Dull versus creative liars—Who deceives better? Fantasy proneness and verifiability of genuine and fabricated accounts

(Article begins on next page)

07 January 2023

RESEARCH ARTICLE

WILEY

Dull versus creative liars—Who deceives better? Fantasy proneness and verifiability of genuine and fabricated accounts

Irena Boskovic[1,2] 🅾 | Ayla Ramakers[1] | Ali Yunus Emre Akca[1]

[1]Faculty of Psychology and Neuroscience, Forensic Psychology Section, Maastricht University, Maastricht, The Netherlands

[2]Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands

Correspondence
Irena Boskovic, Forensic Psychology Section, Maastricht University, PO Box 616, Maastricht 6200 MD, The Netherlands.
Email: irena.boskovic@maastrichtuniversity.nl

[Corrections made on 9 October 2020, after first online publication: Text in Section 2.1 has been updated in this version for clarity]

Abstract

The Verifiability Approach (VA), a lie detection method, postulates that genuine statements contain more verifiable information, whereas fabricated reports include more non-verifiable details. We investigated whether participants low ($n = 19$), medium ($n = 23$) and high ($n = 26$) on fantasy proneness differ in the (non)verifiability of their genuine and fabricated accounts. The results showed that groups did not differ in terms of statements' (non)verifiability. Overall, fabricated accounts included more non-verifiable details, but did not differ in verifiable details from genuine stories. The fabricated accounts from each group were given to legal psychology experts ($N = 13$) who rated accounts' authenticity. Experts more often recognised false accounts from the high fantasy proneness group, hence, high fantasy prone deceivers might be easier to detect than people with lower fantasy engagement. Overall, our results show that the VA is not sensitive to fantasy proneness, however, that experts might be.

KEYWORDS
deception detection, fantasy proneness, (non)verifiable details, verifiability approach

1 | INTRODUCTION

Over the last 30 years, the assumption that fabricated narratives contain different linguistic quality of reported content than truthful accounts (Undeutsch, 1967) has been properly tested and empirically supported (e.g., Nahari, Vrij, & Fisher, 2014; Porter & Yuille, 1996; Sporer, 1997). One of the most known methods for verbal credibility assessment is the Criteria Based Credibility Assessment method (CBCA; Steller & Köhnken, 1989), containing 19 criteria, to which the statements are compared (e.g., logical structure, unusual details and quantity of details). The criteria are coded using a 2-point scale ('0' meaning absent and '1' meaning present), and statements with higher overall scores are judged as more truthful (Steller & Köhnken, 1989). CBCA has received a lot of support among the researchers (e.g., Vrij, 2000) and practitioners and has been commonly applied in court cases in Europe (Köhnken, 2002; Lamers-Winkelman & Buffing, 1996).

However, CBCA was shown to have certain drawbacks. For instance, recent meta-analyses indicated that CBCA summary scores lack reliability, thus, their use in decision making is strongly discouraged (Hauch, Sporer, Michael, & Meissner, 2017). Other authors pointed out that some of the CBCA criteria lack clarity and cultural independence (e.g., 'unusual details'; Vrij, 2008), and demonstrated CBCA's vulnerability to contextual bias when additional, domain-irrelevant information is provided (see Bogaard, Meijer, Vrij, Broers, & Merckelbach, 2014). Also, certain CBCA criteria are biased towards specific demographic, such as criterion 'quantity of details', according to which children's accounts are less credible than those of adolescents', simply due to the limited vocabulary in younger age (e.g., Ernberg, Magnusson, Landstrom, & Tidefors, 2018).

Recently, researchers have developed an alternative method for verbal credibility assessment, the Verifiability Approach (VA; Nahari et al., 2014). The VA, in contrast to the CBCA, does not account for the quantity of details, but rather examines a specific *quality* of the provided information, mainly, its verifiability (Nahari & Vrij, 2014). The rationale behind the VA is that a truthful account will include more *verifiable details*, information that, in principle, can be checked, whereas liars will try to avoid providing such information to avoid being exposed. Thus, liars will report more details closed to verification—*non-verifiable details* (Nahari, 2018; Nahari et al., 2014). Details that can potentially be documented, actions performed with identifiable persons, actions that are witnessed by identifiable persons, or/and actions in which technology was used are coded as verifiable. Details that do not fall within these criteria are labelled as non-verifiable (Nahari & Vrij, 2014).

Researchers behind the VA even suggest that informing participants about the rationale behind the VA prior to the study by introducing the so-called Information Protocol (IP), enhances the detection of liars (Harvey, Vrij, Nahari, & Ludwig, 2017; Nahari & Vrij, 2014; Nahari & Vrij, 2019; Vrij, Nahari, Isitt, & Leal, 2016). Yet, several studies found the VA effect even without employing the IP (e.g., Boskovic, Bogaard, Merkcelbach, Vrij, & Hope, 2017; Jupe, Leal, Vrij, & Nahari, 2017; Nahari et al., 2014), while others in which the IP was used did not find the evidence that it enhances lie detection (e.g., Boskovic et al., 2017; Jupe, Vrij, Leal, & Nahari, 2020). However, assessing the proportion of verifiable details (verifiable details/[verifiable + non-verifiable details]; see Nahari & Vrij, 2019) has been recommended by many researchers as one of the consistent cues to detecting deception (Boskovic, Gallardo, Vrij, Hope, & Merckelbach, 2019; Nahari & Nisin, 2019; Nahari, 2018). Specifically, it has been shown that genuine statements include a higher proportion of verifiable information than fabricated accounts (Nahari et al 2014; Vrij & Nahari, 2019; for a meta-analysis see Verschuere, Bogaard, & Meijer, 2020). One of the strengths behind the application of proportion is that the length (i.e., word count) of the statements is controlled for. Traditionally, it has been established that fabricated accounts include fewer details than genuine statements (DePaulo et al., 2003; Nahari & Vrij, 2014; Vrij, 2000). This is also a commonly held opinion among the lay public and police officers (see Bogaard, Meijer, & Merckelbach, 2016). However, VA studies in the context of malingering showed that longer statements are a trademark of fabricated symptom accounts (Boskovic et al., 2017, 2019a, 2019b). Specifically, an abundance of non-verifiable details, rather than the lack of verifiable details, appeared to be the cause of longer statements among malingerers. This inconsistency in VA findings between lie detection and malingering contexts is attributed to the differences in the objectivity of the

reported events (e.g., alibis vs. symptoms; Nahari, 2019). Thus, lying about subjective experiences such as symptoms might include strategies that are less applicable in, for instance, alibi descriptions (i.e., providing more information).

Another factor that could significantly derail the way in which people report their experiences is personality characteristics. The impact of certain personality features on reports' quality has been well established in the malingering field (see Merckelbach, Boskovic, Pesy, Dalsklev, & Lynn, 2017; Merckelbach, Horselenberg, & Schmidt, 2002; Merckelbach, Prins, Boskovic, Niesten, & Campo, 2018; Peace & Masliuk, 2011), but has yet to receive as much attention in the lie detection context (Merckelbach, 2004; Peace, Brower, Rocchio, 2015). One of the personality traits that has been of particular interest for researchers is *fantasy proneness* (Wilson & Barber, 1983). Fantasy proneness refers to ones' ability to engage in fantasy and daydreaming, and it has been shown to influence the clarity and credibility of individuals' trauma reports (see Peace & Masliuk, 2011).

In one of the rare attempts to test fantasy proneness in the lie detection context, Peace et al. (2015) tested whether people high on fantasy proneness judge credibility of eyewitnesses differently than people low on this trait. Their results did not implicate any relationship between judges' fantasy proneness and their credibility ratings (Peace et al., 2015). However, Merckelbach (2004) demonstrated that the verbal credibility assessment might be obscured due to the examinees' differences in fantasy proneness. He tested participants with high and low levels of fantasy proneness who were asked to produce either only fabricated, or both genuine and fabricated stories, within the predetermined theme. All of the statements were then given to the independent judges who assessed the veracity of the accounts using the CBCA. The statements of participants with high levels of fantasy proneness were judged as more truthful than accounts produced by less fantasy prone participants. Thus, a high level of fantasy proneness was shown to directly interfere with a person's ability to generate convincing accounts.

That finding was supported also by a similar study by Schelleman-Offermans and Merckelbach (2010). These researchers also selected participants on the high and low ends of the fantasy proneness spectrum and gave them a task to write one genuine and one fabricated story. The narratives were then inspected using the CBCA criteria. Overall, it was shown that high fantasy prone participants' accounts, both genuine and fabricated, obtained higher scores on CBCA, meaning that they were better at deceiving than participants with low fantasy proneness. Therefore, it was concluded that assessing this personality trait is necessary before conducting a verbal credibility assessment.

We found only one study in which the relationship between participants' fantasy proneness and the (non) verifiability of provided details was briefly explored (see Bogaard, Meijer, & Van der Plas, 2020). These authors found a positive relationship between fantasy proneness and inclusion of verifiable details in fabricated statements, indicating that fantasy proneness could have an impact on VA outcomes. Thus, in this study, we wanted to closely test the influence of fantasy proneness on VA scores. In order to do so, we employed a similar design as to the one used in Schelleman-Offermans and Merckelbach (2010). However, besides including participants with low and high levels of fantasy proneness, we also included a group with moderate scores. Participants were asked to provide both genuine and fabricated statements. Moreover, as it is unknown how stable the fantasy proneness trait is, we measured the fantasy proneness levels of half of our sample 1 year prior to the study, and for the second half of the sample a couple of days before the study. Our study was conducted online, using Qualtrics. We expected that participants, in disregard of their fantasy proneness tendencies, would overall provide more verifiable details and less non-verifiable information in their genuine narratives than in their fabricated narratives. Furthermore, we also expected a relationship between participants' fantasy proneness levels and the quantity of (non)verifiable details in their statements. Specifically, we expected that participants with high fantasy proneness would provide more verifiable details and less non-verifiable information than medium and low fantasy proneness groups in both types of statements.

In the second part of the study, in order to check whether statements' credibility from different fantasy proneness groups is judged differently by experts, we randomly selected three fabricated statements from each

fantasy proneness group and gave them to the members of our department. We opted for only fabricated statements to ensure that experts' judgements specifically reflect participants' differences in confabulating false accounts, rather than subjects' veracity. Our expectation was that the experts would judge statements from the high fantasy proneness group as more plausible than statements from two other groups.

2 | METHOD

2.1 | Participants

Based on power calculations with $f = 0.40$ ($d = 0.80$; Schelleman-Offermans & Merckelbach, 2010), $\alpha = 0.05$ and $1 - \beta = 0.80$, a sample of 51 was deemed sufficient to investigate our research question using a repeated measures analysis of variance (ANOVA). The total sample consisted of 68 students of Maastricht University, the Netherlands, 54 were women and one participant preferred not to say the gender. The mean age was 21 years (SD = 1.77; range 18–27), and the main nationalities of participants were German (57.4%), Dutch (23.5%) and Belgian (6%). The sample was recruited based on a pre-screening with nine items taken from the Creative Experiences Questionnaire (CEQ; see below). One part of the sample ($n = 32$) completed the pre-screening 1 year ago, while the others ($n = 36$) completed it shortly before they participated in the study. Based on participants' scores, we formed three groups: (1) low fantasy proneness ($n = 19$; 13 women), (2) medium ($n = 23$; 19 women) and (3) high fantasy proneness ($n = 26$; 22 women) group.

For the second part of our study, we recruited 13 colleagues (61.5% female) from our forensic psychology section to judge the authenticity of statements provided by all three groups. The experts were blind to the veracity and origin of the statements. The average research experience of experts was 9 years (SD = 8.90; range 1–30 years). The study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience, Maastricht University (ref. ERCPN_186_03_12_2017).

2.2 | Materials

2.2.1 | Creative Experiences Questionnaire (Merckelbach, Horselenberg, & Muris, 2001)

The CEQ involves 25 items to test the fantasy proneness of participants. The measure includes items such as 'I often confuse fantasies with real memories', to which the participant responds with 'Yes' or 'No'. Sánchez-Bernardos and Avia (2004) conducted a factor analysis on the CEQ and found three factors, of which the first one explains the most variance in fantasy proneness. This first factor is about clarity and vividness of fantasies and includes nine items. These nine items were administered for the pre-screening (see Appendix S1), and based on participants' scores (range 0–9), we recruited low (lower quartile, a score of <3), high (upper quartile, a score of >6) groups of participants and a medium group consisting of participants who obtained a score of (=) 4. During the study, we administered the full version of the CEQ. The Cronbach's alpha for this full measure was 0.763.

2.2.2 | Narrative quality

The students' narratives were coded following the VA. The criteria for coding verifiable details were taken from Nahari and Vrij (2014) and also Boskovic et al. (2017) and those imply that a detail is verifiable if (1) it is documented, (2) it involved an action that was witnessed by one or more identifiable persons, (3) it involved an action that was carried out with one or more identifiable persons, (4) it was recorded, (5) the action involved technology

(i.e., ATM) or (6) it could be checked by blood analysis or medical tests. All details that did not meet any of these criteria were coded as non-verifiable. The content of the narratives was inspected by two independent coders, both blind to the conditions (for an example of coding see Appendix S3). One coder analysed all the statements, while the other analysed a randomly selected 25% ($n = 14$). The inter-rater agreement for the non-verifiable details was good, ICC $= 0.81$, and excellent for the verifiable details, ICC $= 0.91$. After coding, the absolute numbers of verifiable and non-verifiable details, as well as the proportions of verifiable details (verifiable details/(verifiable details $+$ non-verifiable details) were calculated.

2.2.3 | Experts' authenticity judgement

We randomly chose three fabricated narratives from each of the fantasy proneness groups and gave them to the members of the forensic psychology section ($n = 13$). The task was to rate the narratives in terms of authenticity using the 3-point scale (anchors: 1 = 'fully fabricated', 2 = 'not sure (could be fabricated/could be true)' and 3 = 'authentic'). Experts' authenticity judgement scores were summarised for each fantasy proneness group.

2.3 | Procedure

2.3.1 | Part 1

Every year the first-year bachelor students are administered a battery of tests. We included nine items of the CEQ as a part in that battery. After inspecting the data provided by the students, we selected only the students in the lower and higher quartile, and students whose scores were medium. A year later, we invited those students to participate in our study. At the same time, we recruited new participants who filled out the nine CEQ items as a pre-requirement to participate in our project. Again, after inspecting their scores, students in the lower and upper quartile, as well as those who obtained the medium score were invited to join the study. The pre-screening was done by paper and pen. Eligible students then received the Qualtrics link to the study. After giving their consent, participants were asked to fill in demographic information, such as age, gender and field of studies. Then, their task was to write a narrative about a genuine event that should have had happened to them in the last 7 days, and a fabricated event should have never happened to them. Both accounts should have been moderately emotional and around 200 words each. The instructions for these narratives were randomised and counter-balanced. This was to make sure that the writing style of one narrative would not have a systematically distorting effect on the other narrative. In order to further lower the writing effect of the previous narrative, in between these two tasks, they were administered the full CEQ. Then, participants were asked to grade their motivation, to evaluate the clarity of instructions and their proficiency in English. We also asked participants if they used a certain strategy while writing the narratives and if they tried to make the narratives similar. If they responded positively, they were asked to further elaborate on their responses. Finally, participants received a debriefing form.

2.3.2 | Part 2

We invited experts in the field of legal psychology to join our study and to judge the veracity of administered statements. Experts were told that statements could be truthful and/or fabricated. Their task was to grade the veracity and to describe the criteria they used for their gradings (for the instructions, see Appendix S2). After completing the study, experts were disclosed the aim of the study.

2.4 | Data analyses

To closely test the potential differences and interactions in narratives' (non)verifiability between the three fantasy proneness groups, a 3 (groups: low, medium and high) $\times$ 2 (veracity: genuine vs. fabricated) $\times$ 2 (details: verifiable or non-verifiable) repeated measures analyses of variance (ANOVA) was conducted. The proportions of verifiable details were inspected using the multivariate analyses of variance (MANOVA).

In order to test whether experts rated statements of high fantasy proneness group as more or less plausible than statements of other groups, we tested the differences using the *Wilcoxon Signed Ranks tests*.

3 | RESULTS

3.1 | Exit questions

All participants were asked to grade the clarity of instructions, their proficiency in English, experienced difficulty while writing statements and their motivation to provide both reports on a 5-point Likert scale (1 indicating the lowest levels and 5 indicating the highest). Overall, the clarity of instructions and the proficiency in English was graded as high, $M = 4.71$, $SD = 0.55$ and $M = 4.25$, $SD = 0.68$, respectively. Reported level of difficulty was somewhat low, $M = 2.62$, $SD = 1.02$, whereas participants' motivation was moderate for writing the both genuine, $M = 3.35$, $SD = 1.07$, and fabricated statements, $M = 3.76$, $SD = 0.96$. The low, medium and high fantasy proneness groups did not differ on any of the variables, Fs (2, 65) < 1.36, ps > 0.263.

We also asked participants whether they followed a certain strategy while writing the fabricated reports and 42.6% ($n = 29$) confirmed that they did. The majority used details, such as location, of true experiences (e.g., 'trip to Australia'), but fabricated the events that took place there. Others wrote about their dreams, or experiences of people they know, which are not actually experienced events, but contain certain familiarity. Thus, the dominant strategy of participants was to use embedded lies. We also asked them if they tried to make the genuine and fabricated statements comparable, and we found that 23.5% ($n = 16$) of participants confirmed using this strategy by including the same words or both statements of a similar writing style or a theme of a story.

3.2 | Stability of fantasy proneness

In order to check the relationship between the participants' nine-item CEQ pre-screening score and the score they later received on the full CEQ, we used *Pearson's product moment correlation coefficient* (*Pearson's r*). Significant correlations were found for both the participants who were pre-screened a year ago ($n = 32$; *Pearson's r* = 0.619, $p < 0.001$) and for the participants who completed the pre-screening shortly before entering this study ($n = 36$; *Pearson's r* = 0.822, $p < 0.001$).

Using a one-way ANOVA, we investigated the differences between groups on the total CEQ score. The groups scored significantly different on the total CEQ scale, F (2, 67) = 38.20, $p < 0.001$, $\eta_p^2 = 0.54$. Bonferroni post hoc tests confirmed the differences between all three groups ($ps \leq 0.001$). For the means and standard deviations see Table 1.

3.3 | Fantasy proneness and statements' quality

We investigated the relationship between the participants' full CEQ score and the length and verifiable and non-verifiable details of their genuine and fabricated reports, using the *Pearson's product moment correlation coefficient* (*Pearson's r*). The results indicated no significant correlation, *Pearson's rs* < 0.122, ps > 0.300, respectively.

TABLE 1 Total CEQ score among high, medium and low fantasy proneness group

Groups	*n*	Total CEQ score	
		M (SD)	**Minimum–maximum**
Low	19	6.58 (3.44)	1–12
Medium	23	10.09 (2.41)	5–16
High	26	14.27 (2.99)	8–19

Abbreviation: CEQ, creative experiences questionnaire.

3.4 | Statements' length

There were no significant differences *between* low, medium and high fantasy proneness groups in length (i.e., total word count) for genuine, $F_{(2, 65)} = 0.08$, $p = 0.918$, nor for fabricated statements, $F_{(2, 65)} = 0.23$, $p = 0.793$. However, to inspect the length differences *within* each group, *pair t*-tests indicated that, overall, all three groups of participants reported longer accounts when writing about fabricated events than when writing genuine statements (see Table 2).

3.5 | Verifiable, non-verifiable, and proportion of verifiable details among groups

The results of mixed repeated measures ANOVA indicated no significant interaction between groups (low, medium, high), veracity (genuine vs. fabricated) and details (verifiable vs. non-verifiable), $\lambda = 0.92$, $F_{(2, 65)} = 2.57$, $p = 0.084$, $\eta_p^2 = 0.073$.[1] Furthermore, the interaction between groups and veracity, $\lambda = 0.99$, $F_{(2, 65)} = 0.285$, $p = 0.75$, $\eta_p^2 = 0.009$, and the interaction between group and details, $\lambda = 0.99$, $F_{(2, 65)} = 0.161$, $p = 0.851$, $\eta_p^2 = 0.005$, did not reach significance. Similarly, the between-subject main effect of Group was insignificant, $F_{(2, 65)} = 0.32$, $p = 0.725$, $\eta_p^2 = 0.010$.

However, the interaction between veracity and details was significant, $\lambda = 0.92$, $F_{(1, 65)} = 5.28$, $p = 0.025$, $\eta_p^2 = 0.075$. The main effect of veracity, $F_{(1, 65)} = 15.04$, $p < 0.001$, $\eta_p^2 = 0.188$, was significant, while the main effect of details was not, $F_{(1, 65)} = 1.69$, $p = 0.197$, $\eta_p^2 = 0.025$. Closely examining the interaction, the difference driving this interaction effect was the dominance of non-verifiable details compared with verifiable details in fabricated statements, $F_{(1, 65)} = 4.92$, $p = 0.030$, $\eta_p^2 = 0.070$, whereas the difference between types of details was gone in genuine statements, $F_{(1, 65)} = 0.39$, $p = 0.534$, $\eta_p^2 = 0.006$.

The results of MANOVA including groups (low, medium and high) and proportions of verifiable details indicated no significant difference between groups, $\lambda = 0.90$, $F_{(4, 128)} = 1.60$, $p = 0.177$, $\eta_p^2 = 0.048$, in genuine, $F_{(2, 65)} = 1.88$, $p = 0.160$, $\eta_p^2 = 0.055$, nor in fabricated statements, $F_{(2, 65)} = 1.03$, $p = 0.360$, $\eta_p^2 = 0.030$ (for means and standard deviations see Table 3).

3.6 | Experts' authenticity judgements

Overall, experts judged the statements of low, $M = 2.36$, $SD = 0.34$, and medium group, $M = 2.38$, $SD = 0.44$, groups as more authentic than statements of high group, $M = 1.79$, $SD = 0.29$, $z = 2.30$, $p = 0.021$ and $z = 2.52$, $p = 0.012$, respectively.

We asked the experts to elaborate on the cues they used for their judgements and to grade the difficulty of the task and their certainty while making judgements, using a 5-point Likert scale (1 indicating the lowest levels and 5 indicating the highest). Experts reported using the CBCA criteria (e.g., number of details, quoted speech etc.), consistency, and plausibility of reported information as main indicators of veracity. Furthermore, they reported

TABLE 2 Length of genuine and fabricated statements among high, medium and low fantasy proneness group

Groups	Length of statements		Pair t-test	Cohen's d
	Genuine M (SD)	Fabricated M (SD)		
Low	172.11 (40.30)	204.11 (53.22)	t (18) = 3.88*	0.89
Medium	166.22 (57.04)	198.96 (61.70)	t (22) = 2.86*	0.59
High	171.96 (61.59)	211.46 (73.64)	t (25) = 3.39*	0.66

*$p < 0.01$.

TABLE 3 Verifiable, non-verifiable and proportion of verifiable details on genuine and fabricated statements among low, medium and high fantasy proneness groups

Groups	Details					
	Genuine statements M (SD)			Fabricated statements M (SD)		
	Verifiable	Non-verifiable	Verifiable proportion	Verifiable	Non-verifiable	Verifiable proportion
Low	18.21 (7.30)	19.16 (11.11)	0.51 (0.21)	20.84 (13.15)	21.47 (10.81)	0.47 (0.27)
Medium	18.26 (13.01)	18.87 (8.22)	0.46 (0.22)	18.21 (10.93)	23.69 (13.02)	0.43 (0.25)
High	21.77 (8.97)	16.76 (9.81)	0.57 (0.18)	17.84 (13.11)	27.77 (11.00)	0.37 (0.23)
Total	19.58 (10.15)	18.14 (9.62)	0.51 (0.21)	18.81 (12.31)	24.63 (11.79)	0.42 (0.25)

experiencing high difficulty, $M = 4.31$, $SD = 0.75$ and low certainty, $M = 2.38$, $SD = 0.96$, while judging statements' authenticity.[2]

4 | DISCUSSION

We investigated whether different levels of individuals' fantasy proneness influence the way they confabulate genuine and fabricated statements. Specifically, we investigated whether genuine and fabricated accounts differ in terms of verifiable and non-verifiable information, as proposed by the VA (Nahari et al., 2014), between low, medium and high fantasy proneness groups.

The findings of this study can be summarised as follows: (1) Fantasy proneness, measured by the CEQ, was shown to be a stable personality trait over 1 year. (2) However, different levels of this trait were not found to be related to alterations in the length of genuine or fabricated accounts, nor to the differences in their (non)verifiability. (3) The lack of differences in statements' (non)verifiability between the three fantasy proneness groups was supported by our correlational analyses which revealed no significant relationship between fantasy proneness and (non)verifiability of individuals' accounts. Thus, our hypotheses that fantasy proneness and narratives' (non)verifiability are related, as well as that high fantasy prone participants would provide more verifiable information, were not supported by our data. (4) We did, however, find one difference in the features of the narratives, not between the groups, but rather between the two veracity conditions. Specifically, all participants provided longer accounts when reporting about the event that has never happened to them. As discussed in the introduction, this finding opposes the empirically supported view in the lie detection field that truth tellers generate longer accounts than liars (DePaulo et al., 2003; Nahari, Vrij, & Fisher, 2012; Nahari & Vrij, 2019; Vrij, 2000). However, a few studies have shown that fabricated accounts are longer than genuine reports (Akehurst et al., 2017; Boskovic et al., 2017, 2019a, 2019b). The reasons behind this discrepancy were usually attributed to the differences in the tested settings (e.g., Nahari, 2019). However, further research of other possible influences should be conducted. (5) The main analyses further revealed that a likely cause of the differences in length is due to the abundance of non-verifiable

information in fabricated statements. This result partially supports our expectation that fabricated statements would include more non-verifiable details but less verifiable information. However, the difference in the number of verifiable information was not found between the two veracity conditions. All participants provided significantly more non-verifiable information in their fabricated statements than in their genuine stories. This finding is not surprising considering that liars have a known tendency to provide vague information (Boskovic et al., 2019b; Nahari, 2018; Nahari & Vrij, 2014). Looking at the strategies participants reported using for the fabricated stories, such as basing them on dreams, it is understandable that those narratives contained more, although vague, details. Previous research showed that memories of dreams can contain many details (Johnson, Kahan, & Raye, 1984) and can be even misremembered as genuinely experienced events (Mazzoni & Loftus, 1996). Actually, 76% of students reported having realistic dreams, while 8% of participants stated even acting upon a memory originating from a dream, due to source monitoring errors (Rassin, Merckelbach, & Spaan, 2001). (6) However, according to the VA approach, the proportion of verifiable details, rather than the abundance of non-verifiable information, serves as the main indicator of veracity (Nahari, 2018; Vrij, 2016). Yet, our investigation showed a general lack of differences in the proportion of verifiable details between participants' genuine and fabricated statements. It is possible that due to the instructions to focus on a moderately emotional experience, participants were more focused on describing the emotional valance that the event had, rather than describing the happenings that took place. As subjective experiences, such as emotions, oftentimes have unverifiable status (Nahari et al., 2012), the utility of details' proportion should be further tested using a more objective story theme. (7) Finally, our experts, even though they only received fabricated statements, judged the narratives of the low and medium group as significantly more authentic than those of the high fantasy proneness group. These findings go against our hypothesis and previous research in this field indicating the superiority of high fantasy prone subjects in fabricating stories (Merckelbach, 2004). It is possible that, due to the lack of specifically predefined theme for fabricated events (e.g., Vrij, 2008), high fantasy prone individuals in our study generated less plausible stories. Previous research showed that high fantasy prone subjects report more 'out-of-body' experiences and have stronger beliefs in paranormal activities (Gow, Lang, & Chant, 2004). Furthermore, it was also established that participants with high levels of fantasy provide significantly inflated trauma stories (Merckelbach, Muris, Horselenberg, & Stougie, 2000; Peace & Masliuk, 2011). As our instructions to report emotional stories might have led participants to write about trauma-like experiences, it is likely that stories of high fantasy prone group were less plausible, thus, detected at a higher rate.

There are certain limitations that warrant comment. First, our a priori power analysis was based on possibly overestimated (large) effect size, thus, that our sample size could have been insufficient to capture possible differences between participants with different levels of fantasy proneness. Thus, a future investigation of this topic might require a larger sample. Second, it is possible that groups did not mutually differ enough to capture more diverse levels of fantasy proneness (i.e., restriction of range). Furthermore, our participants were students; thus, the generalisability of our findings is limited. Third, the motivation of students to fabricate a convincing story for a study significantly differs from the motivation that would be present in a real-life high stake situation. Thus, the ecological validity of this study is restricted. Fourth, the instructions to include a moderately emotional event served the purpose of creating a wide universal theme of reports. However, this type of instructions allowed for a large variability in interpreting what qualifies as an emotional event (i.e., degrees of freedom; see Vrij, 2000), thus limiting the comparability of statements. It is also likely that such instructions drew participants' attention towards describing the emotional (i.e., non-verifiable) characteristics of events, and away from reporting objective, verifiable aspects of them. Perhaps, if the IP had been included in this study, participants would steer away from subjective description. Thus, further investigation of the individual differences' impact on the VA should incorporate the IP. Finally, the main part of this study was performed online, which enabled participants to, potentially, seek inspiration for their fabrication in different media outlets online, consequently creating more verifiable stories. Therefore, further laboratory studies, in which the availability of other sources of information is controlled, are needed.

In conclusion, the main outcomes of this study showed that generating (non)verifiable information is not sensitive to personality characteristics, such as fantasy and imagination engagement. Yet, it appeared that examinees' high levels of fantasy proneness have an impact on experts' credibility judgements. Thus, although the presence of high fantasy proneness per se does not indicate deception, it seems to enhance the suspicion of deception, which could lead to erroneous outcomes of verbal credibility assessment. Thus, further investigation of the fantasy proneness' impact on credibility assessment is required.

ACKNOWLEDGEMENT

CONFLICT OF INTEREST

The author declares that there is no conflict of interest.

ORCID

Irena Boskovic 🔟 https://orcid.org/0000-0003-1675-1004

ENDNOTES

[1] We also performed the main analyses with English proficiency as a covariate. The results did not differ, and none of the interactions ($Fs < 1.46$, $ps > 0.180$), nor the main effect of English proficiency, $F (1, 64) = 0.909$, $p = 0.344$, $\eta_p^2 = 0.01$, were significant.

[2] Data, syntax and outputs are available on the Open Science Framework platform (https://osf.io/qntg7/).

REFERENCES

Akehurst, L., Easton, S., Fuller, E., Drane, G., Kuzmin, K., & Litchfield, S. (2017). 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

Bogaard, G., Meijer, E. H., & Van der Plas, I. (2020). A model statement does not enhance the verifiability approach. *Applied Cognitive Psychology*, *34*, 96–105. https://doi.org/10.1002/acp.3596

Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Strong, but wrong: Lay people's and police officers' beliefs about verbal and nonverbal cues to deception. *PLoS One*, *11*, e0156615. https://doi.org/10.1371/journal.pone.0156615

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

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., Dibbets, P., Bogaard, G., Hope, L., Jelicic, M., & Orthey, R. (2019a). Verify the scene, report the symptoms: Testing the Verifiability Approach and SRSI in the detection of fabricated PTSD claims. *Legal and Criminological Psychology*, *24*, 241–257. https://doi.org/10.1111/lcrp.12149

Boskovic, I., Gallardo, C. T., Vrij, A., Hope, L., & Merckelbach, H. (2019b). Verifiability on the run: An experimental study on the verifiability approach to malingered symptoms. *Psychiatry, Psychology and Law*, *26*, 65–76. https://doi.org/10.1080/13218719.2018.1483272

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74–118. https://doi.org/10.1037/0033-2909.129.1.74

Ernberg, E., Magnusson, M., Landström, S., & Tidefors, I. (2018). Court evaluations of young children's testimony in child sexual abuse cases. *Legal and Criminological Psychology*, *23*, 176–191. https://doi.org/10.1111/lcrp.12124

Gow, K., Lang, T., & Chant, D. (2004). Fantasy proneness, paranormal beliefs and personality features in out-of-body experiences. *Contemporary Hypnosis*, *21*, 107–125.

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. http://dx.doi.org/10.1037/pas0000426

Johnson, M. K., Kahan, T. L., & Raye, C. L. (1984). Dreams and reality monitoring. *Journal of Experimental Psychology: General*, *113*, 329–344. https://doi.org/10.1037/0096-3445.113.3.329

Jupe, L. M., Vrij, A., Leal, S., & Nahari, G. (2020). Fading lies: Applying the verifiability approach after a period of delay. *Psychology, Crime & Law*, *26*, 327–342. https://doi.org/10.1080/1068316X.2019.1669594

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.2017.13275

Köhnken, G. (2002). German perspective on children's testimony. In H. L. Westcott, M. D. Graham, & R. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 233–244). Chichester, UK: John Wiley & Sons Ltd. https://doi.org/10.1002/9780470713679.ch15

Lamers-Winkelman, F., & Buffing, F. (1996). Children's testimony in The Netherlands: A study of statement validity analysis. *Criminal Justice and Behavior*, *23*, 304–321. https://doi.org/10.1177/0093854896023002004

Mazzoni, G. A., & Loftus, E. F. (1996). When dreams become reality. *Consciousness and Cognition*, *5*, 442–462. https://www.sciencedirect.com/science/article/abs/pii/S1053810096900276

Merckelbach, H. (2004). Telling a good story: Fantasy proneness and the quality of fabricated memories. *Personality and Individual Differences*, *37*, 1371–1382. https://doi.org/10.1016/j.paid.2004.01.007

Merckelbach, H., Boskovic, I., Pesy, D., Dalsklev, M., & Lynn, S. J. (2017). Symptom overreporting and dissociative experiences: A qualitative review. *Consciousness and Cognition*, *49*, 132–144. https://doi.org/10.1016/j.concog.2017.01.007

Merckelbach, H., Horselenberg, R., & Muris, P. (2001). The creative experiences questionnaire (CEQ): A brief selfreport measure of fantasy proneness. *Personality and Individual Differences*, *31*, 987–995. https://doi.org/10.1016/S0191-8869(00)00201-4

Merckelbach, H., Horselenberg, R., & Schmidt, H. (2002). Modeling the connection between self-reported trauma and dissociation in a student sample. *Personality and Individual Differences*, *32*, 695–705. https://doi.org/10.1016/S0191-8869(01)00070-8

Merckelbach, H., Muris, P., Horselenberg, R., & Stougie, S. (2000). Dissociative experiences, response bias, and fantasy proneness in college students. *Personality and Individual Differences*, *28*, 49–58.

Merckelbach, H., Prins, C., Boskovic, I., Niesten, I., & Campo, J. À (2018). Alexithymia as a potential source of symptom over-reporting: An exploratory study in forensic patients and non-forensic participants. *Scandinavian Journal of Psychology*, *59*, 192-197. https://doi.org/10.1111/sjop.12427

Nahari, G. (2018). The applicability of the Verifiability Approach to the real world. In P. J. Rosenfeld (Ed.), *Detecting concealed information and deception* (pp. 329–349). London, UK: Elsevier Inc. https://doi.org/10.1016/B978-0-12-812729-2.00014-8

Nahari, G., & Nisin, Z. (2019). Digging further into the speech of liars: Future research prospects in verbal lie detection. *Frontiers in Psychiatry*, *10*, 56. https://doi.org/10.3389/fpsyt.2019.00056

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. (2019). The Verifiability Approach: Advances, challenges, and future prospects. In R. Bull & I. Blandon-Gitlin (Eds.), *The Routledge International Handbook of Legal and Investigative Psychology* (pp. 212–224). New York, NY: Routledge.

Nahari, G., Ashkenazi, T., Fisher, R. P., Granhag, P. A., Hershkowitz, I., Masip, J., ... Verschuere, B. (2019). Language of lies': Urgent issues and prospects in verbal lie detection research. *Legal and Criminological Psychology*, *24*, 1–23. https://doi.org/10.1111/lcrp.12148

Nahari, G., Vrij, A., & Fisher, R. P. (2014). 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. (2012). Does the truth come out in the writing? SCAN as a lie detection tool. *Law and Human Behavior*, *1*, 68–76.

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., Brower, K. L., & Rocchio, A. (2015). Is truth stranger than fiction? Bizarre details and credibility assessment. *Journal of Police and Criminal Psychology*, *30*(1), 38–49.

Porter, S., & Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior*, *20*, 443–458. https://doi.org/10.1007/BF01498980

Rassin, E., Merckelbach, H., & Spaan, V. (2001). When dreams become a royal road to confusion: Realistic dreams, dissociation, and fantasy proneness. *The Journal of Nervous and Mental Disease*, *189*(7), 478–481.

Sánchez-Bernardos, M. L., & Avia, M. D. (2004). Personality correlates of fantasy proneness among adolescents. *Personality and Individual Differences*, 37, 1069–1079. https://doi.org/10.1016/j.paid.2003.11.015

Schelleman-Offermans, K., & Merckelbach, H. (2010). Fantasy proneness as a confounder of verbal lie detection tools. *Journal of Investigative Psychology and Offender Profiling*, 7, 247–260. https://doi.org/10.1002/jip.121

Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, 11, 373–397. https://doi.org/10.1002/(SICI)1099-0720(199710)11:5<373::AID-ACP461>3.0.CO;2-0

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

Undeutsch, U. (1967). Beurteilung der glaubhaftigkeit von aussagen. *Handbuch der psychologie*, 11, 26–181.

Verschuere, B., Bogaard, G., & Meijer, E. (2020). *Discriminating deceptive from truthful statements using the verifiability approach: A meta-analytic approach* (Working Paper). Retrieved from https://www.researchgate.net/profile/Bruno_Verschuere/project/Discriminating-lie-from-truth-using-the-Verifiability-Approach-A-Meta-analysis/attachment/5ee248584c18f90001294ff5/AS:901257581568000@1591887960054/download/VA+MetaAnalysis+Working-Paper+11.06.2020.pdf?context=ProjectUpdatesLog

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

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*, Chichester, UK: John Wiley & Sons.

Vrij, A. (2016). Baselining as a lie detection method. *Applied Cognitive Psychology*, 30, 1112–1119. https://doi.org/10.1002/acp.3288

Vrij, A., Nahari, G., Isitt, R., & Leal, S. (2016). Using the verifiability lie detection approach in an insurance claim setting. *Journal of Investigative Psychology and Offender Profiling*, 13, 183–197.

Vrij, A., & Nahari, G. (2019). The verifiability approach. In J. J. Dickinson, N. Schreiber Compo, R. N. Carol, B. L. Schwartz, & M. R. McCauley (Eds.), *Evidence-based investigative interviewing* (pp. 116–133). New York, NY: Routledge.

Wilson, S. C., & Barber, T. X. (1983). *Inventory of childhood memories and imaginings*. Framingham, MA: Cushing Hospital.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.