Legal and Criminological Psychology (2019)

The British Psychological Society

www.wileyonlinelibrary.com

Verify the scene, report the symptoms: Testing the Verifiability Approach and SRSI in the detection of fabricated PTSD claims

Irena Boskovic[1,2]*, Pauline Dibbets[1], Glynis Bogaard[1], Lorraine Hope[2], Marko Jelicic[1] and Robin Orthey[1,2]

[1]Maastricht University, The Netherlands,
[2]University of Portsmouth, UK

Purpose. In order to effectively feign post-traumatic stress disorder, a person needs to confabulate an exposure narrative and to fabricate symptoms of high distress. The Verifiability Approach (VA) is a lie-detection method based on the notion that truth tellers' narratives include more verifiable (checkable) information than liars' narratives. The Self-Report Symptom Inventory (SRSI) is a measure of over-reporting, and it includes genuine symptoms and pseudosymptoms that are likely to be endorsed in fabricated symptom reports. In this study, we examined whether the VA can help discriminate the fabricated exposure narratives, and whether the SRSI can aid screening for symptom over-reporting.

Method. One group of participants (truth tellers) witnessed a vehicle crash scene using the Virtual Reality paradigm (n = 22), while the other group (feigners) was instructed to fabricate such an experience (n = 46). All the participants wrote the exposure narratives and completed the SRSI.

Results. Feigners produced non-verifiable (vague) and lengthier narratives than truth tellers, who reported a higher proportion of checkable information. Regarding the symptom reports, feigners endorsed more of trauma-related genuine symptoms and pseudosymptoms than truth tellers.

Conclusion. The non-verifiable details and the proportion of verifiable details, together with the SRSI subscales, can assist explaining the reporting strategies of those feigning negative exposures.

Post-traumatic stress disorder (PTSD) was officially introduced in 1980 (DSM-3; American Psychiatric Association, 1980) and pertains to a broad range of psychological disturbances as a consequence of a trauma experience (Resnick, West, & Payne, 2008). Originally, this type of diagnosis was associated with war veterans (Adamou & Hale, 2003), but it soon became clear that combat exposure is not the only possible trigger for PTSD. In fact, any negative life event can be experienced as traumatic (Resnik *et al.*, 2008), but not every traumatic experience results in PTSD (Bonanno, 2005; Hall, Hall, & Chapman, 2006).

**Correspondence should be addressed to Irena Boskovic, Forensic Psychology Section, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands (email: irena.boskovic@maastrichtuniversity.nl).*

DOI:10.1111/lcrp.12149

Prevalence of traumatic exposure in general public is up to 70%, yet the prevalence of PTSD is approximately 10% (Young, 2016). The highest prevalence of PTSD is among the victims of sexual abuse (up to 80%; Hall *et al.*, 2006) and war veterans (up to 58%; Guriel & Fremouw, 2003), compared with a general population (up to 15%; Hall & Hall, 2007).

As currently described in DSM 5 (American Psychiatric Association, 2013), the diagnosis of PTSD includes eight different criteria (A–H). In order to receive the PTSD diagnosis, a person must have (1) a traumatic experience (A), (2) symptoms that were caused by the traumatic experience: re-experiencing (B), avoidance (C), negative mood and cognition (D), and arousal (E), which (3) last at least 1 month (F), present a severe obstacle for daily functioning (G), and are not a product of medication or alcohol/drug abuse (H). All of these criteria are based on self-report, meaning that a person can easily over-report of fully fabricate his/her trauma-related complaints if incentivized to do so (i.e., malinger). Malingering presents a deliberate fabrication of symptoms in order to gain potential external benefits (American Psychiatric Association, 2013). Indeed, there are many potential financial (e.g., compensation) and legal (e.g., reduced criminal responsibility) benefits available for people suffering from PTSD (Knoll & Resnick, 2006; see also Rassin, Boskovic, & Merckelbach, 2018). Thus, due to the self-evident nature of its key symptoms, incentives surrounding this diagnosis, and professionals who admit their uncertainty in recognizing malingering (Cohen & Appelbaum, 2016), PTSD is one of the most easily and the most frequently feigned psychological disorders (Guriel & Fremouw, 2003; Resnik *et al.*, 2008; Resnick, West, & Wooley, 2018). The estimated prevalence of fabricated PTSD is above 30% (Freeman, Powell, & Kimbrell, 2008; Lees-Haley, 1997).

In order to successfully feign suffering from PTSD, a person needs to provide a convincing story about the exposure (A criterion), and to adequately report symptoms (B–E). Therefore, the detection of such cases must comprise assessment of the validity of the exposure statements, as well as the validity of reported symptoms.

Verifying exposure narratives

Previous research on verbal credibility assessment suggests that truthful narratives contain overall more details than deceptive reports (Johnson, 2006), as well as more specific information, such as time and location (Porter, Peace, & Emmett, 2007). Furthermore, truthful narratives are more emotionally charged and judged as more plausible than narratives about a non-actual event (Peace & Porter, 2011). Currently, there are many methods of credibility assessment that are based on the assumption that truthful narratives include details of certain quality. For example, the most commonly used tool, the Criteria Based Content Analysis (CBCA, Steller & Koehnken, 1989), includes 19 different criteria organized around five major domains, such as general characteristics, specific content, specific features of the content, content related to motivation, and elements of aggression. The presence of each CBCA criteria indicates truthfulness of the statement (Amado, Arce, Farina, & Vilariño, 2016). Two meta-analyses performed on the utility of the CBCA (Amado *et al.*, 2016; Hauch, Sporer, Masip, & Blandón-Gitlin, 2017) provided overall favourable findings. However, findings of Bogaard, Meijer, Vrij, Broers, and Merckelbach (2013) suggested that the CBCA criteria are vulnerable to contextual bias (Bogaard *et al.*, 2013). In other words, the interpretation of the statements and scoring on both CBCA and RM criteria can considerably change if coders are presented with any additional information, such as

evaluee's personality characteristic, which limits the reliability of these methods (Bogaard *et al.*, 2013).

Alternatively, a recently developed technique, the Verifiability Approach (VA; Nahari, Vrij, & Fisher, 2014a), has shown some success in facilitating the detection of fabricated statements (e.g., Harvey, Vrij, Nahari, & Ludwig, 2017; Jupe, Leal, Vrij, & Nahari, 2017; Nahari, 2018). According to this approach, people who are telling truth include more details that are verifiable (i.e., checkable in the real world) in their statements than people who fabricate their accounts. Liars cannot provide verifiable information without the risk of being caught; therefore, they opt to report more information unavailable for external validation (i.e., non-verifiable details; Nahari, 2018). For instance, reporting 'I had coffee with my friend, Brianna' would be considered as verifiable due to the inclusion of an identifiable person that could (in principle) confirm provided information. However, a person who wants to deceive the interviewer would probably keep the statement vague, such as 'I was sitting on a bench alone' (see also Nahari *et al.*, 2014a).

Furthermore, the statements of truth tellers and liars differ even more when the statement-givers are informed, using an 'Information Protocol', that the number of reported verifiable details will serve as an indicator of their veracity. While this instruction elicits more verifiable details from truth tellers, liars are unable to follow through without exposing their lies, thus their reports remain mostly non-verifiable (Harvey *et al.*, 2017). This implies that disclosing the detection strategy actually facilitates the success of the VA (Nahari, Vrij, & Fisher, 2014b).

In previous studies (Boskovic, Bogaard, Merckelbach, Vrij, & Hope, 2017; Boskovic, Gallardo, Vrij, Hope, & Merckelbach, 2018), researchers aimed to test the utility of the VA outside of its original lie-detection context. They investigated whether the VA is effective in the context of symptom reporting and if the use of this approach would enhance the detection of fabricated complaints (i.e., headaches). However, the verifiable details did not aid the detection of fabricated symptoms complaints because people with genuine complaints and malingerers reported similar amounts of verifiable information. Rather, the increased amount of non-verifiable information appeared to be a cue for deceptive symptom reports. Those who malingered provided notably more non-verifiable details, also making their reports significantly lengthier when compared with genuine accounts. However, both of these trademarks, non-verifiable details and length of reports, were lost once the Information Protocol was applied. It is likely that providing participants with the clear instructions on which type of details would indicate their veracity subsequently influenced both groups of participants to write similar statements. This finding goes against presented results from the lie-detection literature.

Although the creators of the VA claim that the utility of this approach is limited in the context of malingering due to the unverifiable status of symptoms (Nahari, 2018), we tend to disagree with this reasoning. Symptoms, such as pain or headache, usually lead to behaviours open for verification (e.g., going to the doctor, taking medication etc.; Boskovic *et al.*, 2018). However, a critical difference between the lie-detection and malingering contexts might be the feasibility of providing false verifiable details. For instance, in typical lie-detection contexts, where people report about an external event likely witnessed by others (e.g., a fight), a production of misleading verifiable information might be a riskier choice than when a person is describing his/her internal state (e.g., pain). Therefore, feigners might have an easier task in providing checkable information when presented with the Information Protocol (Boskovic *et al.*, 2017).

Nevertheless, in case of malingered PTSD, the contexts of lie detection and malingering overlap. Hence, although the detection of fabricated exposure narratives

resembles the lie detection, and therefore, similar methods may apply, the symptom veracity assessment requires a malingering specific approach.

Symptom endorsement in feigned PTSD

Assessing the veracity of the symptoms using the verbal assessment and content analysis (e.g., Akehurst *et al.*, 2015) has not been a dominant approach within the field of symptom validity assessment. Rather, most research focusses on the development of symptom scales that test for over-reporting and exaggerated complaints known as the Symptom Validity Tests (SVTs). The 'over-the-top' way of responding to symptoms inventories was shown to be a trademark of fabricated PTSD reports (e.g., Hall & Hall, 2007; Peace & Masliuk, 2011; Tracy & Rix, 2017). Thus, the idea behind the SVTs is that those who fabricate their symptoms will overendorse items on symptom inventories, even if they contain bizarre/implausible complaints (Merckelbach & Smith, 2003). One such measure is the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997). The SIMS includes 75 atypical items, which are not likely to be true even for genuine patients. For example, items such as 'Sometimes when writing a phone number, I notice that the numbers come out backwards even though I don't mean to do it'. People who endorse such items above a proposed threshold score (e.g., 16, Smith & Burger, 1997) are believed to be over-reporting, and thus, their reports should be viewed with some scepticism. Furthermore, recently a study showed that the tendency to overendorse SIMS items is even more pronounced among people fabricating direct traumatic exposure than indirect aversive experience (Szogi & Sullivan, 2018). However, these atypical items are often obvious to the examinees, a shortcoming that diminishes the reliability of SIMS. Additionally, the SIMS includes items pertaining to complaints such as amnesia, psychosis, and low intelligence, which are frequent within the criminal context, but not in civil medico-legal setting (Merten, Merckelbach, Giger, & Stevens, 2016).

A recently developed measure of over-reporting, the Self-Report Symptom Inventory (SRSI; Merten *et al.*, 2016), may be a better alternative for two reasons. First, the SRSI includes a mix of genuine symptoms and pseudosymptom items divided into two main scales. A typical item for the genuine symptom scale is 'I have nightmares about things that happened to me', while for the pseudosymptom scale, a typical example is 'I can't remember what happened to me, but I constantly dream about it'. Because both types of items are distributed throughout the questionnaire, it is more difficult for an examinee to recognize the real aim of the assessment. Second, the SRSI genuine symptoms scale of symptoms related to (1) cognitive complaints; (2) depression; (3) pain; (4) somatic problems; and (5) anxiety/PTSD. Meanwhile, the pseudosymptoms scale taps into (1) cognitive/memory complaints; (2) neurological motor; (3) neurological sensory issues; (4) pain; and (5) anxiety/depression/PTSD. Thus, each of the two main (genuine symptoms and pseudosymptoms) scales includes five subscales describing the most prevalent complaints within civil medico-legal context (Merten *et al.*, 2016). Research to date suggests that the SRSI is a promising tool for detection of over-reporting, with rates of detection above 77% (Merten *et al.*, 2016). However, the utility of the subscales alone, such as genuine and pseudosymptoms of anxiety-related issues, has not been examined so far.

The current research

The aim of this study was to investigate the quality of narratives and symptom reports between people instructed to fabricate an aversive experience and its consequences and people who were exposed to an aversive experimental manipulation. In this study, we included two groups of participants. One group participated in a (separate) study conducted in the clinical psychology department, in which they were exposed to a Virtual Reality (VR) scene of a vehicle crash as a method of inducing trauma-like symptoms. The second group, participating only in the current study, was not exposed to the VR scene of the vehicle crash, but was instructed to act as if they had witnessed the same VR scene. We then asked all participants to write a narrative about the scene as if they had witnessed it in person. We anticipated that truth tellers would report more verifiable details while feigners would produce more non-verifiable information. Additionally, we asked participants to report their distress caused by the witnessed scene by using the anxiety/PTSD-related subscales of the SRSI. We expected that truth tellers would endorse significantly fewer symptoms of both the genuine symptoms and pseudosymptoms than feigners.

Method

Sample

We initially recruited 102 participants; 21 participants were subsequently excluded (see below), and a further 13 participants withdrew from the study. The final sample consisted of 67 students (77.6% females, Mean age = 22.5, *SD* = 3.83). Of this sample, 22 (truth telling condition) participants were recruited from a separate clinical study focused on inducing PTSD-like symptoms using the VR paradigm. On average, participants joined the current study 6 days (*SD* = 8.00; range 0–25) after the VR exposure. After careful inspection, one participant was additionally excluded because of the extremely long delay after the exposure (80 days), which led to a total of 21 truth tellers. The rest of the participants (n = 46) were a newly collected sample who did not witness the VR scene. Participants in this second group were instructed to feign the experience of witnessing the VR scene as if they experienced it 7–15 days ago (feigning condition). Both studies were approved by the standing ethical committee.

Measures and materials

Jellinek-PTSD Screening Questionnaire (JPSQ, van Dam, Ehring, Vedel, & Emmelkamp, 2013)

Participants in the VR study (truth tellers) were pre-screened for putative PTSD symptoms using the JPSQ. The JPSQ is a short self-report questionnaire and consists of four questions that can be answered with either yes or no. The score is the total sum of positive answers (range 0–4). Only participants with a score of 0 were allowed to participate in the current study. The JPSQ has shown to have high sensitivity (.87) and specificity (.75) (van Dam *et al.*, 2013).

PTSD checklist for DSM-5 (PCL-5; Weathers et al.*, 2013)*

Participants in the feigning condition were pre-screened for any current high distress using the PCL-5. To adhere to ethical restrictions, we excluded any participant (n = 21) who responded with an answer equal to or higher than 3 (*Quite a bit*). The *Cronbach's alpha* of PCL-5, based on the entire initial feigning sample (N = 80), was .94.

Self-Report Symptom Inventory (SRSI; Merten et al.*, 2016)*

The SRSI includes 107 items belonging to two superordinate scales: One that pertains to plausible symptoms and one that pertains to pseudosymptoms. Furthermore, each of the two main scales includes five subscales tapping into different issues. In this study, we focused only on the two subscales, pertaining to the symptoms expected to follow an exposure to an aversive event (i.e., anxiety and PTSD-like symptoms). Thus, we included 22 items, two control items, and 20 items describing the anxiety/PTSD-related complaints. Half of the items present genuine anxiety/PTSD-like symptoms (*Cronbach's alpha* = .88), and the other half tap into exaggerated complaints (e.g., pseudosymptoms; *Cronbach's alpha* = .75). For each symptom, participants indicate whether or not they suffer from it (False/True); thus, the maximum score per scale (genuine and pseudosymptoms) is 10. Although these subscales include a mix of symptoms, among which is also depression, we will address them as 'anxiety/PTSD subscales' for the sake of clarity.

Newspaper article

Before witnessing the VR scene, participants in the clinical study (truth tellers) were given a newspaper article. Modelled on newspaper reports of actual crashes, the article describes the accident, includes information about the victims, and speculates about possible causes of the crash. The article included a picture of the crash scene. The article was 100 words in length and was presented in both English and Dutch (see Appendix). We also incorporated the article into our study design.

Virtual reality scenario

In order to elicit PTSD-like symptoms among the truth tellers group, participants were exposed to a Virtual Reality scenario depicting a vehicle crash between a car and a train. The VR simulation was created and run in Unity 5 on an Oculus Rift DK2 (Development Kit 2). The programming language used was C#, and the graphics were created in Blender 3D. The VR scenario was shot from the first-person perspective, and a participant could look left and right within the immersive scene. The scene is as follows: the crash involving a train and a car has already happened. The participant is on a bike, and in front of her/him, there is a fence and three parked cars. Two people, a woman and a man, are standing next to their car, panicking and trying to call for help. The train is not moving, and a fire starts in the car. The victims, a man and a baby, are loudly screaming from the burning car. No train passengers are visible. The crossing lights and warning sirens are on, and it is raining. The car is consumed in flames, and the victims trapped in the car fall silent. In the last minute of the scene, police and ambulance sirens can be heard approaching.

Procedure

Participants from a separate clinical study served as our truthful comparison group (truth telling condition). This study was concerned with inducing PTSD-like symptoms. Exclusion criteria were a non-zero score on the JPSQ and having witnessed or having been involved in a car crash. For participants in the feigning condition, the exclusion criteria were a PCL-5 score (equal to or higher than three), and having witnessed or having been involved in a car crash, but no participants reported having that experience.

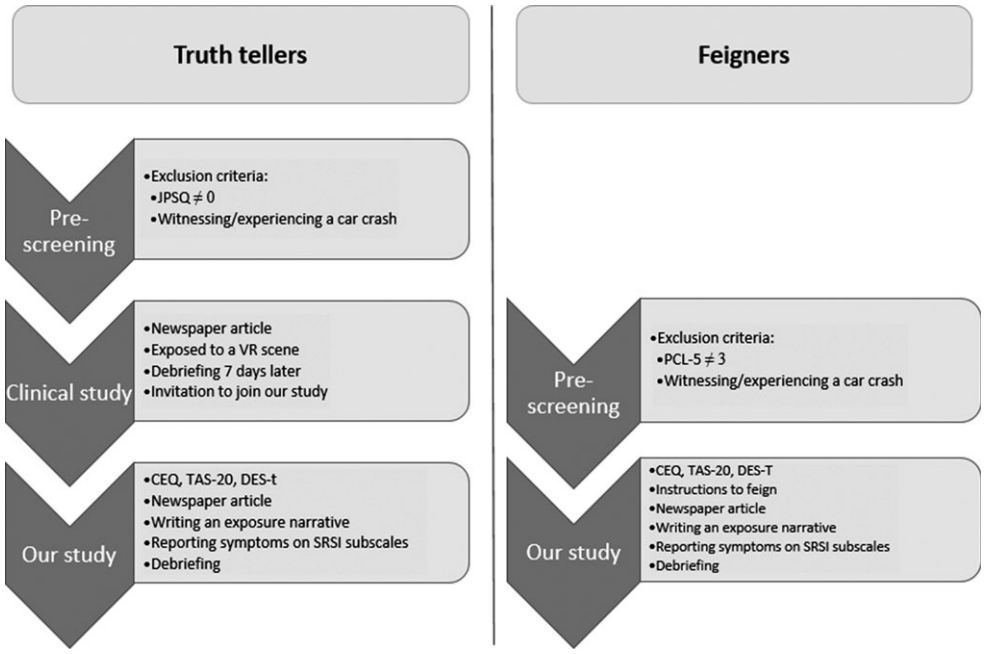

Figure 1. Study procedure.

All of the participants in both conditions received a Qualtrics link, which allowed them to access the study program. After responding to demographic questions (age, gender, and student status), truth tellers were also asked to indicate the delay (in days) since exposure to the VR scene. Then, all participants were given a battery of self-report questionnaires[1] followed by the two main tasks: (1) to write the narrative and (2) to fill out the SRSI subscales. Prior to writing the narrative, participants in both conditions were instructed to read the newspaper article. Truth tellers were instructed to report about the VR scene in as much detail as possible. Participants in the feigning condition received instructions to feign witnessing the crash scene and were asked to imagine receiving financial compensation from the train company if they provided a convincing witness account. Research showed that, besides compensation, revenge is a strong motive for feigning (Peace & Masliuk, 2011); thus, participants were encouraged to think about any negative experience they previously had with a national train company. Participants could opt to write their narratives in either English or in Dutch (see Figure 1). In the truth telling condition, 14 participants wrote the narratives in Dutch (66%), whereas 13 participants opted for Dutch in malingering group (20%). In order to motivate all the participants to write detailed narratives, we informed them that participants who provided detailed accounts would be entered into a prize draw for €20.

Coding using verifiability approach (VA; Nahari & Vrij, 2014)

According to the VA, details (i.e., piece of information) of statements can be coded as verifiable or non-verifiable. In the current study, for a detail to be coded as verifiable, it had to

[1] *For the sake of clarity, hereby we will only focus on the narrative and the SRSI findings.*

meet one of the following criteria: (1) to be documented (recorded, or left actual or virtual trace); (2) was related to occurrences that were carried out together with (an) other identified person(s) rather than alone or with a stranger who could not easily be traced; or (3) pertained to something that was witnessed by an other identified person(s) (Nahari & Vrij, 2014). Details that did not fulfil these requirements were labelled as non-verifiable. Coding was performed by two coders, both blind to the conditions. The primary coder coded all the statements, while the second coded a randomly selected sample of 20% (n = 18) of the statements. The details were coded in four different categories[2] : (1) Verifiable details from the newspapers (VNP; 'There were *two victims*, a *man* and *a baby*'), (2) New verifiable details (V; 'A *man* and a *woman* were *calling the ambulance*'), (3) Non-verifiable from the newspapers (NNP; 'Potential *cause* of the *crash* was a *distracted car driver*'), and (4) New non-verifiable details (NV; 'I *thought* I would faint').

The inter-rater reliabilities between coders were calculated using the intraclass correlation coefficients (ICC), which indicated a very good agreement for VNP details (ICC = .88, 95% CI [.69–.96]) and for V details (ICC = .86, 95% CI [.33–.96]), as well as agreement for the NNV details (ICC = .78, 95% CI [.40–.92]) and excellent agreement for NV details (ICC = .91, 95% CI [.52–.97]).

Statistical approach

The differences in the narratives' quality and symptoms reports between truth tellers and feigners were calculated using the *Welch's t-tests*, due to unequal sample size between groups. For the effect size of our findings, we report *Cohen's d*. To further examine whether the verifiable, non-verifiable details, and the SRSI subscales aid the prediction of the group membership (truth tellers vs. feigners), we performed the *Discriminative function analyses* and calculated the *Areas Under the Curve (AUCs)*.

Results

Exit questions

Participants reported moderate motivation in writing the convincing narratives on a five-point Likert scale (anchors being: 1 = 'Not motivated at all'; 5 = 'Extremely motivated') (M = 3.73, SD = 1.10), and reporting symptoms (M = 3.82, SD = .88).[3]

Narrative reports

Participants of the feigning group provided significantly longer narrative reports than truth tellers, *Welch's* $t(55.77) = 2.03, p = .047, d = 0.49$. Looking into the frequency of the four categories of details, we observed that 98.5% reported at least one verifiable detail from the newspapers, 100% reported new verifiable details, only 15% reported non-verifiable details from the newspapers, and 95.5% added new non-verifiable information into their reports. The feigners and truth tellers differed in the total number of non-verifiable details (NNP + NV), *Welch's* $t(51.05) = 2.64$, $p = .020$, *Cohen's* $d = 0.65$, which might be a

[2] *Underlined parts of the statements were coded as one detail. For instance, the specification of thoughts should not be coded as new details, rather only the indication that a person was thinking something (that did not actually happen) would be coded as one non-verifiable detail.*

[3] *We re-conducted all the main analyses excluding the three participants who reported very low motivation; however, our results did not significantly differ. Therefore, we retained them in the data.*

Table 1. Comparison of means and standard deviations between the truth tellers and feigners on all used measures

	Truth tellers *M* (*SD*)	Feigners *M* (*SD*)	Welch's *t*-test (*df*)[a]	Cohen's *d*
Narratives				
Length	161.09 (111.86)	231.63 (167.10)	$t(55.77) = 2.03$*	0.49
Total of verifiable details	33.81 (19.90)	34.52 (22.80)	$t(44.10) = .130$	0.03
Total of non-verifiable details	11.43 (11.66)	20.50 (15.65)	$t(51.05) = 2.64$*	0.65
Verifiable details from newspapers (VNP)	8.62 (3.61)	8.20 (3.56)	$t(38.29) = .45$	0.11
New verifiable details (V)	25.19 (18.30)	26.33 (21.25)	$t(44.75) = .22$	0.05
Proportion of total verifiable details	.77 (.15)	.64 (.14)	$t(37.36) = 3.37$**	0.90
Non-verifiable details from newspapers (NNP)	.24 (.54)	.17 (.53)	$t(38.20) = .45$	0.13
New non-verifiable details (NV)	11.19 (11.54)	20.33 (15.60)	$t(51.38) = 2.68$**	0.66
Symptom reports				
SRSI PTSD Genuine	2.81 (2.97)	5.85 (3.23)	$t(41.91) = 3.77$**	0.98
SRSI PTSD Pseudo	.71 (1.42)	2.67 (2.18)	$t(56.95) = 4.39$**	1.06

Notes. The proportion of verifiable information was calculated using the next formula: (VNP + V)/((VNP + V) + (NNP + NV)).

[a]All calculations were also performed using a non-parametric, *Mann–Whitney U test*, and the results did not differ.

*$p < .05$; **$p < .01$.

consequence of the difference in the number of new non-verifiable details (NV), *Welch's t* (51.38) = 2.68, $p = .010$, *Cohen's d* = 0.66. No other significant differences were found in the number of other categories (VNP, V, and NNP) of details (see Table 1).

We calculated the proportions of total verifiable information ((V + VNP)/Total details), verifiable details from the newspapers (VNP/Total details), and the new verifiable details (V/Total Details), controlling for length of statements. The proportion of total verifiable details (V + VNP/Total details) was significant, *Welch's* $t(37.36) = 3.37$, $p = .002$, *Cohen's d* = 0.90, with truth tellers having higher proportions of these details ($M = .77$, $SD = .15$) than feigners ($M = .64$, $SD = .14$). Looking separately in the two categories of verifiable information, no significant differences were found. The proportion of the verifiable details from the newspapers (VNP) was not significantly different between truth tellers ($M = .24$, $SD = .19$) and feigners ($M = .19$, $SD = .14$), *Welch's* $t(30.91) = 1.19$, $p = .24$, *Cohen's d* = 0.30. Similarly, the difference in the proportion of the new verifiable details (V) was not significant, *Welch's* $t(41.12) = 1.95$, $p = .057$, *Cohen's d* = 0.55,[4] although there was a higher proportion of new verifiable details among the truth tellers ($M = .52$, $SD = .14$) than among feigners ($M = .44$, $SD = .15$).

We re-analysed the narrative-related characteristics (Length, VNP, V, NNP, NV, Number of Verifiable details, Proportion of Verifiable details, Total number of Non-

[4] *The differences in all three proportions of verifiable details was calculated also using the* Mann–Whitney U *test, and the results remained on the similar level of statistical significance.*

verifiable details) using the two-way ANOVA, with Groups (truth tellers vs. feigners) and Language (Dutch vs. English) as independent variables. Neither the interaction between Groups and Language, nor the main effect of Language was found to be significant, $\lambda = .917$, $F(6, 58) = .873$, $p = .52$, $\lambda = .943$, $F(6, 58) = .590$, $p = .737$, respectively.

Symptom reports using the SRSI anxiety/PTSD subscales

Truth tellers reported significantly fewer genuine symptoms ($M = 2.81$, $SD = 2.97$) and pseudosymptoms ($M = .71$, $SD = 1.42$), than feigners (genuine symptoms, $M = 5.85$, $SD = 3.23$, and pseudosymptoms, $M = 2.67$, $SD = 2.18$). For genuine symptoms, *Welch's* $t(41.91) = 3.77$, $p < .001$, *Cohen's* $d = 0.98$, and *Welch's* $t(56.95) = 4.39$, $p < .001$, *Cohen's* $d = 1.06$, for pseudosymptoms.

Narratives and symptom reports

In order to investigate the relationship between the features of trauma-related narratives (the number and proportion of verifiable details and the number of non-verifiable details) and the quality of symptom reports (endorsement of genuine symptoms and pseudosymptoms), we calculated the *Pearson's r product–moment correlation coefficients*. For truth tellers, none of the correlations reached significance (all *p*s > .05). Additionally, we included the *delay after the exposure* (number of days before joining our study) as a variable, but none of the correlations were significant (all *p*s > .05), indicating no association between the delay and the quality of the truth tellers' accounts. However, among feigners, the non-verifiable details were significantly related to the endorsement of both genuine symptoms and pseudosymptoms, *Pearson's r* being .32 and .31, respectively (*p*s < .05). None of the other correlations were significant (see Table 2).

Diagnostic utility of the (non-)verifiable details and the SRSI subscales

First, we examined general detection accuracy for each of the main measures (number and proportions of total verifiable and non-verifiable details, as well as the genuine and pseudosymptoms), using the Area Under the Curve (AUCs) (see Table 3). All measures, except the number of verifiable details (AUC = .50, $p = .962$), detected feigners better than chance (AUCs > .70).

Table 2. Pearson's correlation coefficients for the features of fabricated narratives and the symptom reports among feigners

	SRSI Genuine symptoms	SRSI Pseudosymptoms
Truth tellers ($n = 21$)		
Verifiable details total	.30	−.04
Verifiable details proportion	−.19	−.41
Non-verifiable details total	.42	.19
Feigners ($n = 46$)		
Verifiable details total	.19	.15
Verifiable details proportion	−.14	−.11
Non-verifiable details total	.32*	.31*

Note. *$p < .02$.

Second, we tested the incremental validity of combining these measures: (1) the VA (proportion of verifiable detail & total number of non-verifiable details), (2) SRSI subscales (genuine anxiety/PTSD symptoms and pseudosymptoms), and (3) VA and SRSI subscales combined. We ran three *Discriminative Analyses* to distinguish between the truth tellers and feigners. The group membership was in all three cases the classifying variable. The analysis including the VA (the proportion of verifiable details and total non-verifiable details) yielded significant discriminative function, $\chi^2(2) = 10.95$, $\lambda = .84$, $p = .004$ (*canonical correlation* = .40). The SRSI subscales also significantly predicted the group classification of participants, $\chi^2(2) = 14.98$, $\lambda = .79$, $p = .001$ (*canonical correlation* = .45). Finally, the discriminant function was also significant for the VA and the SRSI subscales combined, $\chi^2(4) = 20.30$, $\lambda = .72$, $p = .001$ (*canonical correlation* = .52). The sensitivity, specificity, and the overall accuracy are provided in the Table 4.

Discussion

In the current research, we examined whether the combination of different detection strategies helps to expose the report strategies of those who feign PTSD. Specifically, we tested the Verifiability Approach (VA; Nahari *et al.*, 2014a) in the context of PTSD exposure narratives, and the utility of the anxiety/PTSD-related subscales of the Self-Report Symptom Inventory (SRSI; Merten *et al.*, 2016) within symptom validity assessment.

Our results are as follows: First, participants who were asked to feign the aversive exposure produced significantly longer narratives than truth tellers did. Earlier studies also showed that feigners elaborate on their injuries more than truth tellers (Purisch & Sbordone, 1997). Examination of the narratives' content revealed that feigners inflated the length of their reports by overproducing new, non-verifiable details. Hence, they focused on describing their subjective state during the scene (e.g., 'I was shocked') rather than on external circumstances, as shown in previous symptom-focused studies (e.g., Boskovic *et al.*, 2017, 2018). This means that, regardless of the origin of experience in question, feigners overcompensate the lack of truthful information with non-verifiable details. Importantly, every participant reported at least one detail that was, in principle, checkable. However, truth tellers' narratives included more information that was already available and more self-generated information that was checkable. These results align well with findings regarding verifiability within the lie-detection context in which the proportion of verifiable information was found to differentiate between truthful and

Table 3. The Area Under the Curve (AUC), significance (*p*) level, and Confidence Intervals (CI) of numbers and proportions of verifiable, number of non-verifiable details, and SRSI anxiety/PTSD genuine and pseudosymptoms

Measures	AUC	*p*	95% CI
Verifiable details total	.50	.962	.36–.65
Verifiable details proportion	.72	.004	.59–85
Non-verifiable details total	.70	.008	.57–.84
Non-verifiable details proportion	.72	.003	.59–85
SRSI genuine anxiety/PTSD symptoms	.74	.001	.62–87
SRSI anxiety/PTSD pseudosymptoms	.78	<.001	.66–89

Table 4. Sensitivity (true positive), specificity (true negative), and overall accuracy in classifying participants to truth tellers and feigners using narrative features and the subscales of the SRSI

Predictors	Sensitivity (%)	Specificity (%)	Overall accuracy (%)
Narrative			
Verifiable details proportion	47.6	93.5	79.1
Non-verifiable details			
SRSI subscales			
SRSI genuine symptoms	62	82.6	76.1
SRSI pseudosymptoms			
Narrative + SRSI	52	87	76.1

fabricated accounts (e.g., Harvey *et al.*, 2017). Thus far, the majority of the studies that inspected the narratives of real-life trauma exposure (e.g., Peace & Porter, 2011; Porter *et al.*, 2007) did not look into the statements' verifiability, which would be a good direction for future investigation of this topic.

Second, looking at the symptom report results, the anticipated response pattern emerged. Feigners endorsed significantly more symptoms from both genuine-anxiety/PTSD and pseudo-anxiety/PTSD subscales than truth tellers. This pattern of results fits well with previous research when using the complete SRSI questionnaire (Merten *et al.*, 2016), and it confirms the over-generalization strategy feigners exhibit when reporting about their complaints (Merten, Thies, Schneider, & Stevens, 2009). The inflated symptoms reports were also previously reported in other studies investigating feigned victimization claims (e.g., Peace, Porter, & Cook, 2010). Furthermore, we found that feigners' symptom over-endorsement was associated with a higher frequency of non-verifiable details in their narratives. This indicates that both the VA and the SRSI captured the *hyperbolism* that is typically found in fabricated accounts.

Third, the proportion of verifiable details, the total amount of non-verifiable details, and both SRSI subscales were valid predictors of feigning. However, this was not the case for the number of verifiable details, which might not be as an important feature of narratives in the symptom validity assessment (Boskovic *et al.*, 2017). Furthermore, when the VA was included in a discriminative function, the overall accuracy was 79%. The main advantage of the VA was the high true negative rate (specificity), meaning that truth tellers were correctly identified. In contrast, the SRSI subscales exhibited similar overall accuracy of 76%, with improved sensitivity at the cost of specificity. Combined, both measures produced an overall detection accuracy of 76%, with a balanced trade-off in terms of sensitivity and specificity.

A few methodological issues warrant comment. First, we included a healthy student sample, some of whom had been induced with trauma-like symptoms using the VR exposure. The VR paradigm is currently the most sophisticated method for induction of PTSD symptoms because it provides a more immersive environment than a Trauma Film paradigm (Dibbets & Schulte-Ostermann, 2015). Still, the VR exposure cannot fully imitate the real-life exposure; thus, our results have a limited generalizability to actual PTSD patients. Second, our sample was limited in size, which may have led to underpowered results. However, the main findings regarding narratives and symptom reports in two veracity groups correspond well to the results of previous studies in the field (e.g., Boskovic *et al.*, 2017). Third, because truth tellers were free to join the current

study at any time after the VR exposure, they may have chosen to proceed with our study once the effect of the exposure had declined. However, in this study, the correlation analyses indicated that the delay was not associated with the quality of truth tellers' accounts. Furthermore, all the participants in the truth telling condition were debriefed about the previous study before joining our study. There was no way of testing whether/how this influenced their reports, but it is likely that students anticipated our interest in the trauma-related symptomatology. Fourth, we did not apply the full version of the SRSI, as recommended by Merten *et al.* (2016); thus, a closer examination of detection accuracy was limited. Therefore, the findings concerning the SRSI subscales should be taken with caution and further tested using the whole SRSI. Fifth, some of the participants opted to write their narratives in Dutch and others in English. Recent studies showed that language could be an important confounder in symptom validity studies (Nijdam-Jones & Rosenfeld, 2017); thus, the language differences could have had an impact on our findings. However, the effect of language on our dependent variables was not significant. Germane to this is the finding that incentives, rather than language, have a significant impact on the response style in symptom validity assessment (van der Heide & Merckelbach, 2016). Sixth, and related to the previous issue, we offered every participant the opportunity to enter a lottery to win an additional financial reward for writing a 'convincing report'. This might have also motivated the truth tellers to increase the severity of their symptom reports. Feigners, besides the financial incentive, were asked to imagine having a chance to revenge to the train company. Revenge was shown to be a strong motivator (Peace & Masliuk, 2011). Yet, in reality, people claiming PTSD are often confronted with a significantly stronger (financial or emotional) incentive (positive or negative), which can have a different influence on their response style (Peace & Masliuk, 2011; Peace & Richards, 2014; Resnick *et al.*, 2008). Finally, due to the different pre-screening procedures of two groups, we were unable to compare the symptom reports of truth tellers and feigners prior to our study.

To summarize, non-verifiable and lengthier narratives remain a strong cue to fabrication of symptoms. Furthermore, an over-endorsement of both genuine and pseudosymptoms should raise doubt in the truthfulness of the aversive exposure claims. However, in order to validate our results, future research including the PTSD patients or/and accounts of the real-life aversive events is necessary.

Acknowledgements

This research is supported by a fellowship awarded from the Erasmus Mundus Joint Doctorate Program, The House of Legal Psychology (EMJD-LP) with Framework Partnership Agreement (FPA) 2013-0036 and Specific Grant Agreement (SGA) 2015-1610 to Irena Boskovic. The VR study (sample truth tellers) was funded by MIND Netherlands.

References

Adamou, M. C., & Hale, A. S. (2003). PTSD and the law of psychiatric injury in England and Wales: Finally coming closer? *Journal of the American Academy of Psychiatry and the Law Online*, *31*, 327–332.

Akehurst, L., Easton, S., Fuller, E., Drane, G., Kuzmin, K., & Litchfield, S. (2015). An evaluation of a new tool to aid judgements of credibility in the medico-legal setting. *Legal and Criminological Psychology*, *22*, 22–46. https://doi.org/10.1111/lcrp.12079

Amado, B. G., Arce, R., Farina, F., & Vilariño, M. (2016). Criteria-Based Content Analysis (CBCA) reality criteria in adults: A meta-analytic review. *International Journal of Clinical and Health Psychology*, *16*, 201–210. https://doi.org/10.1016/j.ijchp.2016.01.002

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: APA. https://doi.org/10.1176/appi.books.9780890425596

Bogaard, G., Meijer, E. H., Vrij, A., Broers, N. J., & Merckelbach, H. (2014). Contextual bias in verbal credibility assessment: Criteria-based content analysis, reality monitoring and scientific content analysis. *Applied Cognitive Psychology*, *28*, 79–90. https://doi.org/10.1002/acp.2959

Bonanno, G. A. (2005). Resilience in the face of potential trauma. *Current Directions in Psychological Science*, *14*, 135–138. https://doi.org/10.1111/j.0963-7214.2005.00347.x

Boskovic, I., Bogaard, G., Merckelbach, H., Vrij, A., & Hope, L. (2017). The Verifiability Approach to detection of malingered physical symptoms. *Psychology, Crime & Law*, *23*, 717–729. https://doi.org/10.1080/1068316X.2017.1302585

Boskovic, I., Gallardo, C., Vrij, A., Hope, L., & Merckelbach, H. (2018). Verifiability on the run: An experimental study on the Verifiability Approach to malingered symptoms. *Psychiatry, Psychology, and Law*, 1–12. https://doi.org/10.1080/13218719.2018.1483272

Cohen, Z. E., & Appelbaum, P. S. (2016). Experience and opinions of forensic psychiatrists regarding PTSD in criminal cases. *The Journal of the American Academy of Psychiatry and the Law*, *44*, 41–52.

Dibbets, P., & Schulte-Ostermann, M. A. (2015). Virtual reality, real emotions: A novel analogue for the assessment of risk factors of post-traumatic stress disorder. *Frontiers in Psychology*, *6*, 681–689. https://doi.org/10.3389/fpsyg.2015.00681

Freeman, T., Powell, M., & Kimbrell, T. (2008). Measuring symptom exaggeration in veterans with chronic posttraumatic stress disorder. *Psychiatry Research*, *158*, 374–380. https://doi.org/10.1016/j.psychres.2007.04.002

Guriel, J., & Fremouw, W. (2003). Assessing malingered posttraumatic stress disorder: A critical review. *Clinical Psychology Review*, *23*, 881–904. https://doi.org/10.1016/j.cpr.2003.07.001

Hall, R. C., & Hall, R. C. (2007). Detection of malingered PTSD: An overview of clinical, psychometric, and physiological assessment: Where do we stand? *Journal of Forensic Sciences*, *52*, 717–725. https://doi.org/10.1111/j.1556-4029.2007.00434.x

Hall, R., Hall, R., & Chapman, M. (2006). Effects of terrorist attacks on the elderly: Part 2. Posttraumatic stress, acute stress, and affective disorders. *Clinical Geriatrics*, *14*, 17–24.

Harvey, A. C., Vrij, A., Nahari, G., & Ludwig, K. (2017). Applying the Verifiability Approach to insurance claims settings: Exploring the effect of the information protocol. *Legal and Criminological Psychology*, *22*, 47–59. https://doi.org/10.1111/lcrp.12092

Hauch, V., Sporer, S. L., Masip, J., & Blandón-Gitlin, I. (2017). Can credibility criteria be assessed reliably? A meta-analysis of criteria-based content analysis. *Psychological Assessment*, *29*, 819–834. https://doi.org/10.1037/pas0000426

Johnson, M. K. (2006). Memory and reality. *American Psychologist*, *61*, 760–771. https://doi.org/10.1037/0003-066X.61.8.760

Jupe, L. M., Leal, S., Vrij, A., & Nahari, G. (2017). Applying the Verifiability Approach in an international airport setting. *Psychology, Crime & Law*, *23*, 812–825. https://doi.org/10.1080/1068316X.2017.1327584

Knoll, J., & Resnick, P. J. (2006). The detection of malingered post-traumatic stress disorder. *Psychiatric Clinics*, *29*, 629–647. https://doi.org/10.1016/j.psc.2006.04.001

Lees-Haley, P. R. (1997). MMPI-2 base rates for 492 personal injury plaintiffs: Implications and challenges for forensic assessment. *Journal of Clinical Psychology*, *53*, 745–755. https://doi.org/10.1002/(SICI)1097-4679(199711)53:7<745:AID-JCLP13>3.0.CO;2-L

Merckelbach, H., & Smith, G. P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology*, *18*, 145–152. https://doi.org/10.1093/arclin/18.2.145

Merten, T., Merckelbach, H., Giger, P., & Stevens, A. (2016). The Self-Report Symptom Inventory (SRSI): A new instrument for the assessment of distorted symptom endorsement. *Psychological Injury and Law*, *9*, 102–111. https://doi.org/10.1007/s12207-016-9257-3

Merten, T., Thies, E., Schneider, K., & Stevens, A. (2009). Symptom validity testing in claimants with alleged posttraumatic stress disorder: Comparing the Morel Emotional Numbing Test, the Structured Inventory of Malingered Symptomatology, and the Word Memory Test. *Psychological Injury and Law*, *2*, 284–293. https://doi.org/10.1007/s12207-009-9057-0

Nahari, G. (2018). The applicability of the Verifiability Approach to the real world. In J. P. Rosenfeld (Ed.), *Detecting concealed information and deception* (pp. 329–349). London, UK: Elsevier. https://doi.org/10.1016/b978-0-12-812729-2.00014-8

Nahari, G., & Vrij, A. (2014). Can I borrow your alibi? The applicability of the verifiability approach to the case of an alibi witness. *Journal of Applied Research in Memory and Cognition*, *3*, 89–94. https://doi.org/10.1016/j.jarmac.2014.04.005

Nahari, G., Vrij, A., & Fisher, R. P. (2014a). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology*, *19*, 227–239. https://doi.org/10.1111/j.2044-8333.2012.02069.x

Nahari, G., Vrij, A., & Fisher, R. P. (2014b). The verifiability approach: Countermeasures facilitate its ability to discriminate between truths and lies. *Applied Cognitive Psychology*, *28*, 122–128. https://doi.org/10.1002/acp.2974

Nijdam-Jones, A., & Rosenfeld, B. (2017). Cross-cultural feigning assessment: A systematic review of feigning instruments used with linguistically, ethnically, and culturally diverse samples. *Psychological Assessment*, *29*, 1321–1336. https://doi.org/10.1037/pas0000438

Peace, K. A., & Masliuk, K. A. (2011). Do motivations for malingering matter? Symptoms of malingered PTSD as a function of motivation and trauma type. *Psychological Injury and Law*, *4*, 44–55. https://doi.org/10.1007/s12207-011-9102-7

Peace, K. A., & Porter, S. (2011). Remembrance of lies past: A comparison of the features and consistency of truthful and fabricated trauma narratives. *Applied Cognitive Psychology*, *25*, 414–423. https://doi.org/10.1002/acp.1708

Peace, K. A., Porter, S., & Cook, B. L. (2010). Investigating differences in truthful and fabricated symptoms of traumatic stress over time. *Psychological Injury and Law*, *3*, 118–129. https://doi.org/10.1007/s12207-010-9078-8

Peace, K., & Richards, V. (2014). Faking it: Incentives and malingered PTSD. *Journal of Criminal Psychology*, *4*, 19–32. https://doi.org/10.1108/JCP-09-2013-0023

Porter, S., Peace, K. A., & Emmett, K. A. (2007). You protest too much, methinks: Investigating the features of truthful and fabricated reports of traumatic experiences. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *39*, 79–91. https://doi.org/10.1037/cjbs2007007

Purisch, A., & Sbordone, R. (1997). Forensic neuropsychology: Clinical issues and practice. *The Neuropsychology Handbook*, *2*, 309–356.

Rassin, E., Boskovic, I., & Merckelbach, H. (2018). Posttraumatic stress disorder and diminished criminal responsibility as "New Evidence" in criminal revision procedures. *Journal of Forensic Sciences*, *63*, 1911–1913. https://doi.org/10.1111/1556-4029.13789

Resnick, P. J., West, S., & Payne, J. W. (2008). Malingering of post-traumatic disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed.) (pp. 109–127). New York, NY: The Guilford Press.

Resnick, P. J., West, S. G., & Wooley, C. N. (2018). Malingering posttraumatic disorders. In R. Rogers & S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4th ed.) (pp. 22–61). New York, NY: The Guilford Press.

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law Online*, *25*, 183–189.

Steller, M., & Koehnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217–245). New York, NY: Springer Publishing Co.

Szogi, E. G., & Sullivan, K. A. (2018). Malingered Posttraumatic Stress Disorder (PTSD) and the effect of direct versus indirect trauma exposure on symptom profiles and detectability. *Psychological Injury and Law*, *11*, 1–11. https://doi.org/10.1007/s12207-018-9315-0

Tracy, D. K., & Rix, K. J. (2017). Malingering mental disorders: Clinical assessment. *BJPsych Advances*, *23*, 27–35. https://doi.org/10.1192/apt.bp.116.015958

van Dam, D., Ehring, T., Vedel, E., & Emmelkamp, P. M. (2013). Screening for posttraumatic stress disorder in civilian substance use disorder patients: Cross-validation of the Jellinek-PTSD screening questionnaire. *Journal of Substance Abuse Treatment*, *44*, 126–131. https://doi.org/10.1016/j.jsat.2012.03.005

van der Heide, D., & Merckelbach, H. (2016). Validity of symptom reports of asylum seekers in a psychiatric hospital: A descriptive study. *International Journal of Law and Psychiatry*, *49*, 40–46. https://doi.org/10.1016/j.ijlp.2016.05.007

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). The PTSD Checklist for DSM-5 (PCL-5) – Standard [Measurement instrument]. Available from http://www.ptsd.va.gov/

Young, G. (2016). PTSD in court I: Introducing PTSD for court. *International Journal of Law and Psychiatry*, *49*, 238–258. https://doi.org/10.1016/j.ijlp.2016.10.012

Received 5 June 2018; revised version received 8 January 2019

Appendix

Newspaper Article

Two dead in crash between passenger train and car

In a crash between a passenger train and a car on Monday evening two people died. It concerns the passengers of the car, including one male (28) and a 4-month-old baby. There were around 100 passengers on the train. They remained unharmed. The exact cause of the accident remains unclear. Information released by the relatives showed that the man tried to send a picture of his young daughter, using his smartphone, right before the crash. It is possible that the man was distracted and did not notice the railroad lights. The ongoing police investigation will have to offer conclusive information on this accident."