The mind in the courtroom: on forensic mental health reports in judicial decision-making about guilt and sentencing in the Netherlands

Es, R.M.S. van

3 The effects of forensic mental health reports on decisions about guilt in the Netherlands
An experimental approach

ABSTRACT

In the Netherlands, in approximately 25% of the more serious criminal cases, a pre-trial forensic mental health report (FMHR) is requested to inform the court whether a mental disorder was present at the time of alleged crime, whether this disorder affected behavior and decision-making at the time of the offense and how this disorder may affect future (criminal) behavior. While informative for sentencing decisions, information about mental disorders or recidivism risk is irrelevant for the question whether the defendant committed the alleged crime. Yet based on cognitive psychological theory of evidence evaluation and integration, we hypothesized that information in an FMHR would affect the evaluation of evidence as well as the ultimate decision about guilt. Using an experimental vignette study among 200 law and criminology students with manipulation of the presence and content of an FMHR, we found a main effect of the presence of an FMHR report on decisions about guilt. The proportion of guilty verdicts increased with almost 20% when an FMHR was present compared to when this report was absent, irrespective of the type of disorder (schizophrenia or antisocial personality disorder) or level of recidivism risk (low or high) present in the report. We did not find support for our hypothesis that this effect could be explained by assimilation of other available evidence. Implications for further research and practice are discussed.

This chapter is published as: Es, R.M.S. van, Keijser, J.W. de, Kunst, M. J. J. & Doorn, J. van (2022). The effects of forensic mental health reports on decisions about guilt in the Netherlands: An experimental approach. *International Journal of Law and Psychiatry, 80, 101760.*

A pilot study for this chapter was published in:

- Es R.M.S. van, Doorn J. van, Keijser J.W. de & Kunst M.J.J. (2020). Het effect van een pro Justitia-rapportage op de bewijsbeslissing: Een empirische verkenning. *Recht der Werkelijkheid 41*(2), 15-36.
- Es R.M.S. van, Keijser J.W. de, Kunst M.J.J. & Doorn J. van (2020). Het effect van een Pro Justitia rapportage op de bewijsbeslissing: Een pilot experiment. In R. Horselenberg, V. van Koppen & J.W. de Keijser (red.), *Bakens in de rechtspsychologie: Liber amicorum voor Peter van Koppen* (pp. 285-298). Boom criminologie.

3.1 Introduction

Mentally ill people are overrepresented in the criminal justice system and mental illness is more prevalent among prisoners than in the general population (Dirkzwager et al., 2021; Dorn et al., 2014; Favril & Dirkzwager, 2019; Fazel & Danesh, 2002; Fazel et al., 2016). In the United States, for instance, more mentally ill individuals are detained than admitted to mental health hospitals or treatment facilities (Fuller Torrey et al., 2010). Consequently, people with mental illness are profoundly prevalent in criminal procedures and information regarding mental illness can have a significant role in criminal trials. In the Netherlands, in approximately 25% of the more serious criminal cases, a pre-trial forensic mental health report (FMHR) is requested to inform the court (Nederlands Instituut voor Forensische Psychiatrie en Psychologie, 2020; Nederlandse Vereniging voor Psychiatrie, 2013). Forensic mental health experts evaluate whether a mental disorder was present at the time of the alleged crime, whether this disorder affected the defendant's behavior and decision-making at the time of the offense and how this disorder may affect future (criminal) behavior. This evaluation results in advice on criminal responsibility[1] (three degrees: no responsibility, diminished responsibility, full responsibility), an indication of recidivism risk and advice on possible treatment measures (e.g. Hummelen & Van der Wolf, 2018; see Koenraadt, 2010 for an English overview of this evaluation; Van Marle et al., 2013).

In Dutch criminal trials the court first determines the crucial question whether the alleged behavior was committed by the defendant. Second, it is determined whether the conduct constitutes a criminal act. Only if this second question has been answered affirmatively, the court proceeds to determine whether the defendant is criminally responsible, and therefore blameworthy. It is at this third stage and not earlier that mental health information is formally considered relevant (Keiler et al., 2017). Forensic mental health information is thus commonly used for assessing criminal responsibility and subsequently in sentencing decisions.

While not prohibited by Dutch procedural law, in practice forensic mental health information is logically irrelevant for determining the material facts in a case, specifically regarding the question whether the defendant committed the alleged crime (*actus reus*). In the Dutch inquisitorial system, judges rely heavily on the case file containing all evidence collected in the pre-trial investigation. In contrast to adversarial jurisdictions where the evidence is presented during a trial in accordance with strict rules of evidence (e.g. Federal Rule of Evidence 403) in a bifurcated trial, Dutch criminal judges are instantly exposed to all information relevant for both the decision about guilt and for subsequent decisions about criminal responsibility and sentencing. The court

1 In this manuscript 'criminal responsibility' is used to address blameworthiness of the defendant and to what extent a defendant is eligible for punishment and/or treatment.

also has substantial discretion to evaluate and combine the available evidence as they see fit, because Dutch procedural law does not regulate such integration (Dubelaar, 2014).

This unstructured feature of criminal fact finding may prompt psychological pitfalls that can bias decisions about guilt. Such pitfalls are well-documented in the literature (see for example Charman et al., 2019). Studies have shown, for example, that professional judges are susceptible to irrelevant factors (e.g. presented order of evidence, context effects) with respect to evidence evaluation and ultimately decisions about guilt (Rassin, 2017b, 2020). This susceptibility to irrelevant factors is further facilitated by the uncertainty that accompanies the intricate complex binary decision of a guilty verdict versus acquittal (Bodenhausen & Lichtenstein, 1987). Especially when the evidence in a case is weak and/or circumstantial (for example when a suspect denies the allegations) decision-makers may rely (subconsciously) more on experience and intuition to make a decision in addition to evaluation of the pieces of evidence (Epstein, 1994; Gunnell & Ceci, 2010; Kalven & Zeisel, 1966; Tversky & Kahneman, 1974).

As a result of the unstructured criminal fact finding in the Dutch procedure and because judges can be cognitively susceptible to irrelevant factors, it is plausible that information in an FMHR meant for the sentencing phase may inadvertently affect deliberations about guilt. However, very little empirical research has focused on this effect so far (see Chapter 2). Hence, the main research question in the current study is to determine to what extent and in what manner an FMHR affects decisions about guilt[2] in the Netherlands.

3.1.1 Theory

According to theories of holistic evidence evaluation and integration, a categorical decision in a complex criminal case is made by evaluating and integrating individual pieces of information to construct a coherent story that instigates a cognitive shift towards a guilty verdict or an acquittal (e.g. Pennington & Hastie, 1992; Pennington & Hastie, 1993; Simon, 2004).[3] According to the cognitive psychological model of coherence-based reasoning (Simon, 2004), once this decision has been made, a coherent representation of the case bidirectionally affects the evaluation of evidence by further inflation and assimilation of individual pieces of evidence to strengthen the coherency of

2 In this manuscript 'decisions about guilt' means the determination whether the suspect committed the alleged crime (*actus reus*). In the Netherlands, forensic mental health information is rarely used to determine criminal intent in terms of *mens rea* (see Chapter 2). Criminal responsibility is assessed after it is decided whether the suspect is guilty.

3 In the Netherlands the court may decide that the defendant committed the offense as charged only when the judges are *convinced* based on the legal evidence (Section 338 CCP). Whenever the court is not convinced of the defendant's guilt, they must acquit.

the story. This is an automated and thus subconscious decision-making process (Simon, 2004). When the evidence in a case is weak and/or circumstantial (especially when a suspect denies the allegations), irrelevant information can provide context to interpret the evidence. Consequently, in the process of shifting towards a coherent scenario, irrelevant context information can affect the relevant evidence in a case and ultimately affect the conviction decision (Simon, 2004).

Information provided by an FMHR can facilitate this effect of irrelevant contextual information (see Neal & Grisso, 2014) because such a report is specifically aimed at establishing an association between a disorder and the alleged criminal behavior. When information in an FMHR provides an adequate explanation for the alleged crime (e.g. a disorder that can explain sudden aggressive behavior when a defendant is suspected of a violent crime), this information may increase the perceived plausibility and coherence of a guilty scenario and result in a guilty verdict.

3.1.2 Prior research

Prior research on the effects of information in FMHRs on decisions about guilt is scarce (see review in Chapter 2). With regard to research on the effects of the specific presence of an FMHR on decisions about guilt, only one study focusing on the Dutch legal system has been done so far. Van Es et al. (2020) used an experimental vignette study among 155 students to study whether presence of an FMHR affected the incriminating value of evidence, the evaluation of guilt and ultimately the verdict. They tested whether the simple presence of an FMHR affected decisions about guilt or the specific diagnosis of a borderline personality disorder in the defendant accounted for the expected effect. Results showed that the mere presence of an FMHR without a disorder being diagnosed, did not significantly affect the verdict or evaluation of evidence. Yet an FMHR including a diagnosis of borderline personality disorder in the defendant significantly increased guilty verdicts with almost 30% but did not affect the evaluation of other evidence. This study served as a pilot study for the experiment in this chapter.

The focus of most prior research has remained on the effects of specific mental disorders on verdicts (Mossière & Maeder, 2015; Mowle et al., 2016; Rassin, 2017b; Termeer & Szeto, 2021). These studies vary in how they presented information about these mental disorders to their respondents, the type of respondents they used, the type of disorders they studied and legal systems in which the research was done. For example, Rassin (2017b) used an experimental vignette study among a sample of professional judges from the Netherlands to study whether irrelevant information about a diagnosis of antisocial personality disorder and psychopathy provided by a psychiatrist would assimilate the incriminating value of the evidence and ultimately affect

decisions about guilt. Results showed that the presence of the disorders assimilated the evaluation of the evidence and increased the proportion of guilty verdicts with 33% compared to when information about these disorders was absent. A similar effect of information about psychopathy was found by Mowle et al. (2016) in an experimental vignette study among 419 jurors in the United States. They found that expert testimony by a psychologist about psychopathy in the defendant significantly increased guilty verdicts compared to testimony about schizophrenia. A guilty verdict was less likely when the defendant suffered from schizophrenia, but only when jurors had a liberal political orientation. Most recently, Termeer and Szeto (2021) conducted an experimental vignette study among 248 students, in which they compared a defendant with a history of schizophrenia to a defendant with a history of depression or a healthy defendant. They found that a defendant with schizophrenia was less likely to be found guilty compared to a defendant with a history of depression. In addition, they examined whether mental illness affected perceived dangerousness of the defendant based on prevalent negative stereotypes between mental illness and dangerousness and violence in the general population (Angermeyer & Dietrich, 2006; Link et al., 1999; Pescosolido et al., 1999). While perceived dangerousness was positively correlated to a guilty verdict, presence of a mental illness had no effect on these perceptions of dangerousness (Termeer & Szeto, 2021).

Finally, Mossière and Maeder (2015) studied effects of information about mental disorders often associated with violent behavior (i.e. schizophrenia and substance abuse) compared to mental disorders that have no such association (i.e. depression and obsessive-compulsive disorder) and a healthy defendant in two different samples: students and a community sample of jury-eligible Americans. Information about mental illness was presented as part of the alibi of the defendant and not by a forensic expert in the trial. There was no significant effect of any mental illness on the verdict for either sample.

Hence, the limited amount of research that has been carried out showed inconsistent results and because of considerable variability regarding how they presented information about these mental disorders to their respondents, the type of respondents they used, the type of disorders they studied and legal systems in which the research was done, any conclusions based on these results are tentative. Research on effects of FMHRs on decisions about guilt in the Netherlands is still scarce. Therefore, the aim of this study was to explore to what extent and in what manner an FMHR affects decisions about guilt.

3.1.3 Current study

The central research question of this study was to what extent and in what manner an FMHR affects decisions about guilt in the Netherlands. Based on

this research question, the theory and limited prior research we identified three main hypotheses that we explored:

1) Presence of an FMHR in a case with weak and circumstantial evidence will increase the incriminating value of available evidence and result in more guilty verdicts compared to when an FMHR is absent.
2) Presence of an FMHR *with a disorder* (irrespective of its nature) in a case with weak and circumstantial evidence will increase the incriminating value of the available evidence and result in more guilty verdicts compared to when a diagnosis of a disorder is not present in the FMHR.

In accordance with prevalence of specific disorders in the forensic population in the Netherlands (Kempes & Gelissen, 2020; Vinkers et al., 2011) and based on the inconsistent and diverging results in the previously discussed research (Mowle et al., 2016; Rassin, 2017b; Termeer & Szeto, 2021), we further examined the effects of the presence of two specific mental disorders: schizophrenia and personality disorder with antisocial traits. We hypothesized that:

3) Diagnosis of a personality disorder with antisocial traits will increase the incriminating value of the available evidence and result in more guilty verdicts compared to a diagnosis of schizophrenia.

Previous research (Mossière & Maeder, 2015; Termeer & Szeto, 2021) has focused on stigmatization as an explanation of an effect of mental illness on verdicts by studying perceptions of future risk and dangerousness (Termeer & Szeto, 2021) or studying mental disorders stereotypically associated with violence (Mossière & Maeder, 2015). While these studies focused on perceptions of risk, we wanted to explore whether actual information about recidivism risk would contribute to the relation between mental disorder and decisions about guilt. Since indication of recidivism risk is a crucial part of a Dutch FMHR, we explored whether an effect of mental disorder on decisions about guilt varied according to information about recidivism risk. In order to study the research questions and test these hypotheses, we conducted an experimental vignette study among law and criminology students. The experimental design and procedure are explained in the *Method* section (paragraph 3.2) after which the results are presented in paragraph 3.3. Paragraph 3.4 discusses these results and the implications for future research and practice.

3.2 Method

3.2.1 Participants and procedure

Participants were 307 students recruited from law and criminology courses at 9 universities in the Netherlands.[4] A large number of participants (n = 107, 34.9%) was removed because they failed (at least) one of the manipulation checks about the information given in the vignette (see paragraph 3.2.2.3). The final sample consisted of 200 participants.[5] The majority of the sample were young (M = 22.03 years; SD = 4.50), female (85%) criminology students (73%) in their third year of undergraduate studies (35.5%).[6]

Participants were recruited through virtual learning environments and websites of multiple Dutch universities (e.g. Blackboard, Brightspace, Canvas) and via social media (e.g. Facebook and Instagram). The recruitment message provided a Qualtrics[7] link to a 15-minute survey. When participants clicked on the Qualtrics link, they were presented with a consent form. After giving informed consent, participants were directed to the case summary. No incentives were given for participation. This study was approved by the Committee of Ethics and Data of Leiden Law School.

3.2.2 Materials and measures

3.2.2.1 Vignette

All participants received a summary of a case file (approximately 1200 words) resembling an actual case file used in Dutch criminal proceedings (see Appendix A). The vignette was adopted and adapted from a study by De

4 In most inquisitorial legal systems, including the Netherlands, judicial decisions are made by professional judges. However, it is particularly difficult to use professional judges as participants in experimental research, because experimental designs often require large sample sizes. Permission to conduct this experiment among professional judges was unfortunately denied because it would produce an overload on the courts, according to the Council of Judiciary. To obtain a sufficient sample size, we conducted this experiment among law and criminology students. Law and criminology students are more representative for professional judges than other student populations (e.g. psychology students) often used in prior research on this topic, because they are more familiar with the materials and decisions they were required to respond to. In a pilot study by Van Es et al. (2020) we established that these students made similar decisions to professional judges (Rassin, 2017b) with regard to the effects of forensic mental health information on decisions about guilt.

5 A power analysis in G*Power (Faul et al. 2007) suggests a sample size of 200 participants provides 80% power to detect a, relatively small, main effect for each hypothesis (w = 0.2, power = 0.8, α = 0.05; cf. Allen e.a. 2019).

6 First year undergraduate: 27%; second year undergraduate: 4.5%; fourth year or older undergraduate: 11.5%; master's: 21.5%.

7 Qualtrics survey software is a tool used to create and conduct online survey research.

Keijser and Van Koppen (2004, 2007). In this fictitious but realistic case, a male defendant was accused of aggravated assault with serious bodily harm (section 302 CC). The defendant and his girlfriend had broken up on the day of the assault. The defendant went on a night out with two friends. He had multiple beers. While on their way home, they crossed paths with the victim and the victim's girlfriend. The defendant and the victim, as well as the victim's girlfriend, did not know each other. The defendant and the victim had had an argument about something the defendant had said to the victim's girlfriend. After that, the defendant allegedly followed the victim and his girlfriend and attacked the victim. He allegedly kicked the victim against his body and head multiple times. The physical trauma resulted in loss of memory, loss of speech and permanent paralysis in the victim according to a neurologist. Other than the girlfriend, no one witnessed the assault, and the defendant denied all allegations. The case file contained legally sufficient, but relatively weak and circumstantial evidence in order to facilitate doubt whether the suspect committed the alleged crime. This doubt was necessary to determine whether the manipulation of information in the FMHR would affect the evaluation of guilt and ultimately the verdict (see De Keijser & Van Koppen, 2004). The information in the case file consisted of: 1) two interrogations in which the defendant denies all allegations; 2) a statement about the assault by the victim's girlfriend; 3) statements of two friends of the defendant on the situation prior to the assault. They went home before the alleged assault took place; 4) a hesitant identification of the defendant by the victim's girlfriend in a photo line-up; 5) statement by a neurologist on the injuries of the victim. The statements and identification procedure were indicative but inconclusive of guilt. The study by De Keijser and Van Koppen (2004, 2007) among professional judges showed that this vignette, as intended, facilitated doubt about the defendant's guilt since 77% of the judges provided a guilty verdict.

3.2.2.2 Design

After reading the case summary, participants were randomly assigned to one condition in the 2 (Diagnosis: personality disorder with antisocial traits or schizophrenia) x 3 (Recidivism risk: low risk or high risk or no information provided) between-subjects factorial design or to one of the two control conditions (no FMHR or an FMHR without disorder and recidivism risk information

due to refusal to cooperate with the evaluation),[8] making a total of 8 conditions.

In the conditions in which the information about a mental disorder and recidivism risk were manipulated, a fictitious and simplified forensic mental health evaluation (between 330 and 400 words) by both a psychologist and psychiatrist was presented to participants (see Appendix A). Use of language in the reports was based on actual FMHRs to make them as realistic as possible.

The first diagnosis was a personality disorder with antisocial traits. In addition to the diagnosis, symptoms of the disorder (e.g. aggressive impulses, lack of empathy, impairment of impulse control and frustration; American Psychiatric Association, 2013) were described. This facilitated a similar and comparable interpretation of the disorders between participants. The other diagnosis was schizophrenia not otherwise specified [NOS]. Symptoms included impulsive aggression, hallucinations and delusions (American Psychiatric Association, 2013). Descriptions and labels of the disorders were based on actual Dutch FMHRs. Regardless of diagnosis, all evaluations contained information on the contribution of the disorder to the alleged offense, along with a preliminary advice on criminal responsibility (in this case diminished responsibility). Since the defendant denied any involvement in the offense, no adequate treatment advice could be given in any of the conditions.

The second between-subject factor that was manipulated was recidivism risk. In the conditions with a diagnosis, participants either received information indicating low recidivism risk, high recidivism risk (both based on the Historical Clinical Risk Management-20, version 3 [HCR-20v3]; Douglas et al., 2014) or no additional information about recidivism risk.

3.2.2.3 Questionnaire

Participants were able to review the case summary and the forensic mental health evaluation while completing the questionnaire. First, participants rated the factual evidence on a Likert scale ranging from 1 (*Not incriminating at all)* to 10 (*Very incriminating*). Evidence included:

1) Identification of the defendant by the victim's girlfriend;
2) Witness statement by friend no. 1;
3) Witness statement by friend no. 2.

8 This condition was added because in recent years, the number of 'refusers' in clinical forensic mental health assessment in the Netherlands has increased from 23% in 2002 to 43% in 2017 (Nagtegaal et al., 2018). Arguments for this refusal are that information gathered in a forensic mental health evaluation can be used to sanction an offender to extensive treatment measures (in Dutch: 'tbs maatregel') instead of or in addition to a prison sentence (section 37a CC). For most violent crimes, this treatment measure can be enforced for an unlimited period of time and therefore exceed a (maximum) prison sentence imposed for the same offense (section 38e CC).

Next, participants indicated whether they found the defendant guilty or not and indicated how convinced they were of the defendant´s guilt on a scale ranging from 1 (*Not convinced*) to 10 (*Very convinced*). Finally, some questions on demographics (gender, age, study course, university and year of studies) were asked. Factual manipulation checks at the end of the questionnaire were used to ensure participants were sufficiently exposed and attentive to the manipulated factors. Participants were not able to review the case or FMHR once they had reached the manipulation check questions. Three multiple choice questions on the presence and information in the forensic mental health evaluation were presented:

1) Was an FMHR present in the case?

If yes,

2) What disorder was diagnosed with the defendant?
3) What was the predicted recidivism risk?

If participants answered at least one question incorrectly, they were excluded from the analyses. As a result of the strict check which was postponed to the end of the questionnaire where all three questions had to be answered correctly to pass, a substantial proportion of participants (34.9%) was removed from further analyses. This is not uncommon in online experimental research (Thomas & Clifford, 2017).

3.2.3 Analyses

The current study used three main outcome measures: the ultimate verdict, evaluation of guilt and evaluation of evidence. First, Chi square tests were used to determine whether the proportion of guilty verdicts differed between conditions. For explorative interaction effects we used logistic regression analysis. Second, independent sample t-tests and Mann-Whitney[9] tests were used to determine whether there were significant effects on the evaluation of guilt. Analyses of variance were used to explore interaction effects. Finally, (multivariate) analyses of variance were used to test (interaction) effects on the evaluation of the evidence.

9 For the majority of groups, assumptions of normality were violated (based on visual inspection, values of Kolmorogov-Smirnov tests and values of kurtosis and skewness). Therefore, for most analyses non-parametric Mann-Whitney test are reported. The analyses were also performed using independent sample t-tests. Results of the t-tests were not different unless stated otherwise.

3.3 RESULTS

3.3.1 Descriptive and preliminary analyses

There were no significant differences between the conditions regarding sex, age, type of studies and year of studies of the participants. Spearman's rho correlations indicated that the evaluations of the three individual pieces of evidence were all significantly correlated (Spearman's rho = .190-.711, p <.001; α = 0.68). Because of these correlations and sufficient internal consistency of the items, the three combined pieces of evidence is an acceptable measure of the total evidence evaluation. Therefore, we also analyzed the average combined score of the three individual pieces of evidence. Across all conditions the conviction rate was 82% (n =164). Participants who supported a guilty verdict rated all evidence as stronger ($Mdn_{combined}$= 6.00, U = 1173.500, z = -5.67, p <.001) and were also more convinced of the defendant's guilt (Mdn = 7.00, U = 556.000, z = -7.80, p <.001) compared to those who acquitted (Evaluation of evidence: $Mdn_{combined}$ = 4.67; Evaluation of guilt: Mdn = 4.00). The evaluation of guilt was moderately correlated to the final verdict (Spearman's rho = .553, p <.001).

3.3.2 Hypothesis 1: Effect of the presence of an FMHR

3.3.2.1 Verdict

In order to test the first hypothesis whether the presence of an FMHR had an effect on decisions about guilt, we compared the groups with and without the presence of an FMHR. Table 3.1 shows that a guilty verdict was more likely when an FMHR was present (85%), compared to the control condition in which a report was absent (66,7%; χ^2 (1) = 6.295, p =.012, φ = .177). The proportion of guilty verdicts increased with 18.3%.

3.3.2.2 Evaluation of guilt

The evaluation of guilt showed similar results: when an FMHR was present, evaluation of guilt was also higher (M = 6.54, SD = 1.50, Mdn = 7.00) compared to when this report was absent (M = 5.91, SD = 1.84, Mdn = 6.00), but this effect was not significant (U = 2238.000, z = -1.874, p = .081).

3.3.2.3 Evaluation of evidence

Table 3.1 shows the mean scores for each individual piece of evidence as well as a combined score for the three pieces of evidence. No significant effects of the presence of an FMHR on the evaluation of evidence were found, neither

for the individual pieces (F (3, 196) = 1.292, v = 0.019, p = .278, η_p^2 = .019) nor for the combined score (t (198) = 1.153, p = .250, d = 0.22). Participants seem to evaluate the evidence similarly in both conditions.

Table 3.1: Effect of an FMHR

	Condition	
	FMHR absent	FMHR present
N	33	167
Guilty verdict (%)	66.7%	85%*
Evaluation of guilt (M, SD)	5.91 (1.84)	6.54 (1.50)
Evaluation of evidence (M, SD)		
Identification	4.67 (1.85)	5.28 (1.66)
Statement (1)	6.03 (1.59)	6.20 (1.75)
Statement (2)	5.21 (1.54)	5.32(1.54)
Total	5.30 (1.38)	5.60 (1.35)

Note. FMHR = forensic mental health report; *M* = mean; *SD* = standard deviation; * $p < .05$.

3.3.3 Hypothesis 2: Effect of information about any disorder

3.3.3.1 Verdict

Within the FMHR present condition, we proceeded to compare the two groups with information about a disorder (so either schizophrenia NOS or a personality disorder with antisocial traits) with an FMHR without information about a disorder. Table 3.2 shows that there was no significant difference between the two groups (Fisher's exact test, one-sided p = .499, φ = -.028).

3.3.3.2 Evaluation of guilt

The evaluation of guilt showed similar results as no significant effect of the presence of a disorder was found (U = 1652.500, z = -.297, p = .766).

3.3.3.3 Evaluation of evidence

Finally, we found no significant effect of information about a disorder on the evaluation of the evidence, neither for the individual pieces using multivariate analysis of variance (F (3, 163) = 1.592, v = 0.028, p = .193, η_p^2 = .028) nor for the combined score (t (165) = 1.405, p = .162, d = -0.33). Surprisingly, univariate

analysis of variance showed that the evidence of identification of the defendant by the victim's girlfriend was evaluated as significantly more incriminating when no disorder was present (M = 5.96, SD = 1.30) compared to when information about a disorder was present (M = 5.17, SD = 1.69, F (1, 165) = 4.784, p = .030, η_p^2 = .028).

Table 3.2: Effect of information about any disorder

	Condition	
	Disorder absent	Disorder present
N	24	143[a]
Guilty verdict (%)	87.5	84.6
Evaluation of guilt (M, SD)	6.71 (1.27)	6.51 (1.54)
Evaluation of evidence (M, SD)		
Identification	5.96 (1.30)*	5.17 (1.69)
Statement (1)	6.38 (1.71)	6.17 (1.76)
Statement (2)	5.54 (1.79)	5.29 (1.86)
Total	5.95 (1.26)	5.54 (1.36)

Note. [a] = this condition is a combination of the 6 conditions in which any disorder was diagnosed. Therefore the sample size is larger compared to the condition in which the disorder was absent; M = mean; SD = standard deviation; * $p < .05$.

3.3.4 Hypothesis 3: Effect of type of disorder

3.3.4.1 Verdict

In accordance with the third, and final, hypothesis, we studied whether the diagnosis of a personality disorder with antisocial traits leads to more guilty verdicts compared to a diagnosis of schizophrenia (see Table 3.3). We found no significant differences (χ^2 (1) = 0.103, p =.748, φ = .027).

Furthermore, we explored whether there was an effect of recidivism risk within the conditions with a diagnosed disorder. We first analyzed whether the mere presence of information about recidivism risk affected the verdict. The analysis showed that when information about recidivism risk (irrespective of whether this was high or low) was *absent*, guilty verdicts were significantly higher (93.8%) than when this information was *present* (80%)[10] (χ^2 (1) = 4.631, p =.031, φ = -.180). Next, we explored whether there was a significant difference

10 Combination of the high risk and low risk conditions.

between information about low recidivism risk and high recidivism risk. No significant differences were found (χ^2 (1) = 0.096, p =.757, φ = .032). Finally, we found no interaction effect between the type of disorder and presence of information about recidivism risk on the verdict (b = -.190, p = .848, 95% CI [0.118;5.768], χ^2 (2) = 0.140, p =.932).

3.3.4.2 Evaluation of guilt

The evaluation of guilt showed similar results to the verdict: no significant effect of type of disorder (U = 2170.000, z = -1.557, p = .119). Although the evaluation of guilt also showed similar results to the verdict with regard to the effect of information about recidivism risk, this effect was not significant (U = 2015.500, z = -1.157, p = .247). The interaction between type of disorder and recidivism risk was not significant either (F (2, 137) = 0.161, p = .851, η_p^2 = .002).

3.3.4.3 Evaluation of evidence

Finally, no significant effects of type of disorder on the evaluation of the evidence were found, neither for the individual pieces (F (3, 139) = 0.871, v = 0.018, p = .458, η_p^2 = .018), nor for the combined score (t (141) = -.808, p = .413, d = -0.14). We neither found main effects of information about recidivism risk nor interaction effects between type of disorder and recidivism risk, for either the individual pieces of evidence (Main effect: F (6, 272) = 0.618, v = 0.027, p = .716 η_p^2 = .013; Interaction effect: F (6, 272) = 0.756, v = 0.033, p = .605, η_p^2 = .016), or for the combined score (Main effect: F (2, 137) = 1.289, p = .279, η_p^2 =.018; Interaction effect: F (2, 137) = 0.283, p = .754, η_p^2 = .004).

Table 3.3: Effect of type of disorder

	Conclusion	
	Schizophrenia NOS	Personality disorder with antisocial traits
N	76	67
Guilty verdict (%, overall)	85.5	83.6
- Low risk (n = 52)	77.8	80
- High risk (n = 43)	84.2	79.2
- No info on risk (n= 48)	93.3	94.4
Evaluation of guilt (M, SD)	6.66 (1.61)	6.34 (1.44)
- Low risk (n = 52)	6.48 (1.85)	6.20 (1.26)
- High risk (n = 43)	6.10 (1.44)	6.29 (1.55)
- No info on risk (n= 48)	6.73 (1.53)	6.61 (1.58)
Evaluation of evidence (M, SD)		
Identification	5.33 (1.69)	4.99 (1.67)
Statement (1)	6.26 (1.75)	6.06 (1.77)
Statement (2)	5.29 (1.94)	5.28 (1.78)
Total	5.63 (1.35)	5.44 (1.37)

Note. NOS = not otherwise specified; *M* = mean; *SD* = standard deviation.

3.4 DISCUSSION

The main objective of this experimental vignette study was to explore the extent and the manner in which an FMHR affects decisions about guilt in the Netherlands. Based on the theory of coherence-based reasoning in evidence evaluation and integration (Simon, 2004), the organization of the Dutch criminal trial and results from the scarcely available prior research (Mossière & Maeder, 2015; Mowle et al., 2016; Rassin, 2017b; Termeer & Szeto, 2021; Van Es et al., 2020), we hypothesized that 1) presence of an FMHR in a case with weak and circumstantial evidence will increase the incriminating value of available evidence and result in more guilty verdicts compared to when an FMHR is absent; 2) presence of an FMHR *with a disorder* (irrespective of its nature) in a case with weak and circumstantial evidence will increase the incriminating value of the available evidence and result in more guilty verdicts compared to when a diagnosis of a disorder is not present in the FMHR. And finally, 3) diagnosis of a personality disorder with antisocial traits will increase the incriminating value of the available evidence and result in more guilty verdicts compared to a diagnosis of schizophrenia.

Our experiment demonstrated that the mere presence of an FMHR increased the likelihood of reaching a guilty verdict, supporting the first hypothesis.

The proportion of guilty verdicts significantly increased with 18.3% when an FMHR was present compared to when an FMHR was absent. This result should be interpreted with care since the effect size was relatively small ($\varphi = .177$) (Cohen, 1988). No effects of the presence of an FMHR were found on the evaluation of evidence or the evaluation of guilt. We also did not find support for the second hypothesis in which we expected that the presence of any disorder would affect decisions about guilt. Apart from a significant difference regarding evidence of the identification of the defendant, no substantive effects on decisions about guilt were found. Finally, we also found no effect of type of disorder on evaluation of evidence or decisions about guilt. Our exploration of a potential effect of recidivism risk, surprisingly, revealed that when no information on recidivism risk was provided in an FMHR, the number of guilty verdicts was significantly *higher* compared to when information about any recidivism risk was given in the report. This finding suggests that information on recidivism risk in combination with the information about a mental disorder, limits biased decisions about guilt compared to only information about a disorder without presence of any information about future risk. Information about recidivism risk did not affect the evaluation of evidence or the evaluation of guilt and no interaction effects between type of disorder and information about recidivism risk on any of the decisions were found. It is possible that these explorative tests for (interaction) effects were underpowered as a result of smaller sample sizes in these analyses.

Nevertheless, the most important finding remains that when an FMHR is available in a case of a violent crime, its mere presence can bias decisions in favor of a guilty verdict regardless of the content of the report. We did not find support for our hypothesis that any effect could be explained by assimilation of other available evidence. Information about mental illness of a defendant acts as an incriminating context to help construct a coherent guilty scenario, but this does not appear to be reflected in the evaluation of available evidence, as suggested by the theory of coherence-based reasoning (Simon, 2004) or other theories of holistic evidence evaluation (Pennington & Hastie, 1992, 1993). The evidence in our experiment was evaluated in a similar manner across all conditions. A more general context effect of an FMHR may be the underlying mechanism: a mental disorder can provide a fitting explanation for a violent crime and when the evidence is weak or circumstantial in a case, this explanation may, legitimately or not, be considered to support a guilty scenario. The result that the proportion of guilty verdicts in the conditions without information about recidivism risk was significantly higher than in the conditions with information about recidivism (an increase of 13.8%) may provide further support for this stereotypical association of mental disorder and violent behavior (Angermeyer & Dietrich, 2006; Link et al., 1999; Pescosolido et al., 1999). Providing information about recidivism risk may have prompted a minor barrier (proportion of guilty verdicts was still 80%) on the coherency of a guilty scenario.

A context effect by an FMHR may already have its origins in the pre-trial forensic mental health evaluation and consequently result in confirmation bias during trial (Neal & Grisso, 2014; Nickerson, 1998). In Dutch practice, prior to the actual evaluation, forensic mental health evaluators receive information about the crime and the defendant from either the prosecutor or the court (Koenraadt, 2010). In their evaluation experts may therefore be inclined to work with the hypothesis that the defendant is guilty, even if this person denies all allegations (Crombag et al., 2005). Part of the evaluation is, among other things, to discuss the alleged crime with the defendant. The evaluator is then asked to research whether there is a psychopathological explanation for the alleged offense. A report that contains any explanation of how psychopathology in a defendant is related to the alleged criminal behavior can facilitate a confirmation bias towards a guilty scenario during trial (de Ruiter, 2010; Neal & Grisso, 2014; Van Koppen, 2004). This potential cumulative effect should be studied further.

The present study contributes to the existing body of (international) empirical literature by providing further insight into the potentially biasing effects of information in FMHRs on decisions about guilt. The current study has partially confirmed the findings of both Dutch studies by Rassin (2017b) and by Van Es et al. (2020), as our study did find an effect of an FMHR on decisions about guilt, but not yield an effect of an FMHR on evaluation of evidence or an effect of a specific disorder.

The current study also expands upon this literature by additionally exploring the effects of different types of disorder as well as a possible effect of recidivism risk. Contrary to a number of prior studies (Mowle et al., 2016; Termeer & Szeto, 2021), we did not find a significant difference between a diagnosis of schizophrenia and a personality disorder with antisocial traits. An explanation for this finding is that both disorders can be associated with violent behavior (e.g. Fazel et al., 2009; Yu et al., 2012) and consequently provide a plausible explanation for the alleged crime of aggravated assault and thus elicit a similar effect. Furthermore, aggressive symptoms overlap between the two disorders and were described in a similar way in the vignettes (e.g. impulsive aggression and aggressive impulses). In relation to the type of offense (aggravated assault), it is not unlikely that participants placed more emphasis on the symptoms of aggression than on the label attached to these symptoms. In fact, this further emphasizes that it does not matter what type of psychopathology is diagnosed to elicit bias in a criminal case as long as the symptomatology is congruent with the violent behavior (see Berryessa & Wohlstetter, 2019 for a recent meta-analysis on a general labelling effect of mental disorder on punishment outcomes).

3.4.1 Study limitations and strengths

The results in the current study are accompanied by a number of limitations. First, due to multiple strict manipulation checks, a substantial number of participants were eliminated from the study, therefore reducing sample size. A substantial exclusion rate is however not uncommon in online survey research. Moreover, the exclusion may even increase statistical power by reducing statistical noise and without introducing significant sampling bias[11] (Thomas & Clifford, 2017). Nonetheless, some analyses (i.e. interaction effects) may have been underpowered due to smaller sample sizes in these analyses.

Second, the external validity of the current study is limited. The use of a vignette allows us to study complex social situations without confounding variables, thereby enabling us to observe a (causal) effect of information in an FMHR on decisions about guilt (Hughes & Huby, 2004). While this method optimizes internal validity, this is usually at the expense of external validity and especially ecological validity (Atzmüller & Steiner, 2010; Sniderman & Grob, 2003). This vignette study used simplified stimulus materials (i.e. summaries from relevant parts of a fictitious case file) and a research setting (i.e. online with students) which does not correspond with actual practice in the Dutch legal system, in which professional judges decide in a panel of three in more serious cases, such as aggravated assault. We were primarily concerned with maximizing the internal validity of our study (i.e. our ability to minimize the influence of potential confounding factors such as poor comprehension) to study unintended and subconscious effects of an FMHR on decisions about guilt, rather than its direct generalizability to actual trials.

Notwithstanding these limitations, this study does have some notable strengths. First, it is among the first studies to extensively research potential subconscious biasing effects of FMHRs on decisions about guilt in the Netherlands. The focus of most prior research is on cognitive bias at the stage of the pre-trial forensic mental health evaluation (de Ruiter, 2010; Murrie et al., 2013; Neal & Grisso, 2014; Rassin & Merckelbach, 2014; Van Koppen, 2004; Zapf et al., 2018) without studying any subsequent effects on the ultimate judicial decisions. Second, the current study elaborated upon initial indications of unwarranted effects by FMHRs with one type of disorder (Rassin, 2017b; Van Es et al., 2020). We studied multiple different disorders prevalent among the Dutch forensic population (Kempes & Gelissen, 2020; Vinkers et al., 2011) and also explored whether a potential bias by an FMHR varied according to information about recidivism risk. Despite simplified stimulus materials, all information in the vignettes was representative of an actual case file and actual FMHRs. Finally, while a sample of students affects external validity of the results, law and criminology students are more representative for professional judges in

11 Analyses showed no significant demographic differences (gender, age and studies) between participants who passed or failed the manipulation checks.

the Netherlands than other student populations (such as psychology students) that have been used in prior research on this topic.

3.4.2 Implications

Based on the results and limitations of this study a number of recommendations for future research can be made. First, although we found a general biasing effect in favor of a guilty verdict of the mere presence of an FMHR, there are still many unanswered questions related to the underlying mechanism of this effect. We did not find support for a mechanism of assimilation of evidence by the FMHR. Therefore, future research should focus on further exploring the effect of cognitive bias, both in samples of students (as potential jurors, depending on jurisdiction) as well as among professional judges since the processing of information in an FMHR seems to differ between these populations (cf. Rassin, 2017b).

Second, although the effect of different types of mental disorder was examined, this study focused on only one type of (violent) crime. Yet offenders with mental disorders are heterogeneous in types of disorder they suffer from, as well as in types of crime they commit (Vinkers et al., 2011). The current study, as well as most prior research, has focused on severe violent crimes and disorders that are compatible with violent behavior (Mowle et al., 2016; Rassin, 2017b; Termeer & Szeto, 2021). Future research should focus on whether an FMHR still causes bias in decisions about guilt, when a disorder is not, or less, compatible with the type of crime (e.g. a psychotic disorder with certain sex crimes) (Vinkers et al., 2011). Moreover, many individuals in the (Dutch) forensic population suffer from comorbid disorders, often with substance abuse (e.g. Kempes & Gelissen, 2020; Ogloff et al., 2004; Ogloff et al., 2015). There already is much debate both in the literature and in practice about the implication of (comorbid) substance abuse for questions regarding criminal responsibility and subsequent sentencing (e.g. Goldberg, 2022; Kennett et al., 2015; Morse, 2013). Future research should therefore also take potential unwarranted effects and bias by substance abuse into consideration.

Additionally, future studies should consider an interaction between information about a mental disorder and severity of the crime (e.g. violent crimes versus property crimes), thereby exploring whether an unwarranted effect is stronger in case of a severe crime as can be argued by the conviction paradox. This paradox describes the tendency to be satisfied with less evidence to become convinced about a defendant's guilt in a more serious case compared to a less serious offense, because the consequences of a false-negative decision (i.e. a guilty individual is acquitted) are considered more severe for society than in the case of a less serious crime. This situation produces a paradox because one would expect decision-makers to be especially careful when evaluating the evidence in a more serious case, because the consequences of

a false positive decision (i.e. wrongful conviction) for a defendant are very serious in case of a more severe crime (De Keijser & Van Koppen, 2007).

Finally, a biasing effect of an FMHR on decisions about guilt in the Netherlands has now been demonstrated in a number of vignette studies (Rassin, 2017b; Van Es et al., 2020). The next step is to study this effect in a more external and ecologically valid setting with professional judges reading a realistic case file and are allowed to make elaborated decisions with three judges as is practice in severe criminal cases. More ecologically valid (qualitative) research on this potentially biasing effect, will provide more understanding of the extent and underlying psychological mechanisms of this issue and whether judges are aware of such effects.

3.4.3 Conclusion

In this explorative study, we demonstrated that the mere presence of an FMHR, regardless of its content, can bias decisions in favor of guilt. This result generates new possibilities for further research into its underlying cognitive mechanisms and into the generalizability to other types of disorders, types of crime and jurisdictions. Since people with mental illness are frequently present in criminal procedures, research is necessary to determine the extent of potential effects of bias by an FMHR on decisions about guilt to gain more insight in factors that are used in these decisions. This is important to prevent wrongful convictions and increase legitimacy and credibility of judicial decision-making.