Distinguishing between True and False Allegations of Sexual Offenses

Gus Lathouwers
Tilburg University

Gus Lathouwers, S688790, 26-8-2016
Supervisor: K. M. E. Lens
Abstract

Trying to understand how false allegations of sexual offenses differ from true allegations of sexual offenses has long been of interest to researchers and law enforcement alike. The aim of the current study was to review currently available literature on what sets apart true from false sexual offenses. This was done by contrasting statement reports made by victims of true sexual offenses with statement reports that were fabricated by pseudo-victims. A total of eight studies were compared in-depth in regards to their methodology and outcome. Three major clusters of behaviors emerged unique to statements of false allegations. First, in true allegations of sexual offenses perpetrators tended to talk a lot more, while in false allegations the perpetrator was mostly silent. Second, true allegations of sexual offenses had a much wider range of sexual acts, such as foreplay and oral acts. These were rarely present in false allegations of sexual offenses. Third, in true allegations criminal acts such as stealing money or items by the perpetrator were common. In false allegations however, such criminal acts were almost never mentioned. The current study concludes that there are viable markers to indicate whether a statement is more likely to be true, but that the research available is still in its preliminary stages and that follow-up studies are required to paint a more complete picture.

*Keywords*: false rape allegations, characteristics, pseudo-victim, red-flags
Distinguishing between True and False Allegations of Sexual Offenses

False accusations of sexual offenses have been the subject of much controversy, both in professional circles as well as in public discourse. Part of the reason for this is that rape cases are difficult to evaluate and investigate by the police (Gibbon, 1998). In the past, detectives have been scrutinized over their skeptic attitude towards victims of rape and how they handle cases (Gregory & Lees, 1996). This is not in the least due to ubiquitously present rape myths, which can be described as “prejudicial, stereotyped, or false beliefs about rape, rape victims, and rapists” (Burt, 1980, p. 217). Rape myths are also still widely being spread by news agencies (Ardovini-Brooker & Caringella-Macdonald, 2002), who tend to selectively pick and report about sensationalist cases because they comply with stereotypes and common prejudices held by the public (Greer, 2003). This leads to very wrong public perceptions about what actual rape looks like, such as for example the idea that rape is usually violent and that the victims are at fault for provoking the rape (Norton & Grant, 2008). In the Netherlands, a recent government report also concluded that sexual offense cases are sometimes being handled negligently by the Dutch police (Nierop & Eshof, 2008). Victims are not always given proper information about their common proceedings, and sometimes inexperienced officers are used to conduct interviews where more experienced officers are officially required (Nierop & Eshof, 2008).

It is not only the public sphere that is having trouble understanding what the true nature of rape looks like. The scientific research that has been done during the last two decades has led to a very confusing overall imagine as to what exactly constitutes a false allegation of a rape and how sexual offenses in general occur. When it comes to prevalence studies of false rape allegations, outcomes vary as much as between 2% and 90% (Rumney, 2006). As a consequence, many studies seem to arbitrarily declare one of these numbers as truth, some readily accepting figures as high as 40% (Rassin & Sleen, 2005) and others as low as 2% (Norton & Grant, 2008). Many studies are also inconsistent in the terminology and methodology they use. The definition as to what exactly constitutes a false allegation of rape or a false allegation of sexual offense differs from study to study, and can lead to highly
CHARACTERISTICS OF SEXUAL OFFENSES

divergent research outcomes, because different standards are being used to judge the same phenomenon (Saunders, 2012).

The mass misinformation that exists about false rape accusations carries with it very profound consequences. For detectives, wrongfully thinking a rape allegation is fabricated while in fact it is genuine can destroy the bond of trust that exists between police and a victim of a crime (Hazelwood & Burgess, 2009). The reverse also holds true: not recognizing that a rape allegation is fabricated can lead to an innocent person being convicted for a crime he did not commit with dire consequences (Hazelwood & Burgess, 2009). A compounding factor here is that detectives have very little awareness that their judgments may be unfounded, and often have a misplaced trust in their ability to intuitively distinguish true from false allegations of rape (Vrij & Winkel, 1993). In courtrooms, mistaken perceptions of what rape looks like may lead to wrongful convictions of innocent persons (Ellison & Munro, 2009). False accusations of sexual acts have also been employed in family disputes by mothers seeking to get custody of their child, and such allegations have traditionally been met with very little scrutiny from police and judges because they are assumed to automatically be true (Wakefield & Underwager, 1990).

Because false rape allegations are so hard to detect and require a sensitive approach to deal with, identifying objective characteristics that do not rely on heuristics would ostensibly be a boon for law enforcement and judiciaries. Since the 90's, there has been an increased focus on identifying the characteristics that set false allegations of sexual offenses apart from true allegations. These are sometimes referred to as 'red-flags' (Lonsway, Archambault, & Lisak, 2010), and such lists of characteristics have been formulated using different approaches, including Criteria-Based Content Analysis (CBCA) (Steller & Köhnken, 1989) and rape myth based-systems (Norton & Grant, 2008). Other authors choose to corroborate and examine lists originating from other sources (Rassin & Sleen, 2005), or rely on subjective experience to create lists (Hazelwood & Burgess, 2009). When reviewing the existing literature, it is important to carefully define the used terminology. Previously, 'pseudo-victim' has been used to refer to the person making a false allegation of
a sexual offense, while 'victim' refers to someone who was the subject of a sexual offense that is not fabricated (Hazelwood & Burgess, 2009). This same distinction will be used here. In addition, 'rape' will refer to non-consensual sexual intercourse that is committed by physical force, threat of injury, or other methods of duress (FindLaw, 2016). This is distinctly different from 'sexual abuse', which covers any type of sexual activity that a person does not agree to, including physical touching, sexual contact with a child, or indecent exposure (OWH, 2015). Lastly, the term 'sexual offense' refers to any offense of a sexual kind, and will be used in the most global sense here, meaning it refers to all possible sexual criminal acts including rape and sexual abuse.¹

Defining and Investigating False Allegations of Sexual Offenses

Defining false allegations of sexual offenses can be approached in a number of different ways. A most important distinction to be made is how to define rape, which is an important topic in many papers investigating false allegations of sexual offenses. Adshead (1996) includes a number of different possible definitions of 'false allegations of rape', for example using 'false rape allegation' to signify an allegation was filed because of malicious intent (i.e. the pseudo-victim makes the allegation as a conscious choice to get back at someone motivated by revenge). Other authors take a neutral approach, such as Kanin (1994), describing a false rape allegation as “the intentional reporting of a forcible rape by a victim when no rape occurred” (p. 82). For Kanin then, ulterior motives such as material gain are also covered under this definition. Gibbon (1998) provides another set of factors that are important to take into account when formulating a working definition of false allegation of sexual offenses and rape, namely whether to include: allegations that were retraced, allegations that were malicious, allegations that were made but not proceeded with due to pressure from suspect or fear of court, allegations that lack proof (e.g., no offender

¹ The reason 'sexual offense' is a term often used in this paper over other specific terms such as 'rape' or 'sexual assault', is that the literature sometimes uses terms such as 'rape' and 'sexual assault' without clearly defining their boundaries. As such, it is not always clearly understandable to what extent certain papers deal with only rape, or whether they also principally include other forms of sexual offenses in their sample such as sexual assault. Because of this, employing the term 'sexual offense', while broad and somewhat vague, is in some cases deemed more suitable because it prevents misinterpretation of vague subject matters.
CHARACTERISTICS OF SEXUAL OFFENSES

identified), mistaken allegations (i.e. interpreting a sexual experience as rape while it is consensual), and allegations made under coercion.

In order to make the right choice as to which components to include in the definition, it is imperative to understand the motives that people have when making a false allegation. One of the first researchers to address this question was Kanin (1994), who identified three major motives behind fabricating a rape allegation. The first was 'alibi', which was present in more than half of the cases Kanin reviewed. The motive of 'alibi' refers to a pseudo-victim fabricating an allegation to purposely cover up for something, for example when they are caught committing adultery by their spouse (Kanin, 1994). The second motive was 'revenge', which Kanin found was present in a little over a quarter of the cases. The third was sympathy and attention, present in less than 15% of the cases. Though Kanin's findings are corroborated by some sources (e.g., McNamara, McDonald & Lawrence, 2012), contemporary research seems to suggest a more nuanced picture. While all three motives presented by Kanin do seem to be present, other additional ones have also been identified as important contributors, such as regret (Zutter, Horselenberg, & Koppen, 2016a).

Additionally, the prevalence of the motive 'revenge' may be only present in less than one-tenth of the cases instead of a quarter, and 'material gain' seems to be occurring very seldom as a motive to file a false rape allegation (Zutter, Horselenberg, & Koppen, 2016a).

Because malicious intent and revenge are not a primary motivator in many cases (Zutter, Horselenberg, & Koppen, 2016a), defining false rape allegations may be better served by focusing on whether the events that are being claimed factually took place or not. However, this presents a problem in practice. In a great number of field cases processed by law police, there is no objective proof that can be used to verify for certain if an allegation is false or true (Spalding & Bigbee, 2009). Because of this, knowing whether a victim is forging an allegation is difficult to establish. Researchers choosing to rely on police records then have to use their own criteria to determine the veracity of statements. As Saunders (2012) notes, across studies these self-defined criteria can range from very strict (accuser admits to having fabricated the allegation) to very loose (case with was not proceeded with because
there was no physical evidence to support the allegation). This difference in criteria that researches use also contributes to the high variance found in prevalence ratings of false allegations (Rumney, 2006).

In light of these difficulties, some researchers choose to disregard real field cases and instead create their own mock false statements so they are certain that their sample of false rape allegations isn’t contaminated by true allegations (e.g., Norton & Grant, 2008; Zutter, Horselenberg, & Koppen, in press). However, asking participants to create mock statements has obvious repercussions for ecological validity and can produce results that are very different from research that samples false allegations from field cases (Koppen, 2012). These discrepancies in methodology again highlight the importance of accurate definitions and consistent approaches in trying to investigate false rape accusations.

**Investigating the Characteristics of False Rape Allegations**

Trying to identify false allegations by pinpointing their unique characteristics has long been of interest to researchers and law enforcement alike. In the past, when lists of red-flags were not readily existent and awareness about rape myth was still largely absent, detectives had to rely on false stereotypes of what rape looks like to judge the truthfulness of rape allegations (Hazelwood & Burgess, 2009). As such, women who appeared either overly hysterical or too subdued would be judged insincere and their allegations of rape would be met with disdain (Hazelwood & Burgess, 2009). During the 90’s, focus shifted to constructing lists of objective red-flags, such as those provided by Hazelwood & Burgess (2009) (See Appendix A). These kinds of red-flag lists were created by forensic experts based on personal experience in the working field, but were not directly based on empirically evidence. Other related lists include the Baeza False Report Index (BAFRI) and the Dietz and Hazelwood red-flag list, both created by forensic or psychiatric experts based on personal experience and personal consultation (Turvey & Petherick, 2009). Lists such as these are being widely taught as tools to use in investigative handbooks on rape and sex (e.g., Turvey & Petherick, 2009; Hazelwood & Burgess, 2009), but because they have never been empirically verified by independent research it is hard to ascertain at face value to which
Besides experience-based lists such as these, other research has tried to identify rape characteristics based on empirical research. One of these revolves around the concept of rape myths, namely: how the public perception on rape differs from the perception of victims who have actually endured a rape (e.g., Norton & Grant, 2008). The accompanying idea is that people fabricating a rape story do not know what 'real rape' looks like, and as such the stories from these pseudo-victims contain a lot of markers that do not all comply with how rape typically takes place. A big body of research has been devoted to understanding how rape transgresses, such as that it is done more often by familiars than strangers, and that contrary to popular belief there is actually very little physical violence involved (Knight, 1999). Because these realities of rape are different from the myths held by the public, researches can generate mock statements by recruiting people who never experienced rape and compare their stories with those of true victims of rape.

Another way to identify red-flags is using classic lie detection theories and applying them to examine sexual offense allegations. One popular means developed by Steller and Köhnken (1989) is the Criteria-Based Content Analysis system. CBCA is based on the Undeutsch hypothesis (Undeutsch, 1967), which asserts that memories from real experienced events are different from those that are produced from imagination. The CBCA checklist consists of 19 criteria that can be used to analyze characteristics of a statement, such as structure, the amount of superfluous details, and presence of self-deprecation (Steller and Köhnken, 1989). CBCA is also part of Statement Validity Assessment (SVA), which takes CBCA and adds a probabilistic guide to estimate the likelihood that the CBCA analysis is true (Amado, Arce, & Fariña, 2015). As such, SVA is meant as an improved method to make the results from a CBCA assessment more viable in practical applications. Though CBCA can be used in the detection of many different sorts of lies, it has been applied to examine sexual offense allegations specifically (e.g., Parker & Brown, 2000). CBCA has been found to have good reliability and validity (Amado, Arce, & Fariña, 2015), though this is contested by other authors (Zutter, Horselenberg, & Koppen, 2016b). In 2000,
Vrij stated that, if applied correctly, the CBCA can raise the likelihood to correctly identify lies and false statements significantly. However, Vrij himself (2000; 2005) also addressed limitations associated with the CBCA system, such as that the studies that perform research on it are sometimes of poor quality.

When talking about identifying characteristics of false rape allegations, a few more things are important to address. First, some researches choose not to apply a specific method of investigation but simply take a long list of characteristics and then examine how often they show up in true and false allegations (e.g., Rassin & Sleen, 2005; Zutter, Horselenberg, & Koppen, in press). While these kinds of studies typically have overlap with using rape myths to identify false allegations, they are usually much more exploratory in nature and do not rely on a specific system or underlying principle. Second, knowing and understanding where people making false rape accusation come from (demographic characteristics such as education, age) and how they function psychologically (personality characteristics such as extroversion, neuroticism) could theoretically be very beneficial to detecting false allegations. In practice however, as shall be discussed in the results section, there is still far too little data available on this matter to make definitive claims. Of course, given the practical difficulty of getting pseudo-victims to agree to partake in research, this is something that might not ever be susceptible to be studied in-depth.

**Current study**

The aim of the current paper is to compare results from different studies in order to examine whether there exist unique characteristics that can be used to identify false sexual offense allegations. Because researchers implicitly employ different definitions and research approaches, these shall be looked at before describing the outcomes of different studies. In part, it shall be assessed whether the criteria described by Gibbon (1998) are included in the definitions used in studies, namely: allegations that were later retraced, allegations that were malicious, allegations that were made but not proceeded with due to pressure from the suspect or fear of court, allegations that lack proof (e.g., no offender identified), mistaken allegations (i.e. interpreting a sexual experience as rape while it is consensual), and
allegations made under coercion. Both empirical studies as well as lists based on personal experience (e.g., Hazelwood & Burgess, 2009) will be looked at. Afterwards, regarding lists based on personal experience, it shall be looked at whether these are in line with findings from empirical research or not. The research questions used as a basis for this literature review is ‘to what extent do false allegations of sexual offenses differ from true allegations of sexual offenses in the unique characteristics they exhibit?’ Within this research goal, the specific aim will be to examine sexual offenses that are reported shortly from the time they were committed, e.g., rapes that were reported to the police within a short time frame after they were committed. This notably excludes accusations of childhood abuse or accusations otherwise made a significant time after the alleged event took place.

Method

Methods of Study Collection and Selection

Research papers on sexual offense allegations were identified using online search databases and the snowball-technique. For online research engines, three databases were used, namely: PubMed, ScienceDirect, and Google Scholar. Only the Google Scholar searches resulted in studies that were viable to be used in the the review section of this paper. For Google Scholar, 'rape allegations' was used as a search term, resulting in 56,600 article hits. Because Google Scholar sorts hits on relevance, only the first 30 pages of results were inspected; the rest strayed too far from the subject matter. In total, the Google Scholar search resulted in four articles that were used in the review section (Hunt & Bull, 2012; Parker & Brown, 2000; Marshall & Alison, 2006; Norton & Grant, 2008). One article was obtained through personal means of communication: Zutter, Horselenberg, and Koppen (in press). Two other articles used in the review section were obtained using the snowball-technique. For example, inspection of the reference list from Zutter, Horselenberg, and Koppen (in press) lead to the discovery of the study from Rassin and Sleen (2005). Finally, inspecting the reference list from Rassin and Sleen (2005) led to a chapter in the book of

---

2 For an extended overview of Google Scholar’s Ranking algorithm, see Beel and Gipp (2009). In short, Google Scholar uses an algorithm that factors in a number of variables (e.g., citation counts of articles), and then sorts the results based on an aggregate of these individual variables.
Hazelwood and Burgess (2009). See Figure 1 for a schematic overview of how the different studies were obtained.

Though Pubmed, ScienceDirect, and related search engines were used, ultimately only Google Scholar produced results used in this review. Through Google's general search engine (not Google Scholar), many more search queries were also attempted but resulted in no new articles apart from the ones already found. For each of the search engines used (e.g., ScienceDirect, Pubmed, Google), an exhaustive list of search queries was applied, including 'false rape allegations,' 'false rape accusations,' 'false allegations of sexual offenses,' 'sexual offenses,' 'rape accusations,' and 'rape allegations' among others. From these search queries, it was looked at whether the titles of the papers that were found were possibly relevant to the review. In addition, if paper titles were ambiguous, their abstracts were read to understand if they related to false allegations of sexual offenses. It may appear surprising that in the end, so few results (eight papers in total) proved viable to be included in this review study. There may a number of possible reasons to account for this. First, a number of articles that deal with false allegations of sexual offenses were excluded because they differed too much from the original research goal from this paper. For example, allegations done by adults about childhood sexual abuse or assault are not included in this review. These claims have been met with serious skepticism, and their scope is too broad and different from rape police-filing cases to be reviewed here (see Loftus and Davis, 2006). Similarly, allegations done by children towards their parents or other adults are also excluded in this paper. Because the motives and method of conduct of these allegations differ so strongly from those done by adults (Bernet, 1993), they are too dissimilar to be included in the current review.

Another important point that contributes to the low total of papers included in this review is the simple fact that there are not many studies who have investigated this topic yet. Though in recent years there has been an improved interest in compiling lists of 'red-flags' to identify false sexual offense allegations, as it stands most of the research is still very preliminary. The Google Scholar search as such yielded very little results that were closely
related to investigations of false allegations of sexual offenses or rape. An added difficulty that contributed to the low total of eight papers is the fact that some papers identified using the back-tracking method could not be retrieved digitally. These four studies (Rainbow, 1996; Woodhams & Grant, 2004; McDowell, 1992; Winkel, Vrij, Koppelaar, & Van der Steen, 1991) were particularly relevant for this paper's subject matter, and would have probably been included if they could be digitally retrieved. Three of these papers (Rainbow, 1996; Woodhams & Grant, 2004; Winkel, Vrij, Koppelaar, & Van der Steen, 1991) were identified by inspecting the Norton and Grant (2008) list of references used. One such paper (McDowell, 1992) was identified by tracking the reference list of Parker and Brown (2000).

Because of the practical limitation of not having access to these four studies they could unfortunately not be included in the present review. Figure 1 shows a summary of all studies used in the review section, and also those that were deemed relevant but could not be accessed online.

As stated earlier, from the Google Scholar search, four studies emerged that were used in the review section of this paper (Hunt & Bull, 2012; Parker & Brown, 2000; Marshall & Alison, 2006; Norton & Grant, 2008). All of these were particularly valuable because they investigate lists of characteristics to see how often they occurred in false and true allegations. Examples of such characteristics are 'presence of injuries' or 'presence of deviant sexual acts'. If a particular characteristics occurred much more in either true or false allegations, these could then be used as a signifier and ostensibly be added to a 'red-flag' list. Of the total eight papers that were selected to be included in the review, seven were empirical studies and the remainder (Hazelwood and Burgess, 2009) was of subjective non-empirical origin (see Table 1). One article written by Zutter, Horselenberg, and Koppen (in press) was also obtained through means of personal communication, but was publicized online before the final version of this paper was written. Figure 1 summarizes the full data retrieval and study selection process.

**Definitions Utilized and Methodology**

The definitions that are used when studying false allegations of rape and sexual
offenses are of great importance on the subsequent results found. Of the seven studies reviewed in this section (not including Hazelwood and Burgess which is not empirically based), six had a very similar research design. The crux of this design is to take a number of statements fabricated by pseudo-victims (or alternatively mock statements), and then contrast how these differ from statements made by actual victims of sexual offenses\(^3\). Only one study (McNamara, McDonald, & Lawrence, 2012) relied on exclusively false statements in researching false allegations of sexual offenses. In all, four studies used real life field allegations taken from a police database to study characteristics, two studies relied on exclusively mock statements to investigate the characteristics of false rape allegations. One study (Norton & Grant, 2998) used both real false sexual offense allegations and generated mock statements as a basis of their research. All studies that used real cases of false allegations provided some criteria on which they based their selection. Critically, two studies (Marshall & Alison, 2006; Norton & Grant, 2008) did not mention any criteria on which they selected truthful statements from police files. Though selection criteria for truthful cases have not been highlighted in academic papers very much, it is a very important aspect to consider for reasons similar to the need to define false allegations accurately. All results pertaining to the definitions and samples used by the studies are summarized in Table 2.

**False allegation definitions used.** Of the five studies that took real field cases of false allegations as a basis for their analysis, only two used selection criteria that were similar to each other (Norton & Grant, 2008; Hunt & Bull, 2012). When compared with the criteria given by Gibbon (1998), many of the definitions employed in the different studies are relatively loose and vague. The most questionable definition employed is that of Rassin and Sleen (2005). In their design, Rassin and Sleen sent an open survey to police officers throughout the Netherlands and asked them to select a true allegation and a false allegation from police files available to them. Police officers that obliged returned the survey to Rassin and Sleen without the case files included, meaning Rassin and Sleen had no way of

---

3 Both Rassin and Sleen (2005) and Parker and Brown (2000) did not test their individual characteristics for statistical significance, and as such their outcomes should be interpreted with caution. All other findings described in the results section will refer to statistically significant results only, unless otherwise specified.
verifying the veracity of files selected by the police officers. The only criteria Rassin and Sleen applied themselves is that the case files selected by the police officers resulted in a conviction for the offender in the truthfulness condition, and a conviction for the victim in the false allegation condition. As such, when police officers coded statements, they did so on statements they themselves selected. Because they were aware whether an allegation they selected was a true or false one, the coding of these statements could have been seriously biased.

Besides Rassin and Sleen, all other sources did have direct access to the statements they analyzed. Norton and Grant (2008) and Hunt and Bull (2012) both employed the same base criteria, namely: that the victim withdrew their complaint and admitted that it was forged; or that the police filed a prosecution against the victim for wasting police time. While the first criteria (victim withdrew statement with admittance of falsity) is a reasonably viable indicator of a false allegation⁴, for the second criteria (victim was charged with wasting police time) this is not necessarily so. For the second criteria, two of Gibbon's (1998) criteria directly apply. First is the possibility of a false negative, i.e. no physical evidence causes the police to mistakenly assume an allegation is made up and prosecute the victim for wasting time. Alternatively, police and victim may disagree about whether the event truly constitutes a rape, i.e. the victim perceived the events as rape while police judged it as consensual. This could lead the police to perceive the victim as a pseudo-victim that intentionally faked the allegation andwrongfully prosecute them for wasting police time. Additionally, a police report from the UK found that a sizable proportion of the victims being charged with wasting police time were under the influence of drugs or alcohol, either when the sexual act happened or when they made the report (CPS, 2013). Being under the influence of drugs or alcohol when the act happened can contribute to a false memory of the events, making the allegation much more eligible to fall under the category of a 'mistaken' allegation than a 'false' one. The

---

⁴ It should be noted that in some cases it is perceivable that a victim withdraws a statement with admittance of falsity even though it wasn't actually a fabricated allegation. Possible reasons for this may be: the victim is pressured by extraneous circumstances to withdraw the complaint (e.g., family, threats from the perpetrator) (Gibbon, 1998), or the victim decided to withdraw because the prospect of going through with the allegation is too much of an emotional burden.
Hunt and Bull study partly accounted for this by excluding cases that were made as a result of confusion on the victim’s part, and also excluding cases that later turned out to be consensual. The Norton and Grant study however did not.

Parker and Brown (2000) used four different criteria to select their false allegations, namely: the victim withdraws the allegation with an admittance of falsity, the physical evidence found contradicts the allegation, the allegation was the result of a delusion or substance-abuse, or the allegation was made for specific inappropriate reasons. Only one of these criteria needed to be present for the allegation to be deemed as false by Parker and Brown. As is the case with the selection criteria used by Norton and Grant (2008) and Hunt and Bull (2012), these four criteria are not foolproof indicators that the allegation was indeed false and not 'mistaken' or 'unfounded'. The choice to include allegations made on the grounds of 'specific inappropriate reasons' is not expanded upon, and as such it is difficult to ascertain the validity of these allegations. Finally, McNamara, McDonald and Lawrence (2012) examined 30 false allegations submitted to the FBI from various cases throughout the United States that also had a conviction attached to them. They themselves also independently inspected these cases for truthfulness, though without explicit mention of which method of assessment they used for this.

Of the three studies that used mock statements generated by participants, two mention the conditions under which these statements were conceived. Norton and Grant (2008) describe their participants as being female and recruited through newspaper advertisements, all between 18 and 25 year old. However, while these participants were asked to rate their previous sexual experience, there is no mention of screening for rape or sexual abuse in this group. Considering the topic of their investigation concerns rape myths, having actually experienced rape or sexual abuse could have served as a serious confounding variable. Marshall and Alison (2006) provide almost no information about the participant sample they used to generate mock rape statements, neglecting to provide for example their participants’ age, ethnicity, or sexual experiences. It only mentions that their fictitious mock statements originate from a sample produced in a previously unpublished
CHARACTERISTICS OF SEXUAL OFFENSES

study, but no other information is provided. Zutter, Horselenberg, and Koppen (in press) are relatively thorough in their review of their participant sample. In their study, all females were above the age of 18 and with a mean age of 28, closely matching the demographics of the real victims in the truthfulness condition group. Participants were also screened for previous rape experience, and any persons who felt uncomfortable was allowed to withdraw from the study. All three studies however lack data about participants when it comes to things such as ethnicity or social economic status.

True allegation definitions used. Of the six studies that used police files as the basis for their true allegations condition, four mentioned explicit criteria used to select these cases. Rassin and Sleen (2005) included any survey submitted to them by a police officer that resulted in a conviction for the offender. Again, because they had no access to these files directly, they could not have verified the veracity of these statements themselves or see whether the coding by the police agents was done properly. Hunt and Bull (2012) similarly only included cases from police files in which the offender was convicted. Only using the presence of a conviction as a criteria may introduce the problem of false positives being included in the 'truthful allegations' condition. Parker and Brown (2000) used an entirely different set of criteria to include allegations in their truthfulness condition, namely: "[the] presence of convincing evidence of rape corroboration in the legal sense and with either a suspect being identified or charged" (Parker & Brown, 2000; p. 241). While the authors note that all care was taken to ensure accuracy, it is hard to exactly ascertain what this definition means in practice. Zutter, Horselenberg, and Koppen (in press) utilized more strict criteria to select their sample. In their study, allegations were only included in the 'truthfulness' condition if they contained a confession by the perpetrator. An additional requirement to the confession was some sort of physical evidence in the form of: a DNA match, identification by the victim, the offender being caught in the act, strong guilty knowledge presented in the confession by the offender, or the offender having stolen items from the victim. While 'strong guilty knowledge' is somewhat subjective, using a confession of the offender is a marked improvement over only selecting cases based on conviction. Two studies (Marshall & Alison,
Methodology. With the exception of McNamara, McDonald and Lawrence (2012), the six different empirical studies used very similar research designs, making comparisons between them possible (see Table 3). Some of the studies investigated similar characteristics, such as Parker and Brown (2000) and Rassin and Sleen (2005), who both investigated CBCA characteristics. The number of statements analyzed differ wildly between the studies (e.g., $N = 41$ in Rassin & Sleen, 2005; $N = 240$ in Hunt & Bull, 2012), as do the number of characteristics that were analyzed for (e.g., 31 in Parker & Brown, 2000; 187 in Zutter, Horselenberg, & Koppen, in press). Of great importance is the way the different studies coded the statements, i.e. the method they used to assess if a characteristic was present in a certain statement. All studies used dichotomous coding, meaning attributing a '1' if a characteristic was present and a '0' if it was not present. The only exception to this is Rassin and Sleen (2005), who also added a 'do not know/not applicable' option that counted for '0'. However, this study does not mention whether they compensated for the fact that 'not present' and 'not applicable' both got attributed a score of '0'.

The statements were coded by different sorts of parties across the studies. Some studies used independent coders, others did the coding themselves, still others relied on police detectives to code the characteristics. Zutter, Horselenberg, and Koppen (in press) are the only ones that mention using independent coders exclusively to prevent confirmation bias. Norton and Grant (2008) mention that the true and false allegation conditions were coded blindly, but that due to an error the mock statements were not coded blindly. Subsequently they accepted recommendations from a third party that was blind to the nature of these allegations. It is not specified why they did not employ a true impartial independent coder to ensure all coding was done blind instead. Parker and Brown (2000) coded the statements themselves, and as Vrij (2005) notes in a meta-review, this compromises the validity of their results to a considerable extent. In a similar way, Marshall and Alison (2006) also did not use independent evaluators to code their statements, though they did mention
that the coding system they employ only has a very low 3-4% error rate (Marshall and Alison, 2006). As mentioned before, Rassin and Sleen (2005) instructed police officers to select both a true and a false allegation themselves and then code them. Because these agents were completely aware of whether the statement they coded was true or false, characteristics on the survey such as 'egocentricity' and 'emotional problems' could have been met with bias. Lastly, the characteristics in Hunt and Bull's (2012) study were coded by the agency who processed the rape allegations, namely Serious Crime Analysis Section of the United Kingdom's National Crime Agency. However, because many of these characteristics are descriptive in origin (i.e. ethnicity victim, involvement of weapon), validity is less likely to be compromised than in other studies mentioned here.

Results

Before venturing into the results from the empirical studies, some key points described in the list provided by Hazelwood and Burgess (2009) shall be shortly described (see Appendix A for the full list). It is important to note that Hazelwood and Burgess argue that their list is only meant as a guideline, and not a checklist to verify whether a statement is true or false. As such, they only see it as a tool to help raise attention to cases that may merit further scrutinizing. In their list of characteristics, a first theme that emerges is the role of emotionality and mental status of the victim. The list states that pseudo-victims are more likely to exhibit attention-seeking behavior (e.g., in the form of Munchhausen syndrome), to have underlying emotional problems, or a history of psychiatric illness. A second theme that emerges is that pseudo-victims according to Hazelwood and Burgess tend to change their story a lot and recall additional details in subsequent hearings. They also tend to have difficulty describing their assailant and do not want to focus on the facts of the case, but rather the force during the offense and injuries that were sustained. A third theme present in the Hazelwood and Burgess list is that pseudo-victims are not likely to report deviant sexual acts, and that they are more likely to present with injuries and ripped clothing. Lastly, they state that pseudo-victims are more likely to report the claim to the police after a delayed time period (for example a week instead of 24 hours), and that the pseudo-victim is likely to be
Characteristics of False Allegations in Empirical Studies

When describing the results of the seven empirical studies reviewed, a first thing to note is that many authors were fast to acknowledge that it is hard to make a distinction between true and false allegations only based on a list of characteristics. For nearly all of the characteristics described, a great majority tended to show up in both false and true allegations in varying degrees. For example, while rape myths were found to be present much more in false allegations, they were almost always also present in true allegations to a certain extent (Norton & Grant, 2008). This is perhaps best highlighted by the results of the study done by Rassin and Sleen (2005). While they noted that on average, false statements contained a higher number of characteristics from Hazelwood and Burgess' list than true allegations, almost none of the true allegations were totally devoid of all such characteristics. Similarly, in the study done by Marshall and Alison (2006), 33 out of 37 characteristics were present in both the false and the truthful statements they examined. This also highlights the importance of cautiously approaching models aimed at systematically ascertaining the truthfulness of rape allegations. The chance of false positives or false negatives is always present, and thus one characteristic in isolation can never be reliably used to judge the veracity of a statement.

Storytelling mechanisms. Different authors have reviewed how true and false allegations differ in the way their stories are constructed. One finding replicated across a number of the different studies was that true stories tended to contain more coded variables than false stories (Parker & Brown, 2000; Rassin & Sleen, 2005; Marshall & Alison, 2006; Zutter, Horselenberg, & Koppen, in press). In particular, Zutter, Horselenberg, and Koppen (in press) noted that their truthful allegations sample contained a more complex structure (i.e. not a clear starting or ending point to the story, jumping to different places in the story continuously) with more expansive details than their false allegations. They attributed this to the fact that people making false rape allegations tend to be careful to fabricate a story that can be told multiple times in a consistent manner, and so they have to be concise. False
statements also differ from truthful statements in the way the liars typically behave (Zutter, Horselenberg, & Koppen, in press). Zutter, Horselenberg, and Koppen gave the example that pseudo-victims were likely to give a detailed description of the nose of perpetrator when asked because they can fabricate it on the spot (57% of them did so). For the real victims however, only a slight minority was able to provide this when asked (7%). They attributed this to the fact that pseudo-victims are afraid that not answering the question would undermine their credibility, whereas true victims do not have this worry. Similar to the findings from Zutter, Horselenberg, and Koppen; Parker and Brown (2000) also note that truthful allegations were more unstructured, showed less emotional incongruirty with the person telling it, and mentioned more of the victim's experience as well as the offender's experience during the rape. However, this last finding was not reflected in subsequent research done by Rassin and Sleen (2005), who noted no difference in how often a victim or perpetrator's state of mind is mentioned between conditions. Because both studies had a number of methodological flaws associated with them (see previous section; Vrij, 2005) it is hard to accurately ascertain the validity of these different claims.

**Characteristics of the rape act.**

**Verbal utterances.** While talk from the perpetrator during sexual offenses was frequently present in true allegations, in false allegations this was rare to non-existent (Parker & Brown, 2000; Marshall & Alison, 2006; Hunt & Bull, 2012; Zutter, Horselenberg, & Koppen, in press). This outcome was found across all studies examined here and extended to different sorts of utterances. Marshall and Alison (2006) reported that while 39% of true allegations have perpetrators that reveal details about themselves (e.g., background story or hobbies), for false allegations this was only 3%. Similarly, they mention that in true allegations perpetrators could be found talking during sex relatively frequently (54%) in comparison with false allegations (3.3%). Verbal utterances can also take the form of requesting sex acts, making threats, and self-disclosure, which were all found in much higher proportion in true statements than in false ones (Hunt & Bull, 2012). One of the major topics Zutter, Horselenberg, and Koppen (in press) investigated was related to this,
CHARACTERISTICS OF SEXUAL OFFENSES

namely how pseudo-intimate behaviors occurred in false rape statements versus in true rape statements. Pseudo-intimacy describes specific behaviors done by the perpetrator, such as trying to appear gentle and caring during the rape to satisfy the perpetrator's need for intimacy (Zutter, Horselenberg, & Koppen, in press). Such behaviors were found much more in true than false allegations, with for example the perpetrator asking the victim if they were enjoying the sex in one-third of the real allegations but none of the false allegations (Zutter, Horselenberg, & Koppen, in press). Thus, any large number of verbal utterances from the perpetrator in a statement was indicative of a statement being true across all studies.

**Range of sexual acts.** A wide range of sexual acts being present in a statement was found to be highly indicative of a statement being truthful (Parker & Brown, 2000; Rassin and Sleen, 2005; Marshall & Alison, 2006; Hunt & Bull, 2012; Zutter, Horselenberg, & Koppen, in press). Among others, Parker and Brown (2000) mentioned that while 56% of the true allegations included anal rape or vaginal penetration with foreign objects, only 13% of the false allegations contained these. In Rassin and Sleen (2005), foreplay was described in 33% of the true cases but none of the false cases. Marshall and Alison (2006) reported that false statements usually only contained vaginal penetration, while true statements were much more likely to contain fellatio or cunnilingus as well. Hunt and Bull (2012) specified further on this, and from their data true accounts of rape tended to have deviant sexual activities with a median of 3, while false accounts average only a median of 1. Zutter, Horselenberg, and Koppen (in press) also reported fellatio occurring in 53% of true cases and zero of the false cases. The reason Zutter, Horselenberg, and Koppen provided for this wide of variety of sexual acts in true allegations, is that real victims usually want to minimize negative consequences. As such they tend to agree with unusual requests from the perpetrator to avoid punishment. Pseudo-victims however do not think to include these strange sexual acts because they draw on only their own sexual experiences.

**Forensic characteristics.**

**Criminal aspects.** Another finding replicated across different studies was that
criminal acts were much more prevalent in truthful allegations than false ones (Rassin and Sleen, 2005; Marshall & Alison, 2006; Hunt & Bull, 2012; Zutter, Horselenberg, & Koppen, in press). Concerning weaponry, Rassin and Sleen (2005) reported that 11% of the true allegations, but 0% of false allegations made mention of it. In Marshall and Alison (2006), true allegations contained mention of weaponry in 51% of the cases, but for false allegations this was only 13%. For Zutter, Horselenberg, and Koppen (in press), these percentages amounted to 40% for true allegations and 3% for false allegations, and for Hunt and Bull (2012) respectively 43% and 16%. Hunt and Bull also described that in true allegations more precautions are taken by perpetrators to restrict the victim. This includes acts of gagging, binding, or blindfolding the victim. Stealing items from the victim was also much more prevalent in true allegations. Marshall and Alison (2006) mentioned acts of stealing occurring in 27% of the true statements and 0% in the false ones. Hunt and Bull (2012) found that theft was 6.2 times more likely to occur in true statements. Lastly, Zutter, Horselenberg, and Koppen reported that stealing things from the victim was present in a much larger number of true than false allegations (33% versus 6%).

**Physical assault.** While criminal acts seem to consistently occur more in true allegations than false ones, for physical assault and injuries the outcomes are far less consistent when comparing different studies. Parker and Brown (2000) found no discernible difference pertaining to severity of injuries found between true and false allegations. Hunt and Bull (2012) however noted that victim injuries are more likely to be seen in genuine (45%) than false (24%) allegations. Norton and Grant (2008) described more violence occurring in true than false or mock statements, but could not verify this statistically with between-group analyses. Marshall and Alison (2006) reported 'multiple acts of violence' being described more in false (33%) than true (13%) statements. They also mentioned 'single acts of violence' to be more often occurring in true statements, but this last result was not found to be statistically significant. Zutter, Horselenberg, and Koppen (in press) found more physical violence in true statements than false statements. For example, pulling the victim's hair, unnecessarily hurting the victim during sex, holding the victim by the neck,
beating the victim, and kicking the victim were all found more in true statements than false ones. Zutter, Horselenberg, and Koppen explained this difference using the distinction between expressive and instrumental violence. They stated that, expressive violence refers to violence purely for the enjoyment of the perpetrator, while instrumental violence refers to violence used by the perpetrator to dominate the victim. True allegations tend to contain expressive violence while false allegations tend to only contain acts of instrumental violence. They also proposed that pseudo-victims are less keen on including injuries and acts of violence in their story because they are aware these can be forensically verified. There is however no data available to corroborate these ideas.

**Law enforcement.** Expectedly, allegations that are supported by physical forensic evidence were much more likely to be true than false (Parker & Brown, 2000). This was also noted in the study done by Rassin and Sleen (2005), where available physical evidence was the strongest discriminator of all characteristics between true and false statements (48% in true and 0% in false statements). Another interesting area concerns the role of the person making the report to the police. In Hunt and Bull (2012), false allegations were primarily reported to the police by the person themselves (70%). True allegations however, were much more likely to be reported to the police by someone else than the victim, such as a family member or a friend (68%). In McNamara, McDonald and Lawrence (2012), of all the false allegations, 73% were made by the person themselves instead of another party. However, Rassin and Sleen found no difference in this regard; in 67% of the true cases the allegation was done by the victim itself, and in false cases this number was almost the same. Concerning time-frame, Hunt and Bull (2012) noted that a case that was reported to the police the same day as the rape happened was 4.9 times more likely to be genuine. Cases reported after 24 hours were 7.7 times more likely to be false. Zutter, Horselenberg, and Koppen (in press) found something similar, with true allegations being reported sooner than false allegations. However, in their findings the difference was not statistically significant. Thus, while available forensic evidence is a good indicator, it is unclear whether the person making the report matters. Nor is it evident to what extent delayed reporting to the police by
the victim matters.

Miscellaneous. Some studies reported information about demographic characteristics of the victim, but overall these findings are too scattered and few in number to draw serious conclusions from. For instance, in Hunt and Bull (2012), people making false allegations were slightly more likely to be white and unemployed than people making true ones. Though statistically significant, the difference is minor enough to warrant caution when interpreting them. Parker and Brown (2000) noted that 69% of the pseudo-victims had some history of psychiatric illness compared to 13% of the genuine victims. However, in McNamara, McDonald and Lawrence (2012), only 26% of the pseudo-victims were found to suffer from mental illness. Since both studies use relatively small samples, it is hard to discern which percentage is a closer approximation of the truth. Zutter, Horselenberg, and Koppen (in press) did not find any statistical difference on the variables of age, ethnicity, or other related demographic characteristics. Finally, McNamara, McDonald, and Lawrence also noted that 76% of the pseudo-victims have some sort of marital problems or problems with their significant other.

Discussion

The research objective of this review was to investigate to what extent false allegations of sexual offenses differ from true allegations of sexual offenses in the unique characteristics they exhibit. From the findings reviewed here, a number of themes emerged. There were roughly three groups of behaviors that emerged more consistently in true allegations in nearly all studies. First, offenders are reported to talk significantly more in true allegations of sexual offenses than in false ones. In false ones they may sparsely say a word, whereas in true ones they often reveal much about themselves, their desires, and even their hobbies or personal preferences. Second, the range of sexual acts is much higher in true allegations than in false allegations. Not only do true allegations contain more sexual acts (e.g., foreplay, cunnilingus, fellatio, anal), the sexual acts themselves are also more extreme and unusual (e.g., using sexual objects, the perpetrator biting the victim during sex). In contrast, false allegations are much more likely to contain only vaginal penetration. Third,
criminal acts are much more likely to occur in true than false allegations. These include the offender carrying a weapon, the offender binding or gagging the victim, or the offender stealing money or goods from the victim. Pseudo-victims are not nearly as inclined to include these criminal acts in their story. Because all three of these themes (verbal behavior, range of sexual acts, and criminal behavior) seem to be absent in false allegations, this lends support to the idea that people making false allegations draw upon their own knowledge to create their fabricated story (see Zutter, Horselenberg, and Koppen (in press) and their ‘theory of fabricated rape’). When false allegations are made by people who have not experienced real rape, they must rely on their own imagination on how such a rape would occur. In doing so they forget to include components that happen during real sexual offenses.

Besides these three big themes, a number of other interesting findings also emerged from the literature review. For example, some of the studies reviewed here found that stories of true allegations contain more variables than false ones, meaning that they are more expansive in the content they provide. They also tend to be more unstructured, possibly because they are based on spontaneous recollection, instead of false stories that are carefully manufactured beforehand (Parker & Brown, 2000). However, these tendencies have a fair amount of subjectivity to them. It is hard to set objective perimeters to separate 'structured' from 'unstructured' stories, or 'stories with in-depth content containing many variables' from 'stories that are shallow'. Similarly, the extent to which emotionality of the victim plays a role is only tentatively explored in the studies reviewed here, and there is not enough data to say anything about this topic with certainty. This also holds true for injuries present on the victim, or aggressive physical acts by the perpetrator towards the victim. Though many authors reviewed here discuss these topics, they do so in different ways, and their findings are overall too different from each other to generalize to one single guideline. Finally, for behaviors relating to law enforcement (e.g., the time it took the victim to report to the police, whether the victim or someone else reported the sexual offense to the police) results are not uniform enough to generalize. For demographic characteristics of the victim,
this is even less so. For instance, the extent to which pseudo-victims have a history of psychiatric illness ranged from 26% to 69% between two studies examined. One study reported pseudo-victims being on average more Caucasian than real victims, but another study did not. In these areas, there is not enough research done yet to say with certainty which comes closer to the truth.

Earlier on, subjective lists that were created based on personal experience in the field were referenced (see for example Appendix A). Because these are not based on empirical research, it is interesting to examine how they contrast with findings from empirical research. First, while verbal communication is an important indicator of true allegations in empirical research, this is not mentioned in any of the personal-experience lists. This is also true for criminal acts; the absence or presence of these is not dealt with in these personal-experience lists. Lack of foreplay or other deviant acts is however referenced (Appendix A, point 11). Some of the points on the personal-experience lists are in direct opposition with results from empirical literature. For example, the Dietz and Hazelwood red-flag list (Turvey & Petherick, 2009) mentions that stories that are bizarre or sensational are more likely to be false. While they refer to sensational allegations influenced by the media, it should be mentioned that real rape allegations are more likely to contain bizarre sexual acts or extraneous details. Similarly, in many of the personal-experience lists it is mentioned that people making false allegations steer away from talking about the offender and cannot describe the offender. This is the reverse of what Zutter, Horselenberg, and Koppen (in press) present in their study. They found that pseudo-victims are more likely to give descriptions of what the perpetrator looks like, because they are able to fabricate these details on the spot. Real victims however do not do this, because they feel no need to lie. As such, a statement containing a description of the offender would not point towards it being a true allegation but rather a false one. Lastly, many other characteristics from these experience-based lists cannot be falsified nor corroborated by the empirical research reviewed here. This is because there is simply not enough information to say for sure whether they hold up or not. Emotionality of a victim during testimony and history of
psychiatric illness especially cannot be empirically verified by the current literature review, but they are accentuated in several personal-experience lists (e.g., Appendix A, point 1 and 2; Turvey & Petherick, 2009). Because these are very sensitive in nature, they should be approached with caution as well.

The review of methodology from different studies in this paper shows a high variability on how studying false accusations is approached in different studies. This was earlier noted by Saunders (2012), and it pertains to the definitions used of both false and true allegations. A number of authors did not commit to very precise criteria on how they selected false accusations, and a few of them failed to mention any criteria at all on which they selected their true statements. It is very conceivable that the selection criteria used could contribute significantly to the final results that were found in individual studies. For example, some papers included allegations in their false condition if the victim was prosecuted for wasting police time. As described earlier, a police prosecution on its own is not nearly enough to guarantee that an allegation was indeed fabricated. There were also shortcomings in several studies in how they approached their general design. Several did not code all of their statements blindly, and one study did not code any of their statements blindly. Some studies also chose to use mock statements exclusively instead of taking from a real field sample of false statements. Because of all these differences in approaches, outcomes from studies remain hard to be accurately compared.

Though the current paper managed to describe some trends when it comes to characteristics of false allegations of sexual offenses, there are limitations that prohibit large-scale generalizations of the current review. For one, the number of studies that were reviewed, eight, is rather small. The research on false rape allegations is still developing, but this also means that the current research done is still preliminary. For this reason there was only a small pool of studies to select from. Associated with this is the fact that the different studies handled here use different designs and definitions. This makes the comparisons between study outcomes in some situations difficult, and can contaminate the combined aggregate of outcomes described here. For example, some studies relied on real field cases
of false allegations, while others used exclusively mock statements generated by participants. Because these are two very different ways of obtaining sample statements, differences in outcomes could certainly be attributed to this. Notwithstanding these limitations, the outcomes described throughout this literature review point to a number of themes that present themselves with some consistency. For example, there was a marked absence of verbal talk from the perpetrator in false allegations that could be noted throughout nearly all the studies reviewed here. Indicators such as these give rise to the thought that with the help of future studies, a more complete and reliable 'red-flag' list could ultimately be mapped. With the rapid developments in scientific research on false allegations of sexual offenses in recent years, more definitive lists are likely to take shape in the very near future.
References


Nierop, N. M., & Eshof, P. *Misbruik, Misleiding en Misverstanden*. Retrieved from
http://www.ouderverstoting.nl/artikelen/onderzoeksverslag%20misbruik%20misleiding
%20en%20misverstanden_tcm35-423342.pdf

Norton, R., & Grant, T. (2008). Rape myth in true and false rape allegations. *Psychology,


Rassin, E., & Sleen, J., Van Der. (2005). Characteristics of true versus false allegations of
sexual offences. *Psychological Reports PR, 97*(6), 589. doi:10.2466/pr0.97.6.589-598

In Hazelwood & Burgess (Eds.), *Practical Aspects of Rape Investigation* (221-250).
Boca Raton: CRC Press.

65*(01), 128. doi:10.1017/s0008197306007069

Saunders, C. L. (2012). The truth, the half-truth, and nothing like the truth: reconceptualizing
false allegations of rape. *British Journal of Criminology, 52*(6), 1152-1171.
doi:10.1093/bjc/azs036

*Psychomethods in criminal investigations and evidence*. New YorK: Springer.


Criminological and Legal Psychology, 20*, 51-57.

Vrij, A. (2000). *Detecting lies and deceit: the psychology of lying and implications for
professional practice*. Chichester, UK: Wiley.


### Table 1

**Studies Used as a Resource**

<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Authors</th>
<th>Sample Choice</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>Detection of deception: Statement Validity Analysis as a Means of Determining Truthfulness or Falsity of Rape Allegations</td>
<td>Parker &amp; Brown</td>
<td>Field cases of true, false and unsubstantiated allegations (n=43) selected by the researchers themselves</td>
<td>SVA(^b) has practical value in differentiating between false and true allegations, but less so for unsubstantiated allegations. VC(^c) has added value over CBCA(^d).</td>
</tr>
<tr>
<td>2005</td>
<td>Characteristics of True versus False Allegations of Sexual Offences</td>
<td>Rassin &amp; Sleen</td>
<td>Field cases of true and false allegations (n=41) selected by field agents who filled in a survey</td>
<td>There are certain criteria present in false allegations that allow to differentiate them from true allegations to a certain extent.</td>
</tr>
<tr>
<td>2006</td>
<td>Structural Behavioural Analysis as a Basis for Discriminating between Genuine and Simulated Rape Allegations</td>
<td>Marshall &amp; Alison</td>
<td>Field cases of true allegations (n=152) and mock statements written by female participants (n=30)</td>
<td>True allegations report a larger number of behaviors than simulated ones; pseudo-intimate behavior is more likely to be found in true allegations.</td>
</tr>
<tr>
<td>2008</td>
<td>Rape Myth in True and False Rape Allegations</td>
<td>Norton &amp; Grant</td>
<td>Field cases of true (n=19) and false (n=19) allegations selected by the researchers themselves in addition to mock statements generated (n=19)</td>
<td>False allegations generally contain more rape myths than true allegations do.</td>
</tr>
<tr>
<td>2008</td>
<td>False Rape Allegations(^a)</td>
<td>Hazelwood &amp; Burgess</td>
<td>(Inductive)</td>
<td>(List of proposed set of red flags)</td>
</tr>
<tr>
<td>2012</td>
<td>Differentiating Genuine and</td>
<td>Hunt &amp; Bull</td>
<td>Field cases of true allegations</td>
<td>A number of factors can be used to</td>
</tr>
</tbody>
</table>
### CHARACTERISTICS OF SEXUAL OFFENSES

<table>
<thead>
<tr>
<th>Year</th>
<th>Study Title</th>
<th>Authors</th>
<th>Data Description</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>Characteristics of False Allegation Adult Crimes</td>
<td>McNamara, McDonald &amp; Lawrence</td>
<td>Field cases of false allegations (n=30) selected by the researchers themselves</td>
<td>False allegations share common factors such as personality characteristics and motivations.</td>
</tr>
<tr>
<td>2014</td>
<td>Filing False Vice Reports: Distinguishing True from False Allegations of Rape</td>
<td>Zutter, Horselenberg, &amp; Koppen</td>
<td>Field cases of true allegations (n=30) and mock statements written by female participants (n=35)</td>
<td>Fabricated stories of rape have different characteristics than true ones, mostly concerning pseudo-intimate behavior.</td>
</tr>
</tbody>
</table>

---

*a* Hazelwood & Burgess' list of characteristics are not based on empirical research.

*b* Statement Validity Analysis.

*c* Validity Checklist.

*d* Criteria Based Content Analysis.
### Table 2

**Definitions and Samples Used by Researchers**

<table>
<thead>
<tr>
<th>Study</th>
<th>Statements Analyzed <em>(n)</em></th>
<th>Criteria of Data Selection Given</th>
<th>Criteria False Allegation Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parker &amp; Brown (2000)</td>
<td>16 12 -</td>
<td>Yes Yes -</td>
<td>Victim admits allegation was faked, physical evidence contradicts allegation, allegation was result of delusion or substance-abuse, AND/OR allegation was made for specific inappropriate reasons.</td>
</tr>
<tr>
<td>Rassin &amp; Sleen (2005)</td>
<td>27 14 -</td>
<td>Yes Yes -</td>
<td>Any allegation judged to be false by an individual police agent AND having resulted in a conviction for the victim.</td>
</tr>
<tr>
<td>Marshall &amp; Alison (2006)</td>
<td>152 - 30</td>
<td>No - No</td>
<td>(Only mock statements used)</td>
</tr>
<tr>
<td>Norton &amp; Grant (2008)</td>
<td>19 19 19</td>
<td>No Yes Yes</td>
<td>Allegation was withdrawn as false by victim AND/OR police filed prosecution against victim.</td>
</tr>
<tr>
<td>Hunt &amp; Bull (2012)</td>
<td>160 80 -</td>
<td>Yes Yes -</td>
<td>Allegation was retracted by victim with admittance of falsity AND/OR victim was charged with wasting police time. Additional requirement was that allegation wasn't made due to confusion on the victim's part, that the encounter wasn't found to be consensual AND that there was only a single offender identified.</td>
</tr>
<tr>
<td>McNamara, McDonald &amp; Lawrence (2012)</td>
<td>- 30 -</td>
<td>- Yes -</td>
<td>Allegations classified by FBI as false AND the allegation has conviction attached to it AND later corroborated as false subjectively McNamara et al.</td>
</tr>
</tbody>
</table>
Zutter, Horsemelberg, & Koppen (2014)

30 - 35 Yes - Yes (Only mock statements used)


*Refers to whether study provided explicit information about how mock statements were generated.
**Table 3**

**Study Methodologies**

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Coded by</th>
<th>Scoring</th>
<th>Characteristics Investigated</th>
<th>Origin Characteristics Investigated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parker &amp; Brown (2000)</td>
<td>43</td>
<td>Researchers</td>
<td>Dichotomous</td>
<td>31</td>
<td>Derived from CBCA(^a), SVA(^b) and VC(^c).</td>
</tr>
<tr>
<td>Rassin &amp; Sleen (2005)</td>
<td>41</td>
<td>Police officers</td>
<td>Dichotomous plus 'not sure' option</td>
<td>43</td>
<td>Derived from CBCA and RM(^d). Also list of non-verbal characteristics included.</td>
</tr>
<tr>
<td>Hunt &amp; Bull (2012)</td>
<td>240</td>
<td>SCAS(^a)</td>
<td>Dichotomous</td>
<td>62</td>
<td>Descriptive characteristics provided by the SCAS(^a).</td>
</tr>
<tr>
<td>Zutter, Horselenberg, &amp; Koppen (2014)</td>
<td>65</td>
<td>Independent evaluators</td>
<td>Dichotomous</td>
<td>187</td>
<td>Derived from investigating literature on true rape allegations in addition to characteristics that were deemed relevant during the study itself.</td>
</tr>
</tbody>
</table>

*Note. Hazelwood & Burgess (2009) non-empirical study excluded from current list. McNamara, McDonald & Lawrence (2012) study also excluded because it only investigates common factors relating to rape crimes.*

\(^a\) Criteria Based Content Analysis.

\(^b\) Statement Validity Analysis.

\(^c\) Validity Checklist.
CHARACTERISTICS OF SEXUAL OFFENSES

d Reality Monitoring.
* Serious Crime Analysis Section of the United Kingdom's National Crime Agency.
Figure 1. Data Selection Process.

*Note.* All downward arrows denote that an article's reference list was used to obtain a subsequent article. The only exception to this is are the arrows from the top-left two boxes; these denote that Google Scholar produced 56,600 hits from which four articles were ultimately selected.
Appendix A

Red-flag list by Hazelwood and Burgess (2009)

Attention points that signal that an allegation made by a victim may be fabricated (derived from Hazelwood and Burgess, 2009).

1. Victim has underlying emotional problems, e.g., Munchhausen or other dramatic illness.
2. Victim has history of psychiatric illness.
3. Victim repeatedly changing story to make it more believable.
4. Victim continually recalls additional details.
5. Victim is motivated by need for attention and sympathy, anger and revenge, or alibi.
6. Victim is primarily female, Caucasian, age 21-30 years.
7. Victim did not immediately chose to report the rape but instead delayed for a considerable time.
8. The report was not made by the victim itself, but someone else who acted on behalf of the victim.
9. Victim prefers talking about force, resistance, and injuries during the crime.
   Victim avoids talking about offender or facts of the crime.
10. Victim reports a large and overpowering offender.
11. Victim only report sexual acts in which they would normally engage (e.g., foreplay is absent).
12. There is an absence of forensic evidence to support the allegation of the victim.
13. Victim presents with superficial injuries that occurred during self-defense.
14. Victim presents with ripped clothing.
15. Victim has a previous history of reporting rape or assaults.
16. Victim has difficulty describing the offender.