

Detecting the True Nature of Allegations of Rape

André De Zutter¹ · Robert Horselenberg¹ · Peter J. van Koppen²

Published online: 23 June 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract A study was conducted to test whether it is possible to build a model to distinguish true and false allegations of rape based on the theory of fabricated rape. The theory is based on the principle that a false complainant of rape has not been raped and has to fabricate a story while the story of a true victim is based on recollections of the event. Consequently, false complainants will behave as liars do, construct their story based on their own sexual experiences and on mental representations, beliefs of how such a crime would happen (De Zutter et al. in *Eur J Psychol Appl Leg Context*. doi:10.1016/j.ejpal.2016.02.002, 2016). To test the theory and to build a model to discriminate between true and false allegations of rape, a police sample of true and false allegations was studied. A total of 129, 72 true and 57 false, allegations of rape fulfilled the stringent criteria of the current study, among others on ground truth. Fifty-four allegations of rape, 27 true and 27 false, were used to build a prediction model based on the theory of bounded rationality by Gigerenzer (2002). The remaining 75 cases, 45 true and 30 false, were blindly categorised as either true or false based on

the model. The model was able to predict the true nature of the majority of allegations with an accuracy rate of 91 %. Thus, it seems possible to discriminate to a considerable extent between true and false allegations of rape.

Keywords Rape · Allegation · True · False · Characteristics

Introduction

The Problem of False Allegations of Rape

False allegations of rape are a problem for all parties involved. In false allegations of rape, the perpetrator-victim relationship is reversed. The dire consequences of a false allegation of rape are best illustrated by the case of Gary Dotson. He was the first DNA exoneree in the USA (Gross et al. 2005). He became the victim of a false allegation made by Kathleen Crowell Webb and spent 10 years in and out of prison as a consequence. Webb made such a compelling allegation that even when she recanted her allegation and told the truth, not everyone believed that the rape had never occurred (Taylor 1987). It was not until the advent of adequate DNA testing that Gary Dotson was fully vindicated (Gross et al. 2005). In cases where no one is accused directly, no innocent suspects are targeted, false allegations still result in a dissipation of limited police resources and cause public distress. False allegations also raise unnecessary suspicion towards true rape victims (Bohmer and Blumberg 1975). It might, therefore, be helpful to detect false complainants, as soon as possible, before they can cause harm.

The Theory of Fabricated Rape

De Zutter et al. (2016) proposed a new theory based on the literature that might aid to attain that goal, the theory of

✉ André De Zutter
andre.dezutter@maastrichtuniversity.nl

Robert Horselenberg
robert.horselenberg@maastrichtuniversity.nl

Peter J. van Koppen
p.j.van.koppen@vu.nl

¹ Department of Criminal Law and Criminology, Faculty of Law, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

² Department of Criminal Law and Criminology, Faculty of Law, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

fabricated rape. The theory of fabricated rape predicts that a false allegation will bolster detectable differences in comparison with true allegations of rape based on the principle that a false complainant of rape has not been raped and has to fabricate a story while the story of a true victim is based on recollections of the event. That has three consequences. First of all, false complainants are lying and will behave as liars do. If the stories liars present are not the same as the stories truth-tellers present, then the stories of false complainants will diverge from the stories of true victims. Second, false complainants will construct a story based on their own sexual experiences. If the sexual experiences of false complainants do not resemble rape, the fabricated story of rape will bolster detectable differences compared to a true story of rape. Third, false complainants will construct a story based on their beliefs concerning rape. If the beliefs false complainants hold about rape are invalid, then an invalid story of rape will be constructed.

The idea that the stories of truth-tellers differ from the stories of liars stems from Undeutsch (1982). Undeutsch (1982) hypothesised that true statements of children in sexual abuse cases have characteristics that distinguish them from statements of children in which invented events or distortions of experienced events are described. Since then, many studies were conducted to discover the differences between truth-tellers and liars. DePaulo et al. (2003) conducted a meta-analysis to find valid and reliable cues to deception. These researchers found some cues of deception but only small effect sizes. Most differences might therefore be difficult to detect. A common strategy of liars, however, is to keep the story simple and without details (Masip and Herrero 2013; Strömwall et al. 2006). Since false complainants are liars, false complainants will probably adopt the same strategy and construct a concise general story.

McDowell and Hibler (1993) argued that a fabricated story of rape is less detailed than an authentic account that is consistent with the theory of fabricated rape. A false complainant, for instance, does not give an estimate of the duration of the rape or describes how she and the rapist became undressed. On a related note, Woodhams and Grant (2004) studied the speech of offenders as reported in the stories of false complainants and true victims. These researchers studied 22 allegations that were maintained as true, the stories of true victims in their study and 22 allegations that were withdrawn as false, the stories of false complainants in their study. It was found that true stories contained significantly more utterances by the offender than false stories. Thus, false complainants behaved as liars do and presented a concise general story.

False complainants will construct a story that is based on their own sexual experiences. Consensual sexual experiences do not resemble rape. In a field study by Philips (2000), who interviewed 30 women between 18 and 22 years of age to study their sexual life, the participants described a wide array

of sexual experiences, desires, and fantasies. Some women described violent sexual experiences or sexual experiences in which a rape scenario was played out as a sexual fantasy between consenting lovers. Violent sexual experiences are not the same as rape experiences, because the women consented to the violent sexual encounter. Consensual violent sexual experiences are also different from sexual experiences in the context of rape because most rapists do not use violence (Canter 2000, 2004; Canter et al. 2003; Knight 1999; Kocsis et al. 2002; Prentky and Knight 1991). Researchers who studied rapists' or victims' accounts of rape found that if violence is used, the violence is instrumental in almost all cases (Canter et al. 2003; Knight 1999; Kocsis et al. 2002; Prentky and Knight 1991). Instrumental violence is goal-oriented violence; violence used to reach a goal and when the goal is reached, the violence stops. Instrumental violence in the context of rape is the violence that is needed to control the victim. Excessive levels of violence are rare during rapes (Canter et al. 2003; Knight 1999; Kocsis et al. 2002; Prentky and Knight 1991). McDowell and Hibler (1993) suggested that during a rape, the victim is more concerned with survival and submits to the attack with little resistance, while in false allegations, the levels of violence and resistance described by the complainants are much higher.

Sexual experiences in which a rape scenario, as a sexual fantasy, is played out between consenting partners, do not resemble true rape experiences. The rape script is, after all, based on the same invalid beliefs false complainants hold about rape. Moreover, the women who engage in such fantasy gave their consent and can always retract their consent and stop if they feel uncomfortable. In the study by (Philips 2000), a few women told that they were raped as a child, teenager or in early adulthood. The rape experiences the women described did not resemble the consensual sexual experiences described by the same women. Thus, in the case when a woman constructs a rape story based on her own consensual experiences, the story will not resemble a true story of rape.

A woman who is not raped will presumably associate rape with not wanting. False complainants will therefore believe that rape resembles unwanted sex and will use their own unwanted sexual experiences to fabricate their rape story. As part of the sexual experiences of sexually active people, unwanted sex is common (Bay-Cheng and Eliseo-Arras 2008; Erickson and Rapkin 1991; O'Sullivan and Allgeier 1998; Philips 2000). It means that most sexually active persons have experienced unwanted sex at least once.

Unwanted but consensual sexual experiences resemble wanted sexual experiences, but are restricted in the variety of sexual acts that are performed. If false complainants fabricate the rape story based on their own unwanted sexual experiences, the variety of sexual acts that were performed will be limited in fabricated stories of rape. McDowell and Hibler (1993) argued that in a false allegation the sexual acts are more

basic, usually just vaginal intercourse. Parker and Brown (2000) found a wider array of sexual acts in the stories of true victims of rape. For instance, 13 of the 16 stories of true victims described anal intercourse and the insertion of objects. A description of sexual acts other than vaginal intercourse was only reported in 6 of the 17 stories deemed to be false or fabricated stories of rape. Marshall and Alison (2006) compared the stories of false complainants with stories of true victims. These researchers asked women to write down a fabricated story of rape. For the stories of true victims of rape, a police database was used. Consistent with the theory of fabricated rape, Marshall and Alison (2006) found that a significant difference between the stories of false complainants and true victims was the variety of sexual acts and sexual positions that were described in the stories. In a fabricated story of rape, usually only one sexual act and position was described, mainly frontal vaginal penetration. True stories of rape included other sexual acts such as fellatio and cunnilingus.

In cases of allegations of rape differences between true and false allegations become apparent because only rape victims can rely on recollections of the event and will report details that are not commonly associated with rape but in reality frequently are part of the offence such as a wide variety of sexual acts that were performed. News media form the beliefs held by lay-persons concerning a phenomenon (Greer 2003). Portrayals of rape in the media are consistently atypical and general. As a consequence, a prototype of rape arises in the public domain and in the mind of lay people that does not correspond with the reality of rape, at least in most cases of rape. News agencies influence people's beliefs and perceptions of rape in an invalid way, thereby reinforcing misconceptions about rape and cultivating invalid rape stereotypes (Ardevini-Brooker and Caringella-MacDonald 2002).

Representations of rape in news media lack details and are biased. The sensational and unusual types of rape cases are the cases that are most frequently covered in the news media (Greer 2003; Soothill and Walby 1991). For instance, pseudo-intimate behaviour is never described in news media (Greer 2003; Soothill and Walby 1991), while a lot of rapists exhibit pseudo-intimate behaviour (Canter et al. 2003). Pseudo-intimate behaviour is behaviour that is commonly exhibited in the context of consensual sex and mimics a caring relationship. Kissing is, for example, pseudo-intimate behaviour. Lay people do not believe that pseudo-intimate behaviour is exhibited by rapists. Such is demonstrated in a study by Ellison and Munro (2009) on mock jury deliberations after a rape trial. These researchers varied several parameters and had actors and barristers act out nine different rape trials. It was found that jurors believed that an allegation of rape was false in case the rape was preceded by kissing. The jurors especially deemed the allegation to be false in case the kissing was requested by the rapist and consented to by the victim. Jurors

believed that rapists would not be asking for a kiss if they intended to rape someone. Thus, false complainants will file an allegation that bolsters detectable differences with a true allegation of rape because false complainants rely on invalid beliefs of how such an event would be enacted.

To test the validity of the theory of fabricated rape, De Zutter et al. (2016) studied 65 allegations of rape. A quasi-experimental design was used. True allegations were obtained from files of convicted rapists. The convicted rapists had confessed to the crime, and the stories of rape were corroborated by evidence. To obtain false allegations, participants were asked to fabricate an allegation. These researchers constructed a list of 187 variables that was in part theory-driven, constructed based on the theory of fabricated rape, and in part data-driven, retrieved from the 65 allegations of rape that were studied. The variables were coded dichotomously: 0 for absent, 1 for present. The results were consistent with the theory of fabricated rape. First, false complainants behaved as liars and presented a concise and less detailed story than true victims did. Second, false complainants based their fabricated story of rape on their own sexual experiences. Therefore, false allegations of rape included a restricted array of sexual activities and sexual positions compared to true allegations of rape. For example, anal intercourse was not included in the stories of false complainants while 39 % of the stories of true victims included anal intercourse. Third, false complainants constructed, as expected, their story based on invalid beliefs about rape. Fabricated rapists, for instance, did not exhibit pseudo-intimate behaviour. True rapists caressed, kissed and complimented the victim. Foreplay was included in 70 % of the stories of true victims while 14 % of false complainants included foreplay in their story. True rapists asked personal questions, tried to discover the identity and address of the victim and stayed longer with the victim than necessary. In 30 % of the stories of true victims, the rapist apologised afterwards, whereas no fabricated rapist in the stories of false complainants did. In 40 % of the stories of true victims, the rapist was friendly afterwards and in 53 %, the rapist reassured the victim.

A Tool to Discriminate Between True and False Allegations of Rape

Consistent with the theory of fabricated rape, researchers reported differences between true and false allegations of rape (De Zutter et al. 2016; Hunt and Bull 2012; Kanin 1994; Marshall and Alison 2006; McDowell and Hibler 1993; Parker and Brown 2000; Rassin and Van der Sleen 2005). Thus, it might be possible to develop a tool that uses this knowledge of known variations between true and fabricated allegations of sexual assault to predict the true nature of an allegation (De Zutter et al. 2016; Hunt and Bull 2012; Parker and Brown 2000).

Parker and Brown (2000) used statement validity analyses (SVA) to classify rape allegations into true and false allegations of rape. SVA is based on the aforementioned Undeutsch hypothesis (Undeutsch 1982). SVA contains criteria-based content analysis (CBCA) and a validity checklist. SVA is developed to assess the veracity of child testimony and is not without controversy (Rassin 2001). Parker and Brown (2000) analysed 43 allegations, 16 were classified as true, 15 as unsubstantiated and 12 as false. These researchers found that using the SVA method, 100 % of true allegations and 91.7 % of false allegations could be correctly classified. In the Parker and Brown (2000) study, predictive validity was compromised, since the researchers who coded the SVA criteria were not blind for the true nature of an allegation. A confirmation bias, a tendency to confirm the hypothesis unconsciously, cannot be excluded and could in part explain the almost perfect hit rate, correct classification rate. A similar conclusion was made by Vrij (2005) in his qualitative review of studies of SVA. He concluded that due to methodological flaws, the results of Parker and Brown (2000) should be disregarded. Since, to our knowledge, this is the only study where CBCA was used to differentiate between true and false allegations, we cannot compare the results with other studies that used the same stimulus material.

In other studies in which CBCA was used to differentiate between true and false statements, however, error rates were much higher, between 27.40 and 40 % (Sporer 1997; Vrij et al. 2000). In a qualitative review of 37 studies on the validity or reliability of SVA, Vrij (2005) reported overall accuracy rates ranging from 55 to 90 %. A conclusion of Vrij (2005) was that because of the high error rate, SVA was not accurate enough to be presented as scientific evidence in the legal arena. At best, according to Vrij (2005), it could be used in the beginning of a police investigation to form a rough indication whether it concerned a true or a false account.

Recently, another way of differentiating true from false allegations of rape was introduced by Hunt and Bull (2012). They compared 160 true allegations with 80 false allegations and scored the allegations on absence or presence of certain behaviours. These researchers constructed a list of 62 variables (e.g. theft, cunnilingus, victim masturbating offender, victim injured; see Hunt and Bull 2012 for an overview of the list of variables). Subsequently, these researchers coded the variables as either present or absent. These researchers found that the coding of 44 out of the in total 62 variables differed significantly between true and false allegations. A true allegation of rape included other offence behaviours such as theft. In their sample, an allegation involving theft was 6.2 times more likely to be a true allegation than an allegation not involving theft. A true allegation of rape also included a wide variety of sexual acts and a lot of verbal victim-offender interaction. A true allegation was also marked by pseudo-intimate behaviour such as kissing, cuddling, fondling and

cunnilingus. These researchers used the differences to build a model. By means of backward stepwise logistic regression analysis, a regression equation was constructed. The regression equation consisted of five predictors and a constant. To validate their model, 12 allegations, eight true allegations and four false allegations, of rape were blindly categorised as either a true or a false allegation. The overall hit rate was 83 %; ten allegations of the 12 allegations of rape were correctly classified. Six out of eight (75 %) true allegations were correctly classified, and all four (100 %) false allegations were correctly classified.

The same method was employed by the previously described study by De Zutter et al. (2016). These researchers identified characteristics that were typical for true allegations of rape as well as characteristics that were typical for false allegations of rape. Thus, it might be possible to build a tool to discriminate between true and false allegations of rape based on the differences found by these researchers.

Ground Truth

In studies on truthfulness of allegations of rape is important to establish ground truth. Ground truth is a term used to define what actually happened (Horowitz 2009). It means that allegations classified as false are, in fact, false allegations of rape, while allegations classified as true are, in fact, true allegations of rape. In that sense, false negatives, true allegations in the sample of false allegations, as well as false positives, false allegations in the sample of true allegations should be avoided as much as possible.

Researchers use different concepts to represent ground truth in their studies. For example, some researchers use the no-crime definition as ground truth (Rumney 2006). It means that allegations are deemed true unless they receive the no-crime label by the police. That classification, though, relies on police decision making, which therefore might not fulfil the concept of ground truth. Police officers sometimes use the no-crime label for marital rape or in case of various evidence problems regardless of the true nature of the allegation (Gregory and Lees 1996). Another approach to represent ground truth is to use judicial outcome as a substitute (Rassin and Van der Sleen 2005). That is not a correct representation of ground truth either, since sometimes guilty people are exonerated and innocent people get convicted, as in the case of Gary Dotson (Gross et al. 2005). A final approach is to take a retraction by the claimant as proof of a false allegation (Kanin 1994). Sometimes claimants, though, retract their allegation due to police pressure, e.g. when they are not believed or told that there is no possibility to obtain a conviction (Haket 2007). In conclusion, it is not easy to obtain ground truth. Therefore, stringent criteria should be used in studies on allegations of rape to avoid false negatives as well as false

positives. In the current study, we took several precautions which will be explained in the method section.

Current Research

We tested whether it is possible to build a model to distinguish true and false allegations of rape based on the theory of fabricated rape. The theory of fabricated rape predicts that a false allegation will entail detectable differences in comparison with true allegations of rape based on the principle that a false complainant of rape has not been raped and has to fabricate a story while the story of a true victim is based on recollections of the event. The current research tries to replicate the results of the study conducted by De Zutter et al. (2016) and will test whether it is possible to build a model to distinguish true and false allegations of rape based on the aforementioned characteristics.

Method

Definition of True and False Allegations of Rape

The concepts of the current research are true allegations of rape and false allegations of rape. Definitions of rape vary widely (Gannon et al. 2008). Core concepts, however, are sexual intercourse and lack of consent (Shields and Shields 1983). Thus, in the current study, a true allegation of rape is defined as the actual unlawful compelling of a person through physical force or duress to have sexual intercourse.

To define false allegations of rape is not as straightforward as it may seem. Criminal justice professionals tend to count an allegation as false when the account of the rape is not entirely true, the complainant has lied on some aspects, or made mistakes (Saunders 2012). It is not ideal to define false allegations as such, since allegations of victims who are in fact raped but did lie about some aspects of their story are also considered to be false allegations. A victim may, for example, conceal or distort information in an attempt to conform to cultural standards (Bletzer and Koss 2004).

A logical alternative definition seems to be that an allegation of rape is false when the complainant has not been raped. Sometimes people suffering from sexual hallucinations think they have been raped while in fact they had not been raped (Balasubramaniam and Park 2003). In other cases, complainants believe they were raped while asleep or intoxicated, but changed their opinion in light of the subsequent investigation and retracted the allegation (Kelly et al. 2005). These complainants, however, do not have the intention to mislead police officers. As a consequence, it seems better to include malicious intent in the definition. Some researchers consider malicious intent as a prerequisite to consider an allegation to be

false (Greer 1999; Gregory and Lees 1996; Kanin 1994; Rassin and Van der Sleen 2005; Rumney 2006).

A false allegation of rape in the current study is therefore defined as intentionally reporting to have been raped while no rape has occurred. To constitute a false allegation of rape, no sexual intercourse has taken place or the sexual intercourse was consensual and not the consequence of physical force or duress. In addition, the complainant is aware of the fact that she is filing a false allegation. In other words, the false allegation is not a consequence of a confused state of mind of the complainant.

Sources of Cases and Criteria for Ground Truth

True and false allegations of rape were studied. The study was limited to male perpetrators and female victims of rape. Allegations where the complainant was male or under the age of 14 were excluded from the sample. The male allegations were excluded because we think the story of a male rape varies too much on essentials of the story with a female rape story to study them together. Allegations of complainants under the age of 14 were excluded because in the Netherlands, people under the age of 14 are lawfully unable to consent to any sexual activity. There was no restricted time limit posed on the elapsed time between the occurrence of the rape and the reporting of the rape, the only limitation was that the complainant was not under the age of 14 at the time of the event. All allegations of rape, true as well as false allegations, were provided by the National Unit of the Dutch National Police (NU). Permission to study the files and to gather data was granted by the Minister of Justice of The Netherlands. The files were retrieved from the Violent Crime Linkage System (ViCLAS). ViCLAS is a software program developed by the Royal Canadian Mounted Police (Aldred 2007). The software program creates a database. The database is used to analyse violent crimes to detect patterns and catch serial offenders. In the Netherlands, the goal is to enter all murders and sexual offences in ViCLAS to create a national database.

Law enforcement agencies across the Netherlands send criminal files of murders and sexual offences to the NU to enter in ViCLAS. Since 2002, all entries are made by trained NU officers on the basis of a structured questionnaire. We noticed, however, marked differences between the files and the entries in ViCLAS. As a consequence, we decided to study the original files. ViCLAS was only used to identify potentially relevant files. We complied to all conditions in the permission for the study. Thus, no demographic data were collected, all raw data were anonymised to protect the identity and secure confidentiality of all parties involved. Files were only identifiable through a number, and all files were studied and coded at the headquarters of the NU in Zoetermeer, the Netherlands.

According to the formal NU definition, a case is categorized as false if the investigation showed that the case was in fact false and the criteria of our definition are fulfilled. While studying the files, we discovered that NU officers were using a different definition to label an allegation as false. The NU officers deem an allegation to be false if the complainant's story changes in light of the investigation even if the complainant has in fact been raped.¹ Such cases, cases in which the complainant insisted that she was raped although the police discovered that the complainant had been lying about some aspects of the case, were excluded from the current study. To establish ground truth, all case files were studied in full. If the complete file was not available and could not be obtained the case was excluded from the study.

For the current study, stringent criteria for a true and false allegation were set. A case was categorized as false if the complainant retracted the allegation and told the police that the allegation was in fact false and no rape whatsoever had occurred. But that was not enough. We also wanted that the alternative no-rape-scenario was supported by corroborating and conclusive evidence.² The stringent criteria were used to avoid to a reasonable extent the possibility that false allegations of rape would pollute the sample of true allegations.

For our sample of true allegations, we used criminal files of convicted suspects. All criminal files that did not contain a full confession of the accused were excluded. To eliminate the possibility of false confessions, a confession alone was not sufficient to be included in the sample. Only case files were included that also contained at least one of the following independent pieces of evidence: a DNA match, identification by the victim in a valid line-up, caught in the act, the confession contained strong guilty knowledge, or possessions of the victim were retrieved from the defendant.

¹ One complainant told the police that she was abducted and raped in a basement. The investigation revealed that she was not abducted. Instead she met the perpetrator on the Internet, and MSN conversations between them revealed that she had agreed to meet him voluntarily. When the girl was confronted with the MSN conversations, she admitted that she was not abducted but she insisted that the boy had raped her. She told the police that she had invented the abduction to make her rape allegation more believable. She was afraid that when she told the truth, that she voluntarily entered the basement and engaged in consensual kissing before she was raped, the police would think that the sex was consensual.

² One complainant said that she was raped in an alley. Forensic examination revealed vaginal seizures and the presence of semen. A male DNA profile was obtained from the semen. The DNA profile did not match with any of the profiles in the Dutch DNA database. The alley, however, was equipped with surveillance cameras. The police examined the footage and did not see the complainant. When the complainant was confronted with the footage, she retracted her allegation. She admitted that she had not been raped by a stranger but had had consensual sex with her boyfriend at another location. The DNA profile did match with the DNA sample, provided by her boyfriend. She stated that she had invented the rape out of fear for her father. She told the police that her father was a racist and did not approve of her relationship with her boyfriend.

Sampling Method

All allegations of rape were identified by use of queries in ViCLAS. The queries were based on the definitions of true and false allegations given above. All allegations that were classified as false in ViCLAS from April 1997 to August 2011 were studied, 91 cases in total. Twenty cases that were classified as false allegations of rape were incomplete. It means that it was not possible to establish ground truth on the basis of the available information. Additional information for the files was sought in BlueView, a search engine to retrieve data of all police districts and Royal Military Police Corps. When there was no additional information available in BlueView, the local police district was contacted with a request for additional information. Additional information was given for nine files. Seven of these allegations met the criteria of the definition of a false allegation of rape that was used in the current study and two did not. Twenty-two additional files were excluded from the sample, since the allegations did not meet the criteria of the definition of a false allegation of rape used in the current study. One false allegation misclassified as a true allegation was added to the sample of false allegations.³ The total sample of false allegations of rape that was finally coded consisted of 57 files.

To obtain a control group of true allegations, another query was constructed. The query on true allegations of rape yielded 258 results. All 258 allegations received a number ranging from one to 258. A random sample was drawn following the random numbers sequence published by Moore et al. (2012). In total, 114 files were studied. Seventy-two files met the inclusion criteria for a true allegation of the current study. One false allegation of rape was misclassified as a true allegation in ViCLAS.³ The allegation was added to the sample of false allegations. The result of the selection procedure was 57 cases of false allegations and 72 cases of true allegations.

Materials and Coding

A list of 187 variables was used. The list was developed in a study by De Zutter et al. (2016). All variables were coded dichotomously: 0 for absent, 1 for present. All variables held very straightforward descriptions of behaviours, so coding posed little problems (e.g. 'biting victim', 'stealing something',

³ One complainant claimed to have been raped in a car. The suspect claimed that he could prove that the sex was consensual. He had received a police fine for sex in public. The police officer testified in court. He testified that he had indeed fined the man for having sex in public. He had seen the girl while she was having sex and interrupted the sexual activity. The girl did not seem to be in distress and did not alert the police officer that she was being raped. After the man was fined, he drove off with the girl in his car. The accused testified that he drove off to another place to continue the sexual encounter. The girl had fabricated the rape out of shame for the consensual sexual encounter with a stranger, she said.

'bushes', 'condom use offender' and 'fellatio'). All allegations were coded blind by one of two independent evaluators. The independent evaluators did not know whether an allegation was true or false during the coding phase. Both independent evaluators were trained legal psychologists. The first author was one of the independent evaluators. No training on the coding of the variables was provided for the current study. Fifteen allegations were coded by both independent evaluators to assess reliability.

Results

Reliability

Cohen's measure of agreement, kappa, was calculated for all 187 variables on the coding schemes of 15 allegations that were coded by both coders (Cohen 1960). Only incidental differences between the evaluators were found. Cohen's kappa's ranged from Cohen's kappa = .40 to Cohen's kappa = 1.00. Cohen's kappa could not be calculated for the coding of 78 variables, because the variables were coded by both independent evaluators as either absent ($n = 76$) or present ($n = 2$) in all 15 allegations. The coding of 86 variables was in perfect agreement, Cohen's kappa = 1.00. The measure of agreement on the coding of one variable was .40. According to Landis and Koch (1977), the kappa's calculated for the current study range from fair ($k = .40$) to almost perfect ($k = .80$ – 1.00). The measure of agreement on the coding of 21 variables was between .61 and .80 which is considered substantial by Landis and Koch (1977). The measure of agreement on the coding of one variable was between .41 and .60 which is considered moderate by Landis and Koch (1977). The scale by Landis and Koch (1977) is the scale that is commonly cited when interpreting Cohen's kappa's (Viera and Garrett 2005). Since no calculated kappa was unacceptable, no corrective measures were taken (Sim and Wright 2005; Viera and Garrett 2005).

Model Building

The theoretical underpinning of our rationale for building our model is the theory of bounded rationality by Gigerenzer (2002). He argued that people have limited resources to reach decisions. People in everyday situations simply are unable to perform the complex computational processes that are required to reach decisions in complex situations with multiple regression or Bayesian networks. According to Gigerenzer (2002), people rely on heuristics instead. A heuristic is a simple rule that ignores information and leads to quick decisions based on only a small piece of information. The decisions based on heuristics can be more accurate than decisions based on complex statistical calculations, known as the less is more effect (Gigerenzer and Gaissmaier 2011). Especially police

officers are confronted with time constraints and limited resources. Therefore, we believe that a less is more approach might be practical, reliable and valid in the context of police investigations.

All 129 allegations, 72 true allegations and 57 false allegations, received a number. The random sample of 27 true allegations and 27 false allegations were drawn from the total sample following the random numbers sequence published by Moore et al. (2012). The random sample of 54 allegations was used to replicate the findings of the study of De Zutter et al. (2016) and to build the model. In the study by De Zutter et al. (2016), ecological validity might have been compromised, since false complainants did not really file a false allegation at a police station. But ground truth was high in the study because all false complainants were certainly not raped. In the current study, ecological validity was high, but ground truth, internal validity, might have been compromised. Therefore, the validity of the model increases if it is built on both independent samples. Building the prediction model on two independent samples also increases the validity of the model and reliability of the results upon which it was built. Therefore, the current results were compared to the results of De Zutter et al. (2016; see Table 1 for an overview). The coding of the variables in true allegations was compared to the coding of the variables in false allegations. All variables that were coded absent in either all true allegations or all false allegations in both independent samples were selected.

Only variables that were coded absent in either all true allegations or all false allegations in both studies were used to build the decision matrix, because we wanted to construct a flow chart that is based on heuristics and returns a decision every step of the process (see Table 1). Therefore, it was imperative that the heuristic on which the step is built is expected to be a perfect predictor in the population of either true or false allegations. Since a variable was coded absent twice in two independent samples, we assumed that it could be a perfect predictor. The decision matrix is a flow chart built up by clear closed questions, heuristics (Gigerenzer and Gaissmaier 2011). The answer to the closed question yields a decision on the allegation and determines the next step, either a decision on the nature of the allegation and an exit of the flowchart or the next question of the flow chart. If no decision is reached on the nature of the allegation after the last question in the flow chart, the allegation receives the label 'Undecided'.

Our final decision matrix consisted of 12 steps. Every step of the decision matrix yielded a decision on the nature of the allegation, true or false. Using the decision matrix, the first author classified 24 allegations, five false allegations and 19 true allegations (see Table 2). It must be stated, however, that it is off course a meaningless and circular result, because the same sample was used to build the decision matrix. Only the first step yielded the decision that the allegation was false which explains why more true allegations than false

Table 1 Proportions coded present of the variables of which the coding of the variables differed significantly between true and false allegations of rape in both independent samples the current study and the study conducted by De Zutter et al. (2016)

Variable	Current study		De Zutter et al. (2016)	
	True %	False %	True %	False %
Condom use offender	0	19	0	15
Kissing afterwards	11	0	43	0
Complimenting afterwards	4	0	17	0
Apologising afterwards	26	0	30	0
Offender asking if victim loves him	4	0	7	0
Offender sexually commenting on himself	7	0	53	0
Spanking	4	0	3	0
Drug use	7	0	40	0
Demanding goods or money	30	0	30	0
Trying to steal something	15	0	3	0
Longer than 30 min	7	0	37	0
Longer than 1 h	0	0	23	0
Victim participating	74	15	83	3
Other evidence	44	7	80	6
Fellatio	70	11	53	3
Victim telling spontaneous events post-rape	93	67	100	17
Instructing to stop	59	26	90	17
Threatening prior to the incident	52	26	63	3
Vaginal penetration behind	41	11	60	6
Friendly afterwards	44	0	40	3
Touching/fondling crotch, without penetration	33	4	67	17
Longer than necessary	44	4	37	0
Holding victim by neck	30	0	53	6
Self-initiative report rape	70	22	73	46
Personal questions	52	11	33	0
Offender carrying a weapon	59	26	40	3
Offender masturbating himself	19	0	57	6
Inconsistencies narrative	15	52	0	31
Sexually commenting on victim	33	4	67	0
Offender giving personal info	44	11	47	14
Longer than 15 min	11	0	57	3
Offender reassuring victim	22	4	53	6
Discovering address	41	11	37	3
Discovering identity	30	7	40	0
Victim masturbating offender	26	0	40	6
Stealing something	30	0	33	6
Stealing personal item	30	0	27	6
Anal penetration	15	4	37	0
Relevant info on multimedia	19	44	10	31
Stealing identifiable item	26	0	20	3
Complimenting before/during sex	15	0	27	6
More than 3 days between rape and report	0	6	10	31

Proportion of coded present = 1; Variables that were coded absent in either all true allegations or all false allegations in both samples are in boldface and were used to build the decision matrix

^aStatistics, *p* values, Odd's ratio's and 95 % confidence intervals of the variables used in the prediction equation are displayed in Table 3

Table 2 Decision matrix and prediction outcome on the building sample and on the blind-coded test sample

Decision matrix			Prediction	
			Building Model ^a	Testing Model
			<i>n</i> = 54	<i>n</i> = 75
Question/heuristic	Coding	Decision	<i>n</i>	<i>n</i>
Did the rapist use a condom?	I do not know	False	<i>n</i> = 5	<i>n</i> = 9 ^c
	Yes or no	Next question	<i>n</i> = 49	<i>n</i> = 66
Did the rapist kiss the victim afterwards?	Yes	True	<i>n</i> = 3	<i>n</i> = 4 ^d
	No	Next question	<i>n</i> = 46	<i>n</i> = 62
Did the rapist compliment the victim afterwards?	Yes	True	<i>n</i> = 1	<i>n</i> = 1
	No	Next question	<i>n</i> = 45	<i>n</i> = 61
Did the rapist apologize afterwards?	Yes	True	<i>n</i> = 6	<i>n</i> = 3
	No	Next question	<i>n</i> = 39	<i>n</i> = 58
Did the offender ask if the victim loved him?	Yes	True	<i>n</i> = 0	<i>n</i> = 0
	No	Next question	<i>n</i> = 39	<i>n</i> = 58
Did the rapist give sexual comments on himself?	Yes	True	<i>n</i> = 1	<i>n</i> = 4
	No	Next question	<i>n</i> = 38	<i>n</i> = 54
Did the rapist spank the victim?	Yes	True	<i>n</i> = 1	<i>n</i> = 1
	No	Next question	<i>n</i> = 37	<i>n</i> = 53
Did the rapist use drugs?	Yes	True	<i>n</i> = 1	<i>n</i> = 6
	No	Next question	<i>n</i> = 36	<i>n</i> = 47
Did the rapist demand goods or money?	Yes	True	<i>n</i> = 4	<i>n</i> = 4
	No	Next question	<i>n</i> = 32	<i>n</i> = 43
Did the rapist try to steal something?	Yes	True	<i>n</i> = 0	<i>n</i> = 1
	No	Next question	<i>n</i> = 32	<i>n</i> = 42
Was the duration of the assault longer than one hour?	Yes	True	<i>n</i> = 0	<i>n</i> = 0
	No	Next question	<i>n</i> = 32	<i>n</i> = 42
Was the duration of the assault longer than 30 min?	Yes	True	<i>n</i> = 2	<i>n</i> = 0
	No	Undecided	<i>n</i> = 30	<i>n</i> = 42

n all allegations, true and false, *n* absolute number of allegations

^aPrediction outcome of the sample on which the model, decision matrix, was partially build

^bBlind-coded prediction outcome of the random sample that was drawn to test the model, decision matrix

^c*n* = 1 misclassified as a false allegation while in fact it was a true allegation, a false-negative

^d*n* = 1 misclassified as a true allegation while in fact it was a false allegation, a false-positive

allegations could be classified using the decision matrix (see Table 2).

Chi-square tests were used to detect differences in coding of the remaining variables. The results of the current study were, again, compared to the results obtained in the study conducted by De Zutter et al. (2016; see Table 3 for an overview). The variables that were not used in the decision matrix were used to form a prediction equation that was again based on the principles formulated by Gigerenzer (2002). For each behaviour that was present in true allegations 1 was added, while for each behaviour that was typically present in false allegations 1 was subtracted. The higher the total score of an allegation, the more likely it becomes that the allegation is true. A total score lower than three was in our sample

indicative of a false allegation, while a total score higher than seven was in our sample indicative of a true allegation (see Table 4). Because true as well as false allegations received total scores between three and seven, three false allegations out of 27 (11.11 %) and five true allegations out of 27 (18.51 %) allegations of rape could not be classified. Based on the decision matrix and the following prediction equation, the first author was able to accurately classify 46 out of 54 (85.19 %) allegations. With a cut-off of three all the allegations could be classified as either a true or a false allegation leaving no allegations undecided (see Table 4). The consequence, however, would be that either the false-positive rate or the false-negative rate would rise considerably. The false-positive rate, allegations classified as true while the allegations

Table 3 Statistics of the variables used in the prediction equation

Variable	Current study				De Zutter et al. (2016)			
	<i>p</i>	<i>OR</i>	95 % CI		<i>p</i>	<i>OR</i>	95 % CI	
			<i>LL</i>	<i>UL</i>			<i>LL</i>	<i>UL</i>
Victim participating	<.0001	0.06	0.02	0.24	<.0001	0.00	0.00	0.05
Other evidence	.002	0.10	0.51	0.02	.049	0.43	0.32	0.57
Fellatio	.001	0.05	0.01	0.23	<.0001	0.03	0.00	0.21
Victim telling spontaneous events post-rape	.010	0.21	0.06	0.72	.033	0.26	0.07	0.94
Instructing to stop	.013	0.02	0.08	0.76	<.0001	0.02	0.01	0.10
Threatening prior to the incident	.049	0.33	0.10	1.02	<.0001	0.01	0.00	0.14
Vaginal penetration behind	.013	0.18	0.04	0.76	<.0001	0.04	0.01	0.20
Friendly afterwards	<.0001	0.36	0.24	0.54	<.0001	0.04	0.01	0.37
Touching/fondling crotch, without penetration	.005	0.08	0.01	0.66	<.0001	0.10	0.03	0.33
Longer than necessary	.001	0.05	0.01	0.41	<.0001	0.35	0.25	0.51
Holding victim by neck	.002	0.41	0.29	0.58	<.0001	0.05	0.01	0.26
Self-initiative report rape	<.0001	0.12	0.04	0.35	.022	.031	0.11	0.87
Personal questions	.001	0.12	0.03	0.48	<.0001	0.36	0.26	0.52
Offender carrying a weapon	.013	0.02	0.08	0.76	<.0001	0.05	0.01	0.37
Offender masturbating himself	.019	0.45	0.33	0.61	<.0001	0.05	0.01	0.23
Inconsistencies narrative	.004	6.25	1.70	25.00	.003	2.22	1.64	2.94
Sexually commenting on victim	.005	0.08	0.01	0.66	<.0001	0.22	0.13	0.38
Offender giving personal info	.006	0.16	0.04	0.65	.004	0.19	0.06	0.63
Longer than 15 min	.050	0.47	0.35	1.33	<.0001	0.02	0.00	0.19
Offender reassuring victim	.043	0.14	0.02	1.21	<.0001	0.05	0.01	0.26
Discovering address	.013	0.18	0.04	0.76	<.0001	0.05	0.01	0.42
Discovering identity	.036	0.19	0.04	1.00	<.0001	0.34	0.23	0.49
Victim masturbating offender	.005	0.43	0.31	0.59	<.0001	0.09	0.02	0.45
Stealing something	.002	0.41	0.29	0.58	.005	0.12	0.02	0.61
Stealing personal item	.002	0.41	0.29	0.58	.022	0.17	0.03	0.86
Anal penetration	.050	0.47	0.35	1.33	<.0001	0.35	0.25	0.51
Relevant info on multimedia	.040	3.57	1.03	12.50	.030	4.30	1.07	17.32
Stealing identifiable item	.002	0.43	0.31	0.59	.033	0.12	0.01	1.04
Complimenting before/during sex	.038	0.46	0.34	0.62	.020	0.17	0.03	0.86
More than 3 days between rape and report	.009	2.27	1.67	3.13	.026	4.50	1.12	18.16

CI confidence interval, *OR* odds ratio, *LL* lower level, *UL* upper level

p values, Odd’s ratio’s and 95 % confidence intervals, of significant differences in the coding of the variables in both independent samples, the current study and the study conducted by De Zutter et al. (2016)

should have been classified as false, would become three, or the false-negative rate, allegations classified as false while the allegations should have been classified as true, would become seven. Based on the results, we decided that allegations with a score between three and seven would be labelled ‘Undecided’.

Testing the Validity of the Decision Matrix

The second author, being blind to the nature of the allegations, coded the remaining 45 true allegations and 30 false allegations using the decision matrix. The second

author classified 32 allegations as either a true or a false allegation of rape based on the decision matrix, 43 allegations received the label ‘Undecided’ (see Table 2). Thirty of the 32 (93.75 %) allegations were correctly classified as either a true or a false allegation of rape based on the decision matrix. Twenty-three allegations were classified as true allegations. Nine allegations were classified as false allegations. The false-positive rate was one allegation out of 23 allegations (4.35 %). The false negative rate was one allegation out of nine allegations (11.11 %; see Table 2).

Table 4 Ground truth of the allegation and outcome of the prediction equation on the building sample

Number	Ground truth	Total score	Predicted	Cut-off 3
1	False	1.00	False	False
2	False	2.00	False	False
3	False	-2.00	False	False
4	False	0.00	False	False
5	False	-2.00	False	False
6	False	-2.00	False	False
7	False	1.00	False	False
8	False	2.00	False	False
9	False	0.00	False	False
10	False	3.00	Undecided	<i>True</i>
11	False	2.00	False	False
12	False	-3.00	False	False
13	False	-3.00	False	False
14	False	-3.00	False	False
15	False	1.00	False	False
16	False	0.00	False	False
17	False	5.00	Undecided	<i>True</i>
18	False	1.00	False	False
19	False	2.00	False	False
20	False	3.00	False	False
21	False	7.00	Undecided	<i>True</i>
22	True	8.00	True	True
23	True	11.00	True	True
24	True	5.00	Undecided	True
25	True	6.00	Undecided	True
26	True	7.00	Undecided	True
27	True	4.00	Undecided	True
28	True	16.00	True	True
29	True	3.00	Undecided	True
30	True	10.00	True	True

Correct predictions are in boldface. Incorrect predictions are in italics

Testing the Validity of the Equation

The second author blindly classified the remaining 43 allegations, the undecided allegations, based on the prediction equation (see Table 5). Of these, 24 allegations (55.81 %) could be classified as either a true or a false allegation of rape based on the outcome of the prediction equation. The accuracy rate was 19 out of 24 allegations (79.17 %). The false-positive rate was three out of 14 allegations (21.43 %). The false-negative rate was two allegations out of 11 allegations (18.18 %; see Table 5). Overall, 56 allegations were classified based on the decision matrix and the equation combined (74.67 %). Nineteen allegations received the undecided label (25.33 %; see Table 5). The error rate was seven allegations out of 75 (9.33 %; see Table 5).

Post hoc Analyses

Post hoc analyses were performed to see what had caused the errors in prediction and to test whether it was possible to increase the accuracy of the classification based on the decision matrix and equation. The false-positive rate of the decision matrix was one out of 23 allegations (see Table 2). The error was caused in the second step of the decision matrix ('Did the offender kiss the victim afterwards?'; see Table 2). Four allegations were classified as true allegations based on the second step of the decision matrix. If we would eliminate the second step, the false-positive rate would drop to zero but the number of allegations that could be classified as either a true or a false allegation of rape based on the decision matrix would also drop to 28 instead of 32.

The false-negative rate was one out of nine allegations. That particular error was caused by the first step of the decision matrix ('Did the offender use a condom?'; see Table 2). Nine allegations were classified as false allegations based on the first step of the decision matrix. If the first step is eliminated, the false-negative rate drops to zero but the number of allegations that could be classified as either a true or a false allegation of rape based on the decision matrix would also drop to 23 instead of 32. In case the first and the second step of the decision matrix would be eliminated, the error rate would drop to zero but the number of allegations that could be classified as either a true or a false allegation of rape based on the decision matrix would also drop to 19 instead of 32. It means that additional 13 allegations could not be classified as either a true or a false allegation of rape based on the decision matrix and that 56 out of 75 allegations would remain undecided.

Based on the building sample, we know that with a cut-off point of three all allegations could be classified, but the error rate would rise. In our prediction sample, a cut-off score of three would also increase the amount of allegations that could be classified. It would become possible to classify all 43 remaining allegations based on the new criterion (see Table 5). The error rate, however, would also increase. Twelve out of 43 (27.91 %) allegations would be misclassified based on the new criterion; the false-positive rate would become ten out of 21 allegations (47.62 %) and the false-negative rate would become two allegations out of 22 allegations (9.09 %; see Table 5). Overall, the accuracy rate would become 61 allegations out of 75 allegations (81.33 %). Post-hoc analyses revealed that it was not possible to improve the overall accuracy rate of the decision matrix and of the prediction equation.

Discussion

It is possible to predict the true nature of an allegation of rape. The second author, who was blind to the nature of the allegations, was successful in classifying the majority of allegations as either a false or a true allegation of rape. The overall error rate of

Table 5 Ground truth of the allegation and outcome of the prediction equation

Number	Ground truth	Total score	Predicted	Cut-off 3
1	False	1.00	False	False
2	False	1.00	False	False
3	False	-1.00	False	False
4	False	-2.00	False	False
5	False	1.00	False	False
6	False	1.00	False	False
7	False	2.00	False	False
8	<i>False</i>	<i>9.00</i>	<i>True</i>	<i>True</i>
9	<i>False</i>	<i>5.00</i>	<i>Undecided</i>	<i>True</i>
10	<i>False</i>	<i>4.00</i>	<i>Undecided</i>	<i>True</i>
11	<i>False</i>	<i>6.00</i>	<i>Undecided</i>	<i>True</i>
12	<i>False</i>	<i>6.00</i>	<i>Undecided</i>	<i>True</i>
13	<i>False</i>	<i>11.00</i>	<i>True</i>	<i>True</i>
14	False	1.00	False	False
15	False	-1.00	False	False
16	False	1.00	False	False
17	<i>False</i>	<i>5.00</i>	<i>Undecided</i>	<i>True</i>
18	False	-2.00	False	False
19	<i>False</i>	<i>5.00</i>	<i>Undecided</i>	<i>True</i>
20	<i>False</i>	<i>9.00</i>	<i>True</i>	<i>True</i>
21	<i>False</i>	<i>5.00</i>	<i>Undecided</i>	<i>True</i>
22	<i>True</i>	<i>4.00</i>	<i>Undecided</i>	True
23	True	10.00	True	True
24	<i>True</i>	<i>7.00</i>	<i>Undecided</i>	True
25	True	9.00	True	True
26	True	12.00	True	True
27	<i>True</i>	<i>5.00</i>	<i>Undecided</i>	True
28	True	10.00	True	True
29	<i>True</i>	<i>4.00</i>	<i>Undecided</i>	True
30	<i>True</i>	<i>5.00</i>	<i>Undecided</i>	True
31	<i>True</i>	<i>3.00</i>	<i>Undecided</i>	True
32	True	8.00	True	True
33	<i>True</i>	<i>2.00</i>	<i>False</i>	<i>False</i>
34	True	9.00	True	True
35	<i>True</i>	<i>4.00</i>	<i>Undecided</i>	True
36	<i>True</i>	<i>4.00</i>	<i>Undecided</i>	True
37	<i>True</i>	<i>7.00</i>	<i>Undecided</i>	True
38	True	15.00	True	True
39	True	9.00	True	True
40	<i>True</i>	<i>2.00</i>	<i>False</i>	<i>False</i>
41	<i>True</i>	<i>6.00</i>	<i>Undecided</i>	True
42	<i>True</i>	<i>3.00</i>	<i>Undecided</i>	True
43	<i>True</i>	<i>5.00</i>	<i>Undecided</i>	True

Correct predictions are in boldface. Incorrect predictions are in italics

the prediction is lower than 10 %. It seems that the decision matrix is a better tool to predict the nature of an allegation than

the prediction equation is. Based on the decision matrix alone, the second author accurately classified 94 % of 32 allegations as either a true or a false allegation. Based on the prediction equation alone, the second author accurately classified 79 % of 24 allegations as either a true or a false allegation. Since the decision matrix was the first step and the prediction equation was the second step and both were used on the same sample of 75 allegations of rape, it makes sense to interpret the decision matrix and the prediction equation together.

The use of a decision matrix and prediction equation to predict the nature of allegations of rape is to our knowledge new. Other researchers used either a regression equation (Hunt and Bull 2012) or CBCA (Parker and Brown 2000). In the current study, the decision matrix was a practical and simple tool to predict the nature of an allegation. Almost half of the total sample of allegations could be classified based on the decision matrix. Each step of the decision matrix consisted of a simple heuristic, a decision rule, upon which a decision concerning the nature of an allegation was based. In case a decision concerning the allegation could not be reached based on the heuristic, the next heuristic was taken into consideration until all 11 heuristics were taken into consideration. The prediction equation provided a sum score on which a decision concerning the nature of the allegation was reached.

With a sample size that was more than six times larger as the sample size of the study conducted by Hunt and Bull (2012), the decision matrix and prediction equation combined produced an error rate of 9.33 % in the current sample while the regression equation in the study by Hunt and Bull (2012) produced an error rate of 16.67 %. In the current study, decisions concerning the nature of an allegation were based on 46 predictors while the regression equation in the study by Hunt and Bull (2012) included five predictors. The current findings, however, do not contradict the findings of Hunt and Bull (2012). The same differences between true and false allegations were found in both studies. True allegations of rape include other offence behaviours such as theft, a wide variety of sexual acts and pseudo-intimate behaviour. In false allegations, the fabricated rape usually includes only one sexual act and position and has a short time span.

The current findings were in line with other studies and seem to confirm the theory of fabricated rape. False complainants have not been raped and do not fully grasp the phenomenology of rape. False complainants have to fabricate an event of rape and need to rely on other sources. Since false complainants do not know how such an event takes place, the complainants present a narrative that deviates significantly from a narrative of a true rape. False complainants commonly did not include theft in the false allegation, while theft has frequently been associated with rape (Canter et al. 2003; Kocsis et al. 2002). Theft was also one of the five predictors in the study by Hunt and Bull (2012).

Pseudo-intimacy was not reported by false complainants while true victims often reported being kissed, cuddled, fondled, reassured or complemented by the rapist (Canter et al. 2003;

Kocsis et al. 2002). That rapists often try to mimic a loving, caring, consensual relationship is counterintuitive and not commonly associated with a violent offence such as rape. False complainants who fabricate a rape therefore do not include pseudo-intimate behaviour. Rapists often regret having committed the offence which is not consistent with the image of an offender that does not respect the physical integrity of women. Because of the inconsistency between the offence and the regrets of the offender, false complainants do not report that the fabricated rapist apologized or was friendly afterwards. Being friendly afterwards could sometimes be described as pseudo-concern. In one of the true allegations, the rapist walked the victim home because, according to the rapist, the neighborhood was too dangerous for a woman to walk the streets alone. Such behaviour is not included in the mental representations and beliefs people have about rape. Consequently, if one has to fabricate a rape, pseudo-intimacy or pseudo-concern is not included.

The current study tried to overcome methodological flaws of other studies in the field of allegations of rape and the difference between true and false allegations (e.g. small sample sizes, poor definitions and lack of ecological and predictive validity; Hunt and Bull 2012; Lisak et al. 2010; McDowell and Hibler 1993; Norton and Grant 2008; Parker and Brown 2000; Rassin and Van der Sleen 2005; Rumney 2006). In research, a validity trade-off often is inevitable. If in a study ecological validity is maximized, another validity such as internal validity is often decreased (Brehm et al. 2005). In the current study, the ecological validity was at its maximum because all allegations were, regardless of their true or false nature, real allegations. Stringent criteria were used to firmly establish ground truth and to prevent that internal validity was compromised. It cannot, though, be excluded that false allegations corrupted the sample of true allegations of rape and vice versa. Another issue that might compromise the validity of the current study is the fact that no restricted time limit was posed on the elapsed time between the occurrence of the rape and the reporting of the rape. Thus, memory issues might, in part, explain the differences between true and false allegations of rape. Memory problems would, however, not affect false allegations of rape because, as stated before, false allegations are not based on recollections of the event. For true victims of rape, the effects might be minimised, since researchers have shown that stress leads to better memory consolidation (Schwabe et al. 2010).

The current study confirmed the hypothesis that false complainants, because they have invalid mental representations and false beliefs about rape, fabricate a rape story that does not resemble a true rape. The current results are consistent with the idea that there are salient and detectable differences between true and false allegations of rape. The decision matrix and the prediction equation that were built on the differences between true and false allegations of rape were able to predict the true nature of the majority of allegations in the current study with an error rate of 9.33 %. Post hoc analyses could not improve the accurate prediction rate of the current model

meaningfully. It makes sense in light of the current findings not to change the decision matrix and prediction equation when testing its validity in the field.

Misclassification of an allegation as either a true or a false allegation causes distress. Therefore, it seems valuable to classify allegations in a valid and reliable manner. Some researchers have tried to develop such a tool and made valuable contributions to the field of allegations of rape (Hunt and Bull 2012; Parker and Brown 2000). In both studies, predictive validity was compromised due to methodological flaws. Parker and Brown (2000) were not blind for the true nature of an allegation when they classified the allegations in their study and Hunt and Bull (2012) used a limited sample of 12 allegations, eight true allegations and four false allegations, of rape to test the predictive validity of their model.

We concealed the ground truth, the true nature of the allegations, in our study, and categorised the allegations blind. We also used a larger sample. We used a sample of 75 allegations, 45 true allegations and 30 false allegations, of rape to test the predictive validity of the decision matrix and prediction equation. The aim of our study was to develop a valid and reliable method to classify allegations of as either a true or a false allegation of rape based on the theory of bounded rationality by Gigerenzer (2002). We used heuristics to build our decision matrix and prediction equation, a procedure known as heuristic decision making (Gigerenzer and Gaissmaier 2011). The second author was able to classify the majority, 56 out of 75, of allegations based on the developed decision matrix and prediction equation with an accuracy rate of 91 %. The results indicate that the decision matrix and prediction equation might be a useful tool to aid police officers when investigating allegations of rape.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aldred K (2007) A memory for detail: ViCLAS database can provide missing link. *Gazette* 69(1):10
- Ardevini-Brooker J, Caringella-MacDonald S (2002) Media attributions of blame and sympathy in ten rape cases. *Justice Professional* 15:3–18. doi:10.1080/08884310212826
- Balasubramaniam B, Park GR (2003) Sexual hallucinations during and after sedation and anaesthesia. *Anaesthesia* 58:549–553. doi:10.1046/j.1365-2044.2003.03147.x
- Bay-Cheng LY, Eliseo-Arras RK (2008) The making of unwanted sex: gendered and neoliberal norms in college women's unwanted sexual experiences. *J Sex Res* 45:386–397. doi:10.1080/00224490802398381
- Bletzer KV, Koss MP (2004) Narrative constructions of sexual violence as told by female rape survivors in three populations of the southwestern United States: scripts of coercion, scripts of consent. *Med Anthropol* 23:113–156. doi:10.1080/01459740490448911

- Bohmer C, Blumberg A (1975) Twice traumatized: the rape victim and the court. *Judicature* 58:391–399
- Brehm SS, Kassir S, Fein S (2005) *Social psychology*. Houghton Mifflin Harcourt, Boston
- Canter DV (2000) Offender profiling and criminal differentiation. *Leg Criminol Psychol* 5:23–46. doi:10.1348/135532500167958
- Canter DV (2004) Offender profiling and investigative psychology. *J Investig Psychol Offender Profiling* 1:1–15. doi:10.1002/jip.7
- Canter DV, Bennell G, Alison LJ, Reddy S (2003) Differentiating sex offences: a behaviorally based thematic classification of stranger rapes. *Behav Sci Law* 21:157–174. doi:10.1002/bsl.526
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20:37–46
- De Zutter AWEA, Horselenberg R, Van Koppen PJ (2016) Filing false vice reports: distinguishing true from false allegations of rape. *Eur J Psychol Appl Leg Context*. doi:10.1016/j.ejpal.2016.02.002
- DePaulo BM, Lindsay JJ, Malone BE, Muhlenbruck L, Charlton K, Cooper H (2003) Cues to deception. *Psychol Bull* 129:74
- Ellison L, Munro VE (2009) Of 'normal sex' and 'real rape': exploring the use of socio-sexual scripts in (mock) jury deliberation. *Soc Leg Stud* 18:291–312. doi:10.1177/0964663909339083
- Erickson PI, Rapkin AJ (1991) Unwanted sexual experiences among middle and high school youth. *J Adolesc Health* 12:319–325
- Gannon TA, Collie RM, Ward T, Thakker J (2008) Rape: psychopathology, theory and treatment. *Clin Psychol Rev* 28:982–1008. doi:10.1016/j.cpr.2008.02.005
- Gigerenzer G (2002) The adaptive toolbox. In: Gigerenzer G, Selten R (eds) *Bounded rationality: the adaptive toolbox*. MIT Press, Cambridge, pp 36–50
- Gigerenzer G, Gaissmaier W (2011) Heuristic decision making. *Annu Rev Psychol* 62:451–482
- Greer E (1999) The truth behind legal dominance feminism's 'two percent false rape claim' figure. *Loyola Los Angeles Law Rev* 33:947–972
- Greer C (2003) Crime in the press: sex offending and the media in a divided society. Willan, Cullompton
- Gregory J, Lees S (1996) Attrition in rape and sexual assault cases. *Br J Criminol* 36:1–17
- Gross SR, Jacoby K, Matheson DJ, Montgomery N, Patil S (2005) Exonerations in the United States 1989 through 2003. *J Crim Law Criminol* 95:523–560
- Haket V (2007) *Veranderende verhalen in het strafrecht: De ontwikkeling van verhalen over verkrachting in het strafproces*. [Changing narratives in criminal law: The development of narratives concerning rape during criminal proceedings]. (Doctoral dissertation). Universiteit Leiden, Leiden, The Netherlands
- Horowitz SW (2009) Direct mixed and open questions in child interviewing: an analog study. *Br Psychol Soc* 14:135–147
- Hunt L, Bull R (2012) Differentiating genuine and false rape allegations: a model to aid rape investigations. *Psychiatry Psychol Law* 19:682–691. doi:10.1080/13218719.2011.615815
- Kanin EJ (1994) False rape allegations. *Arch Sex Behav* 23:81–92. doi:10.1007/BF01541619
- Kelly L, Lovett J, Regan L (2005) A gap or a chasm? Attrition in reported rape cases. Home Office Research, Development and Statistics Directorate, London
- Knight RA (1999) Validation of a typology for rapists. *J Interpers Violence* 14:303–330. doi:10.1177/088626099014003006
- Kocsis RN, Cooksey RW, Harvey JI (2002) Psychological profiling of offender characteristics from crime behaviors in serial rape offences. *Int J Offender Ther Comp Criminol* 46:144–169. doi:10.1177/0306624X02462003
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
- Lisak D, Gardinier L, Nicksa SC, Cote AM (2010) False allegations of sexual assault: an analysis of ten years of reported cases. *Violence Against Women* 16:1318–1334. doi:10.1177/1077801210387747
- Marshall BC, Alison LJ (2006) Structural behavioural analysis as a basis for discriminating between genuine and simulated rape allegations. *J Investig Psychol Offender Profiling* 3:21–34. doi:10.1002/jip.42
- Masip J, Herrero C (2013) 'What would you say if you were guilty?' Suspects' strategies during a hypothetical behavior analysis interview concerning a serious crime. *Appl Cogn Psychol* 27:60–70. doi:10.1002/acp.2872
- McDowell CP, Hibler NS (1993) False allegations. In: Hazelwood RR, Burgess AW (eds) *Practical aspects of rape investigation: a multi-disciplinary approach*. CRC Press, London, pp 275–300
- Moore DS, McCabe GP, Craig BA (2012) *Introduction to the practice of statistics*. W. H. Freeman, New York
- Norton R, Grant T (2008) Rape myth in true and false rape allegations. *Psychol Crime Law* 14:275–285. doi:10.1080/10683160701770286
- O'Sullivan LF, Allgeier ER (1998) Feigning sexual desire: consenting to unwanted sexual activity in heterosexual dating relationships. *J Sex Res* 35:234–243. doi:10.1080/0022449809551938
- Parker AD, Brown J (2000) Detection of deception: statement validity analysis as a means of determining truthfulness of falsity of rape allegations. *Leg Criminol Psychol* 5:237–259. doi:10.1348/135532500168119
- Philips LM (2000) *Flirting with danger: young women's reflections on sexuality and domination*. New York University Press, New York
- Prentky RA, Knight RA (1991) Identifying critical dimensions for discriminating among rapists. *J Consult Clin Psychol* 55:643–661. doi:10.1037/0022-006X.59.5.643
- Rassin E (2001) Het bepalen van geloofwaardigheid: De methode van de alternatieve scenario's. [Determining credibility: the method of alternative scenarios]. *De Psycholoog* 36:348–355
- Rassin E, Van der Sleen J (2005) Characteristics of true versus false allegations of sexual offences. *Psychol Rep* 97:589–598
- Rumney PNS (2006) False allegations of rape. *Camb Law J* 65:128–158. doi:10.1017/S0008197306007069
- Saunders CL (2012) The truth, the half-truth, and nothing like the truth: reconceptualizing false allegations of rape. *Br J Criminol* 52:1152–1171. doi:10.1093/bjc/azs036
- Schwabe L, Wolf OT, Oitzl MS (2010) Memory formation under stress: quantity and quality. *Neurosci Biobehav Rev* 34:584–591
- Shields WM, Shields LM (1983) Forcible rape: an evolutionary perspective. *Ethol Sociobiol* 4:115–136
- Sim J, Wright CC (2005) The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 85:257–268
- Soothill K, Walby S (1991) *Sex crime in the news*. Routledge Keegan Paul, New York
- Sporer SL (1997) The less traveled road to the truth: verbal cues in deception detection in accounts of fabricated and self-experienced events. *Appl Cogn Psychol* 11:373–397
- Strömwall LA, Hartwig M, Granhag PA (2006) To act truthfully: nonverbal behaviour and strategies during a police interrogation. *Psychol Crime Law* 12:207–219. doi:10.1080/10683160512331331328
- Taylor J (1987) Rape and women's credibility: problems of recantation and false accusations echoed in the case of Cathleen Crowell Webb and Gary Dotson. *Harv Women's Law J* 10:59–116
- Undeutsch U (1982) The development of statement reality analysis. In: Yuille JC (ed) *Credibility assessment*. Kluwer, London, pp 101–120
- Viera AJ, Garrett JM (2005) Understanding interobserver agreement: the kappa statistic. *Fam Med* 37:360–363
- Vrij A (2005) Criteria-based content analysis: a qualitative review of the first 37 studies. *Psychol Public Policy Law* 11:3–41. doi:10.1037/1076-8971.11.1.3
- Vrij A, Edward K, Roberts KP, Bull R (2000) Detecting deceit via analysis of verbal and nonverbal behavior. *J Nonverbal Behav* 24:239–263
- Woodhams J, Grant T (2004) Investigating differences in reported rapists' speech between maintained as true and false allegations, Unpublished Manuscript. School of Psychology, University of Birmingham, Birmingham