narrative information, and it has been defined by some authors a “psychometrically unproven measure” (Rogers & Shuman, 2000). Borum and Grisso found that MSE is actually employed in insanity assessments by 0% of forensic psychologists and 2% of forensic psychiatrists (Borum & Grisso, 1995). The R-CRAS, the most acknowledged instrument to guide insanity evaluations, underwent a three-phase validation process and shows a good interrater reliability. According to its authors, it was developed as a method to “quantify essential psychological and situational variables at the time of the crime and to implement criterion-based decision models for criminal responsibility” (Rogers, 1984, p. 1). It represents a useful protocol with a detailed manual and an aid for forensic evaluators during the insanity assessment (Rogers et al., 1984). The CRS investigates salient aspects of subjects' capacity to understand (i.e., “the notion of culpability, liability, and responsibility” and “awareness of the act's illegality (prohibitive character) and criminal nature”) and of self-determination (i.e., “presence of intent and animus”, and “deliberation, decision, and execution”) and provides a manual (Meyer et al., 2020—Appendix A). However, by employing a hypothetical scenario, a clinical vignette portraying the brief history of an offense, the evaluation is not performed in relation to the alleged crime, but only on the effective decisional capacity of the subject at the *time of the forensic evaluation*. This, in our opinion, may constitute a limitation of the above-mentioned instrument, as the insanity evaluation entails a retrospective evaluation of the state of mind at the *time of the crime*, and can consequently undermine its use in the insanity assessment. The RSCRs evaluates core features of the cognitive dimension (i.e., “Realistic motivation” and “Impairment of reality testing”) and of the control

dimension (i.e., “Concealing the truth during the inquest” and “Place selectivity of crime”) of the state of mind at the time of the crime. A discrimination analysis disclosed that the instrument is capable of correctly classifying 88.90% of defendants despite a few misclassifications in the diminished responsibility group (Cai et al., 2014). Despite seeming a useful instrument to structure insanity evaluations, the authors are cautious in promoting its use in forensic practice, mentioning that it was employed in five agencies for less than 2 years, and that its reliability and validity warrant further testing and confirmation (Cai et al., 2014).

Finally, the DIASS is an 8-items scale, designed to assess both the cognitive and behavioral components (examples on how to perform the evaluation are provided in the manual, see Appendix A in Supplementary materials of Parmigiani et al., 2022). It shows good internal consistency, concurrent validity and substantial inter-rater reliability. While having been initially validated through hypothetical forensic cases, it is currently in the process of having a validation of the developed scale from actual forensic cases. In this second phase of validation, the authors have slightly modified the scale, which was initially composed of 8 binary items (present/absent). The new version of the DIASS is rated on 4 positions (present, partially present, absent, not evaluable).

With these caveats in mind, we deem that the merits of these instruments, all of which do not request a specific training but provide a structure to perform the evaluation, are, first, to make sure that relevant aspects are taken into account, and second, to provide a standardization that may reduce some of the current heterogeneity of these assessments. In fact, they can contribute to increase the objectivity, reliability and transparency of criminal responsibility assessments in order to achieve a greater standardization of forensic methodology, thus leading to an improvement in the entire criminal justice process. Considerations for practitioners when choosing an instrument can be, first, whether the instrument covers the insanity test in the relevant jurisdiction. Second, the level of validation of the instrument can be taken into account. In this respect, it should be noted that the R-CRAS has undergone an extensive process of validation. At this point, in our view at least, the RSCRs and the DIASS require further studies and confirmation. The doubts expressed about the MSE and the CRS in this paper may also be taken into account. For practitioners who collaborate in a group (e.g., in the context of intervision), it may be advisable to use the same instrument as it may facilitate the exchange of ideas. Still, a valid question that remains is: do these instruments actually measure legal insanity? There is no gold standard—like a laboratory test in other areas of medicine—to determine whether the outcome of the instrument is ‘correct’. As already stated, legal insanity is, ultimately, a *legal* category, and one might argue that a psychiatric assessment tool does not provide the final answer about what is, and what is not, legal insanity. Therefore, as mentioned in the introduction, in many legal systems, the expert is *not allowed* to make statements about the defendant’s sanity. This is an important limitation to take into account.

Furthermore, as we have seen, the criteria for legal insanity differ between jurisdictions. So, while the instrument may take into account behavioral control, the legal standard may not include a control component; for instance, the M’Naghten Rule—used in many jurisdictions—does not include such a control component. This means that, legally, the behavioral expert’s findings concerning a lack of control are not relevant. Furthermore, in Norway, the legal criterion does not refer to any influence of the disorder on the crime—while the instrument may include such a reference, like “appreciation of the crime”. Finally, in The Netherlands, no legal criterion for insanity has been formulated, which leaves the question of what is and what is not relevant “open” to the extent that in an individual case, it is up to the behavioral expert and the court to determine what makes a defendant legally (in)sane.

Still, in spite of these limitations, using a tool may help to structure the expert’s thinking about legal insanity in an individual case and increase the chances that relevant aspects are not missed, and it may also facilitate communicating findings (and doubts) to other—behavioral or legal—professionals. In general, we do not consider the tools as restricting the expert’s evaluation, but rather as ways to enrich it, to broaden and structure one’s scope. Eventually, in our view at least, the expert is responsible for what they consider relevant for the individual case at hand and for the eventual conclusions that are drawn and the advice that is given.

One could propose to develop a specific assessment tool for each criterion/jurisdiction. On one hand, we understand such an approach; on the other hand, we feel that one of the advantages of these instruments is that they

“transcend” individual jurisdictions and can be applied in different legal systems and countries: this makes interesting comparisons possible, especially in scientific research. We would like to emphasize the value of international research on insanity, and using similar tools across jurisdictions may be valuable in this respect.

Clearly, some of the challenges/problems that have led to debates about the quality of forensic assessments of insanity (see Introduction)—in particular, the fact that the evaluation is in retrospect and the risk of faking good or bad in forensic settings—are not solved by structured assessment tools. However, also here, structuring one’s findings, thoughts, (preliminary) conclusions, and potential doubts may be helpful to determine the next steps in the evaluation, and eventually, to enhance the quality of one’s report.

5 | CONCLUSION

The insanity defense has—or may have—far-reaching consequences for defendants, victims and the community. Meanwhile, doubts exist regarding the objectivity, reliability and transparency of insanity evaluation. Structured instruments have been developed to guide and support insanity evaluations—but they are rarely used in forensic practice. Still, their employment may represent a step towards standardization and transparency of these assessments. In addition, such tools may promote the exchange of ideas and research findings across jurisdictions and disciplines. However, even though such an exchange is facilitated, legal insanity is ultimately a legal matter—and there is no behavioral gold standard to determine whether the outcome is (legally) correct. This constitutes an inevitable limitation of these assessment tools. Still, the development—and use, also in research contexts—of the tools can be valuable for an area that is of considerable medical, legal, and societal importance but that continues to be relatively understudied.

ACKNOWLEDGEMENT

Open Access Funding provided by Universita degli Studi di Roma La Sapienza within the CRUI-CARE Agreement.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

ENDNOTES

- ¹ At the moment, only New Hampshire and the Virgin Islands use this test.
- ² A few states combine it with a cognitive M’Naughten arm as part of their insanity test.
- ³ Articolo 88: “non è imputabile chi, nel momento in cui ha commesso il fatto, era, per infermità, in tale stato di mente da escludere la capacità di intendere e di volere” (Codice Penale, 1930).
- ⁴ Articolo 89: “chi, nel momento in cui ha commesso il fatto, era, per infermità, in tale stato di mente da scemare grandemente, senza escluderla, la capacità di intendere e di volere, risponde del reato commesso, ma la pena è diminuita” (Codice Penale, 1930).
- ⁵ “É isento de pena o agente que, por doença mental ou desenvolvimento mental incompleto ou retardado, era, ao tempo da ação ou da omissão, inteiramente incapaz de entender o caráter ilícito do fato ou de determinar-se de acordo com esse entendimento.
Parágrafo único – A pena pode ser reduzida de um a dois terços, se o agente, em virtude de perturbação de saúde mental ou por desenvolvimento mental incompleto ou retardado não era inteiramente capaz de entender o caráter ilícito do fato ou de determinar-se de acordo com esse entendimento” (Código Penal, 1940).
- ⁶ “Den som på handlingstidspunktet er under 15 år, er ikke strafferettslig ansvarlig. Det samme gjelder den som på handlingstidspunktet er utilregnelig på grunn av:
(a)sterkt avvikende sinnstilstand.
(b)sterk bevissthetsforstyrrelse eller.
(c)høygradig psykisk utviklingshemming.

Ved utilregnelighetsvurderingen etter annet ledd skal det legges vekt på graden av svikt i virkelighetsforståelse og funksjonsevne." (Lov om straff, straffeloven—Lovdata, 2020).

⁷ Artikel 39: "Niet strafbaar is hij die een feit begaat, dat hem wegens de psychische stoornis, psychogeriatrische aandoening of verstandelijke handicap niet kan worden toegerekend" (Wetboek van Strafrecht, 1881).

REFERENCES

- American Law Institute. (1962). *Model penal Code*. American Law Institute.
- American Academy of Psychiatry and the Law (AAPL). (2014). AAPL practice guideline for forensic psychiatric evaluation of defendants raising the insanity defense. *The Journal of the American Academy of Psychiatry and the Law*, 42, S3–S76.
- Arin, N., & Mengchuay, J. (2022). Assessing symptom exaggeration of psychopathology in incarcerated individuals and mentally ill offenders within forensic contexts. *Behavioral Sciences and the Law*, 1–17. (Online first). <https://doi.org/10.1002/bsl.2603>
- Beckham, J. C., Annis, L. V., & Gustafson, D. J. (1989). Decision making and examiner bias in forensic expert recommendations for not guilty by reason of insanity. *Law and Human Behavior*, 13(1), 79–87. <https://doi.org/10.1007/BF01056164>
- Bijlsma, J. (2016). *Stoornis en strafuitsluiting. Op zoek naar een toetsingskader voor ontoerekenbaarheid*. Oisterwijk. Wolf Legal Publishers.297
- Borum, R., & Grisso, T. (1995). Psychological test use in criminale forensic evaluations. *Professional Psychology: Research and Practice*, 26(5), 465–473. <https://doi.org/10.1037/0735-7028.26.5.465>
- Brasil. Decreto-lei No. 2.484/40 (Código Penal). (1940). http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Accessed March 3 29, 2023.
- Butcher, J. D., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemer, B. (1989). *Minnesota multiphasic personality inventory-2 (MMPI-2): Manual for administration and scoring*. University of Minnesota Press.
- Cai, W., Zhang, Q., Huang, F., Guan, W., Tang, T., & Liu, C. (2014). The reliability and validity of the rating scale of criminal responsibility for mentally disordered offenders. *Forensic Science International*, 236, 146–150. <https://doi.org/10.1016/j.forsciint.2013.12.018>
- Codice, P. Regio Decreto 19 ottobre 1930, n. 1398. https://www.gazzettaufficiale.it/dettaglio/codici/codicePenale/85_1_1. accessed March 3 29, 2023.
- Commons, M. L., Miller, P. M., & Gutheil, T. G. (2004). Expert witness perceptions of bias in experts. *The Journal of the American Academy of Psychiatry and the Law*, 32, 70–75.
- Dattilio, F. M., Commons, M. L., Adams, K. M., Gutheil, T. G., & Sadoff, R. L. (2006). A pilot rasch scaling of lawyers' perceptions of expert bias. *The Journal of the American Academy of Psychiatry and the Law*, 34, 482–491.
- Felthous, A. R. (2022). Rational capacity and criminal responsibility in the USA. *International Journal of Law and Psychiatry*, 83, 101809. <https://doi.org/10.1016/j.ijlp.2022.101809>
- Gowensmith, W. N., Murrie, D. C., & Boccaccini, M. T. (2013). How reliable are forensic evaluations of legal sanity? *Law and Human Behavior*, 37(2), 98–106. <https://doi.org/10.1037/lhb0000001>
- Gröning, L. (2021). Has Norway abandoned its medical model? Thoughts about the criminal insanity reform post 22 July. *Criminal Law Review*, 191–202.
- Gröning, L., Haukvik, U., Meynen, G., & Radovic, S. (2020). Constructing criminal insanity: The roles of legislators, judges and experts in Norway, Sweden and The Netherlands. *New Journal of European Criminal Law*, 11(3), 390–410. <https://doi.org/10.1177/2032284420950>
- Guarnera, L. A., & Murrie, D. C. (2017). Field reliability of competency and sanity opinions: A systematic review and metaanalysis. *Psychological Assessment*, 29(6), 795–818. <https://doi.org/10.1037/pas0000388>
- Homant, R. J., & Kennedy, D. B. (1986). Judgment of legal insanity as a function of attitude toward the insanity defense. *International Journal of Law and Psychiatry*, 8(1), 67–81. [https://doi.org/10.1016/0160-2527\(86\)90084-1](https://doi.org/10.1016/0160-2527(86)90084-1)
- Kalis, A., & Meynen, G. (2014). Mental disorder and legal responsibility: The relevance of stages of decision making. *International Journal of Law and Psychiatry*, 37(6), 601–608. <https://doi.org/10.1016/j.ijlp.2014.02.034>
- Knoll, J. L. I. V., & Resnick, P. J. (2008). Insanity defense evaluations: Toward a model for evidence-based practice. *Brief Treatment and Crisis Intervention*, 8(1), 92–110. <https://doi.org/10.1093/brief-treatment/mhm024>
- Lov om straff, straffeloven – Lovdata. (2020). https://lovdata.no/dokument/NL/lov/2005-05-20-28/KAPITTEL_1_3#%C2%A720. Accessed March 3 29, 2023.
- Meyer, L. F., Leal, C. C. S., Omena, A. A. S., Mecler, K., & Valenca, A. M. (2020). Criminal responsibility scale: Development and validation of a psychometric tool structured in clinical vignettes for criminal responsibility assessments in Brazil. *Frontiers in Psychiatry*, 11, 579243. <https://doi.org/10.3389/fpsy.2020.579243>
- Meyer, L. F., & Valença, A. M. (2021). Factors related to bias in forensic psychiatric assessments in criminal matters: A systematic review. *International Journal of Law and Psychiatry*, 75, 101681. <https://doi.org/10.1016/j.ijlp.2021.101681>
- Meynen, G. (2010a). Free will and mental disorder: Exploring the relationship. *Theoretical Medicine and Bioethics*, 31(6), 429–443. <https://doi.org/10.1007/s11017-010-9158-5>

- Meynen, G. (2010b). Free will and psychiatric assessments of criminal responsibility: A parallel with informed consent. *Medicine, Healthcare & Philosophy*, 13(4), 313–320. <https://doi.org/10.1007/s11019-010-9250-7>
- Meynen, G. (2011). Autonomy, criminal responsibility, and competence. *Journal of the American Academy of Psychiatry and the Law*, 39, 231–236.
- Meynen, G. (2012). An ethical framework for assessments of criminal responsibility: Applying susan wolf's account of sanity to forensic psychiatry. *International Journal of Law and Psychiatry*, 35(4), 298–304. <https://doi.org/10.1016/j.ijlp.2012.04.011>
- Meynen, G. (2013). A neurolaw perspective on psychiatric assessments of criminal responsibility: Decision-making, mental disorder, and the brain. *International Journal of Law and Psychiatry*, 36(2), 93–99. <https://doi.org/10.1016/j.ijlp.2013.01.001>
- Meynen, G. (2016). *Legal insanity: Explorations in psychiatry, law, and ethics*. Springer.
- Meynen, G. (2022). Legal insanity in The Netherlands: Regulations and reflections. The insanity defence. In *International and comparative perspectives, warren brookbanks and ronnie mackay*. Oxford University Press.
- Miller, H. A. (2001). *M- fast: Miller forensic assessment of symptoms test and professional manual*. Psychological Assessment Resource: Odessea.
- Millon, T. (1983). *The Millon clinical multiaxial inventory manual* (3rd ed.). National Computer Systems.
- Millon, T. (1987). *Manual for the Millon clinical multiaxial inventory - II* (2nd ed.). National Computer Systems.
- Millon, T. (1994). *The Millon clinical multiaxial inventory - III manual* (3rd ed.). National Computer Systems.
- M"Nighten"s Case 10 C&F 200. (1843).
- Morey, L. C. (1991). *Personality assessment inventory: Professional manual*. Psychological Assessment Resources, Inc.
- Morse, S. J. (1998). Excusing and the new excuse defenses: A legal and conceptual review. *Crime and Justice*, 23, 329–406. <https://doi.org/10.1086/449273>
- Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. A. (2013). Are forensic experts biased by the side that retained them? *Psychological Science*, 24(10), 1889–1897. <https://doi.org/10.1177/0956797613481812>
- Parmigiani, G., Mandarelli, G., Meynen, G., Carabellese, F., & Ferracuti, S. (2019). Translating clinical findings to the legal norm: The defendant's insanity assessment support scale (DIASS). *Translational Psychiatry*, 9(1), 278. <https://doi.org/10.1038/s41398-019-0628-x>
- Parmigiani, G., Mandarelli, G., Meynen, G., Tarsitani, L., Biondi, M., & Ferracuti, S. (2017). Free will, neuroscience, and choice: Towards a decisional capacity model for insanity defense evaluations. *Rivista di Psichiatria*, 52(1), 9–15. <https://doi.org/10.1708/2631.27049>
- Parmigiani, G., Mandarelli, G., Roma, P., & Ferracuti, S. (2022). Validation of a new instrument to guide and support insanity evaluations: The defendant's insanity assessment support scale (DIASS). *Translational Psychiatry*, 12(1), 115. <https://doi.org/10.1038/s41398-022-01871-8>
- Rogers, R. (1992). *Structured interview of reported symptoms*. Psychological Assessment Resources.
- Rogers, R., Dolmetsch, R., & Cavanaugh, J. L. (1981). An empirical approach to insanity evaluations. *Journal of Clinical Psychology*, 37(3), 683–687. [https://doi.org/10.1002/1097-4679\(198107\)37:3<683::aid-jclp2270370343>3.0.co;2-f](https://doi.org/10.1002/1097-4679(198107)37:3<683::aid-jclp2270370343>3.0.co;2-f)
- Rogers, R., & Shuman, D. W. (2000a). *Conducting insanity evaluations, second edition* (2nd ed.). The Guilford Press.
- Rogers, R., & Shuman, D. W. (2000). The "mental status at the time of the offense" measure: Its validation and admissibility under daubert. *Journal of the American Academy of Psychiatry and Law*, 28, 23–28.
- Rogers, R., Wasylw, O. E., & Cavanaugh, J. L. (1984). Evaluating insanity. A study of construct validity. *Law and Human Behavior*, 8(3/4), 293–303. <https://doi.org/10.1007/BF01044697>
- Simon, R. J., & Ahn-Redding, H. (2006). *The insanity defense, the world over*. Bowman and Littlefields.
- Slobogin, C., Melton, G. B., & Showalter, C. R. (1984). The feasibility of a brief evaluation of mental state at the time of the offense. *Law and Human Behavior*, 8(3/4), 305–320. <https://doi.org/10.1007/BF01044698>
- Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the structured inventory of malingered Symptomatology (SIMS). *Journal of the American Academy on Psychiatry and Law*, 25(2), 183–189.
- van Wetboek, S. (1881). Wet van 3 maart. <https://wetten.overheid.nl/BWBR0001854/2023-01-01>. Accessed March 3 29, 2023.
- Wechsler, D. (1997). *Wechsler Adult intelligence scale* (3rd ed.). Psychological Corporation.

How to cite this article: Parmigiani, G., Mandarelli, G., Meynen, G., & Ferracuti, S. (2023). Structured instruments for insanity defense evaluations: Opportunities and limitations. *Behavioral Sciences & the Law*, 41(5), 432–444. <https://doi.org/10.1002/bsl.2618>