Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Personality and Individual Differences 37 (2004) 1371–1382
PERSONALITY AND INDIVIDUAL DIFFERENCES
www.elsevier.com/locate/paid

Telling a good story: Fantasy proneness and the quality of fabricated memories

Harald Merckelbach *

Department of Experimental Psychology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

Received 19 June 2003; received in revised form 2 December 2003; accepted 21 January 2004
Available online 21 February 2004

Abstract

In two studies, we examined whether fantasy prone people are superior storytellers. In study 1, participants high or low on fantasy proneness ($N = 25$) were instructed to fabricate a memory about an aversive childhood event. Independent judges rated stories of high fantasy prones as more emotional, more plausible, and richer in Criteria Based Content Analysis (CBCA) elements than those of low fantasy prone people. In study 2, high and low fantasy prone participants ($N = 38$) wrote down a true and a fabricated story about a negative event. Although the stories of the two groups did not differ in terms of emotionality or plausibility ratings, both true and fabricated stories of high fantasy prones were rated as being richer in CBCA elements than those of low fantasy prone controls. Taken together, the two studies show that fantasy proneness affects CBCA ratings. Forensic experts employing this tool would be well advised to take into account the potential confounding influence that fantasy proneness might have on CBCA evaluations.

Keywords: Fantasy proneness; Fabricated memories; CBCA; Malingering; Dissociation

1. Introduction

Many researchers (e.g., Pezdek & Taylor, 2000; Undeutsch, 1982) assume that fabricated and truthful stories have quite different characteristics. In fact, this assumption is not without empirical credence (e.g., Porter, Yuille, & Lehman, 1999; Sporer, 1997) and, in any event, it has served as the starting point for the development of various forensic tools. A case in point is the Criteria Based Content Analysis (CBCA; Steller & Koehnken, 1989), which aims at evaluating the

* Tel.: +31-43-388-1945; fax: +31-43-388-4196.
E-mail address: h.merckelbach@psychology.unimaas.nl (H. Merckelbach).

0191-8869/$ - see front matter
doi:10.1016/j.paid.2004.01.007

truthfulness of a story by matching its content against a set of 19 credibility criteria. These criteria focus on qualitative aspects like coherence of the story and amount of superfluous and unusual details (see, for an extensive listing of the CBCA criteria, Ruby & Brigham, 1997). Although the CBCA was originally developed for assessing children's accounts of sexual abuse, more recent research has examined whether this tool might contribute to a more accurate discrimination between adults' truthful and false statements. Over the past few years, evidence has accumulated supporting the effectiveness of CBCA criteria in distinguishing between true and false stories. In their scholarly review, Ruby and Brigham (1997, p. 729) summarize this evidence, but also conclude that "a number of empirical and theoretical issues must be addressed before the CBCA or similar techniques can legitimately be proposed as reliable and valid and to provide powerful evidence that is of probative value in evaluating the fate of a defendant in a sexual abuse case." One of the problems that Ruby and Brigham (1997) identify is that we do not know how well the CBCA fares with people who are talented storytellers.

That stories play an important role in the legal domain has long been recognized. For example, Bennett and Feldman (1981, p. 1) argued that the "criminal trial is organized around storytelling". This is supported by the pioneering work of Penington and Hastie (1988) who showed that participants are more likely to find a defendant guilty of murder when prosecution evidence is ordered in story form and defence evidence is not. Quite the opposite happens when there is a good story about the defendant, while prosecution evidence lacks the coherence of a story. Under such conditions, participants are less likely to find a defendant guilty.

Thus, stories do matter and this is not only true for the legal arena, but also for the medical domain. A good example in this context is provided by the phenomenon of malingering. Purisch and Sbordone's (1997, p. 350) noted that "malingerers do not present with la belle indifference. Their intent is to convey their symptoms to justify their disability along with their concern and distress about being so afflicted. As a consequence, they are likely to describe their injury in a highly detailed manner, which is usually offered as proof of their claims of disability. The forensic neuropsychologist should not be tempted to interpret litigants' highly detailed histories as proof of credibility of their symptoms and disability".

People differ in their storytelling abilities. One individual difference dimension that might play a crucial role in this context is fantasy proneness (Wilson & Barber, 1983). Individuals who are high on fantasy proneness typically spend a large part of their time fantasizing and daydreaming. Because of their vivid and detailed fantasies, they are talented in role playing and pretending to be someone else. Wilson and Barber (1983) describe the example of a fantasy prone individual who introduced herself as an Eskimo to the person sitting next to her in the bus and then went on to tell this stranger a detailed, but entirely fabricated story about her life in Alaska. A related feature of fantasy prone individuals is that they are able to experience physical symptoms when they fantasize about illnesses (Candel & Merckelbach, 2003; Wilson & Barber, 1983). Given these qualities, one would predict that individuals high on fantasy proneness are better at fabricating a plausible story than individuals low on fantasy proneness. To the present author's knowledge, this prediction has not yet been subjected to empirical testing. This is surprising given the fact that fantasy proneness is a trait that strongly overlaps with dissociative symptoms (e.g., Merckelbach & Muris, 2001; Merckelbach, Muris, Horselenberg, & Stougie, 2000; Merckelbach, Rassin, & Muris, 2000). Clinical literature offers many anecdotal reports of patients who feign such symptoms (e.g., Brick & Chu, 1991; Dinwiddie, North, & Yutzy, 1993). Some researchers have

also noted that adults with recovered memories about childhood abuse tend to score high on both dissociation and fantasy proneness (e.g., McNally, Clancy, Pitman, & Schacter, 2000). Discussions surrounding the phenomenon of recovered memories touch upon legal issues of credibility as well as the possibility of symptom malingering. With these considerations in mind, the current studies examined whether people scoring high on fantasy proneness are better at fabricating a memory of a negative event than are control participants.

2. Study 1

In this study, undergraduates high or low on fantasy proneness were instructed to fabricate a recovered memory about physical abuse by a teacher. In addition, they were asked to complete a list of items derived from post-traumatic stress and schizophrenia self-report scales and to feign symptoms on these items in a convincing way. Two independent judges who were blind as to participants' fantasy proneness scores, evaluated emotionality and plausibility of the fabricated stories. Two other independent and blind judges rated the stories in terms of relevant CBCA criteria. We expected that stories of fantasy prone participants would be rated as more emotional and plausible and would meet more CBCA criteria of credibility than those of control participants. As to the symptom items, we predicted that compared to control participants, fantasy prone individuals would tend to endorse both post-traumatic stress items and schizophrenia items. Thus, we anticipated that fantasy prone persons would exhibit an exaggerated and "over-the-top" endorsement of items (e.g., Porter et al., 1999). This expectation was based on previous research showing that fantasy proneness is related to a positive response bias on questionnaires sampling highly diverse autobiographical experiences (Merckelbach, Muris, et al., 2000) or scales listing bizarre and atypical symptoms (Merckelbach & Smith, 2003).

3. Method

3.1. Participants

During a mass testing session, 203 undergraduates were administered the Creative Experiences Questionnaire (CEQ; Cronbach's alpha = 0.89; Merckelbach, Horselenberg, & Muris, 2001) and the Dissociative Experiences Questionnaire (DES; Cronbach's alpha = 0.93; Bernstein & Putnam, 1986). The CEQ is a 25-item dichotomous self-report measure of fantasy proneness. Findings summarized by Merckelbach et al. (2001) indicate that CEQ scores demonstrate adequate test–retest stability and internal consistency and are strongly associated with concurrent measures of fantasy proneness and related constructs (e.g., absorption), with correlations ranging from 0.70 (with the 34-item Tellegen Absorption Scale; Tellegen & Atkinson, 1974) to 0.77 (with the 44-item Inventory of Childhood Memories and Imaginings version described by Myers, 1983). Eleven months after the mass testing session, attempts were made to contact 20 participants scoring in the upper and lower deciles of the CEQ distribution. When participants were successfully contacted, they were invited to participate in an experiment on story telling. Participants were not informed about their CEQ status. This procedure resulted in a high ($n = 11$; two men) and a low fantasy

proneness ($n = 14$; three men) subgroup. Mean age of the participants was 19.9 years (SD = 3.1). Mean CEQ scores of high and low fantasy proneness participants were 14.00 (SD = 1.55) and 2.71 (SD = 1.20), respectively: t (23) = 20.52, $P < 0.01$. High fantasy proneness participants also reported more dissociative symptoms on the DES than did low fantasy proneness individuals, means being 28.22 (SD = 14.01) and 13.01 (SD = 7.50), respectively, t (23) = 3.48, $P < 0.01$.

3.2. Procedure

Participants were provided with written instructions that asked them to imagine the following story. "You have difficulties with understanding other people's speech. So, you go to see a doctor and he examines your ears. The doctor concludes that your left inner ear has been damaged. You ask the doctor about the causes of this damage. He tells you that there are several possibilities. The doctor then gives a detailed description of one of his adult patients who developed serious hearing difficulties because during his primary school, an angry teacher had beaten him up. When you come home, you start to have memories about a similar incident during your primary school." Participants were asked to write down a plausible one-page story about these memories. After having written down their story, participants completed a list of 29 items. Seventeen were PTSD Symptom Scale-Self-Report items (PSS-SR; Foa, Riggs, Dancu, & Rothbaum, 1993). These items relate to typical Post-Traumatic Stress Disorder (PTSD) symptoms such as "have you had upsetting thoughts or images about the assault that came into your head when you didn't want them to?" Items were scored on a 4-point scale anchored 0 (*not at all*) and 3 (*very much*). PSS-SR items were randomly mixed with the 12 items of the Launay-Slade Hallucination Scale (LSHS; Launay & Slade, 1981) The LSHS is a widely used instrument for measuring the disposition to hallucinate. It consists of 12 statements that refer to hallucinatory experiences. Sample items are "Sometimes my thoughts seem as real as actual events in my life" and "I have been troubled by hearing voices in my head". For the present purpose, participants scored each item on a 4-point scale (0 = *not at all*; 3 = *very much*). Following the procedure of Lees-Haley (1989), participants were instructed to imagine that they had filed a lawsuit for emotional and physical damages associated with the teacher's misbehavior and that they had been sent to a psychologist for an independent examination. Their goal was to complete the items so as to feign symptoms in a manner that would convince the psychologist that they suffered from psychopathology as a result of the aversive childhood experience. Participants were not informed about the characteristics of PTSD.

3.3. Ratings

The stories were presented in a random order to two independent judges (one forensic psychologist and one teacher) who rated them in terms of emotionality and global plausibility, using 10-point scales (1 = *not at all emotional/plausible*; 10 = *extremely emotional/plausible*). Raters were explicitly instructed that all stories were fabricated and that they should give their ratings on a purely intuitive basis. Two other independent raters (research psychologists familiar with the CBCA literature) evaluated the randomly presented stories in terms of nine relevant CBCA criteria, using 4-point scales (e.g., Porter & Yuille, 1996; 0 = *criterion not present*; 3 = *criterion strongly present*). The criteria were: logical structure, quantity of details, contextual embedding,

interactions, reproduction of speech, unusual details, superfluous details, subjective experiences, and attribution of perpetrator's mental state (see for a complete listing, Ruby & Brigham, 1997). Raters were provided with a definition of each criterion taken from Ruby and Brigham (1997) and they knew that each story was fabricated.

4. Results and discussion

Stories of fantasy prone individuals contained somewhat more words than those of low fantasy prone individuals, means being 171 (SD = 78.9) and 140 (SD = 51.6), respectively. However, this difference in verbosity did not attain significance: t (23) = 1.18, $P = 0.25$. Emotionality and plausibility ratings of the two independent raters were moderately correlated with each other (emotionality: $r = 0.45$, $P < 0.05$; plausibility: $r = 0.46$, $P < 0.05$). For the final analysis, emotionality and plausibility ratings were averaged across raters. CBCA ratings had moderate (rater 1: Cronbach's alpha = 0.66) to good (rater 2: Cronbach's alpha = 0.82) internal reliabilities. The correlation between summed CBCA scores of raters 1 and 2 was 0.52, $P < 0.01$. For the final analysis, summed CBCA ratings were averaged across the two raters. Participants' scores on the self-report items were summed for PSS-SR and LSHS items separately and expressed as percentages of the maximum possible scores (51 and 36, respectively). Table 1 shows mean emotionality, plausibility, and CBCA ratings for high and low fantasy prone participants' stories. Participants' mean percentage scores on PSS-SR and LSHS items are also shown.

As can be seen, stories written by high fantasy prone individuals received higher emotionality and plausibility ratings than those written by low fantasy prone individuals, t (23) = 3.66, $P < 0.01$ and t (23) = 2.10, $P < 0.05$, respectively. As well, CBCA ratings were higher for stories of high fantasy prone individuals than for those of low fantasy prone individuals, t (23) = 2.19, $P < 0.05$. A 2 (high vs low fantasy prone individuals) × 2 (PSS-SR vs LSHS items) Analysis of Variance (ANOVA) with repeated measures on the last factor showed that high and low fantasy prone participants did not differ with regard to their scores on post-traumatic stress (PSS-SR) or hallucination (LSHS) items, $F(1, 23) < 1.0$. Moreover, both groups had higher scores on PSS-SR items than on LSHS items, $F(1, 23) = 111.40$, $P < 0.01$. The interaction of group and type of item (PSS-SR vs LSHS) failed to reach significance, $F(1, 23) < 1.0$.

Table 1
Mean emotionality, plausibility, and CBCA ratings for fabricated stories of high ($n = 11$) and low ($n = 14$) fantasy prone participants. Mean percentage scores (maximum score = 100) on the Post-Traumatic Stress Symptoms-Self-Report (PSS-SR) and the Launay–Slade Hallucination Scale (LSHS) are also shown. Standard deviations are given between parentheses

	High Fantasy Prone	Low Fantasy Prone
Emotionality*	6.64 (0.67)	5.57 (0.76)
Plausibility*	6.14 (1.27)	5.21 (0.94)
CBCA*	14.00 (4.27)	10.50 (3.73)
PSS-SR	62.75 (14.65)	59.94 (15.05)
LSHS	28.54 (14.65)	30.56 (16.63)

*$P < 0.05$, two-tailed.

The current findings demonstrate that high fantasy prone individuals are better at telling a good story than are low fantasy prone individuals. That is, stories fabricated by high fantasy prones are rated as more emotional and plausible than are the stories of low fantasy prone persons. Most importantly, stories of high fantasy prone participants were richer in CBCA elements than those of low fantasy prone participants. Thus, it seems that fantasy prone people have better storytelling abilities and this enhances the credibility of their stories, at least when the CBCA is used to evaluate credibility. Meanwhile it was not the case that fantasy prone people displayed an overendorsement of symptoms when they were instructed to feign traumatic psychopathology. This indicates that fantasy prone people are not only good at fabricating stories, but are also good at faking a pattern of symptoms that is plausible and lacks an "over-the-top" quality.

Note that differences in emotionality, plausibility, and presence of CBCA elements between stories of fantasy prone and those of control individuals were significant, but by no means dramatic. On the other hand, participants had no time to prepare their fabricated stories and the theme and format of the stories was fixed. Also, our raters were well aware of the fact that all stories were invented. This might have produced the opposite of the well-documented truth bias (DePaulo, Charlton, Cooper, Lindsay, & Muhlenbruck, 1997), i.e., the tendency to believe that others are telling the truth. Literature on the CBCA clearly shows that this set of criteria encourages a truth bias in raters such that CBCA evaluations more often lead to falsely classifying fictitious statements as truthful than to falsely classifying true statements as fabricated (see, for a review, Ruby & Brigham, 1997). However, the relatively low (i.e., close to mid-point) ratings for emotionality, plausibility, and CBCA elements obtained in the current study suggest that, if anything, raters displayed a suspicion rather than a truth bias. Thus, it might well be the case that differences between fantasy prone and control individuals in storytelling abilities would have been more substantial had we allowed participants to prepare stories of both true and fictitious events and without imposing any restrictions as to the story themes. Study 2 was conducted along these lines.

5. Study 2

In this study, undergraduates high or low on fantasy proneness were invited to think for a while about a real and a fictitious event in which they had been the victims of other people's actions. Next, they were instructed to write down the events. As in study 1, independent and blind judges evaluated stories in terms of emotionality, plausibility, and CBCA elements. Given the fact that participants had the opportunity to select and prepare their own stories, we expected that differences between stories of high and low fantasy prone participants would be more substantial than in study 1. Apart from the fact that each participant was instructed to describe a true and a false story, study 2 differed from study 1 in the following ways. To begin with, for psychometric reasons, participants were asked to complete the DES-C (Wright & Loftus, 1999) rather than the DES. Secondly, a subsample of participants completed the CEQ twice: once during a mass testing session and 6–8 months later, just prior to the experiment proper. With these test–retest data we were able to explore whether participants' assignment to the high or low fantasy prone group was valid. Thirdly, unlike study 1, we did not include PSS-SR and LSHS ratings because it would be

unclear whether such ratings would pertain to true or false stories. Fourthly, given that participants were free to select their true and false story themes, a critical factor in study 2 might be the content of the stories. That is, fantasy prones' stories might be more compelling than those of controls because the former tend to come up with more dramatic content. Therefore, study 2 also examined whether fantasy prone and control participants differ with regard to the types of themes they select when describing true and false events.

6. Method

6.1. Participants

During two consecutive mass testing sessions with fresh samples, 342 psychology undergraduate students completed the CEQ (Cronbach's alpha = 0.74) and the DES-C (Cronbach's alpha = 0.94), which in non-clinical populations has a less skewed distribution than the conventional DES (Wright & Loftus, 1999). After 6–8 months, respondents scoring in the highest or lowest deciles of the CEQ distribution were contacted and invited to participate in a follow-up study. As female students were in the majority during our mass testing sessions (80%) and because we wanted to keep our final samples as homogeneous as possible, we decided to include only women. Twenty high fantasy prone participants and 18 low fantasy prone participants were willing to come to the lab. Their mean age was 19.5 years (SD = 1.3 years). Mean CEQ scores of high and low fantasy prone participants were 13.85 (SD = 1.22) and 2.27 (SD = 0.83), respectively: $t(36) = 33.72$, $P < 0.01$. High fantasy prones had significantly higher dissociative levels than low fantasy prone persons, mean DES-C scores being 45.41 (SD = 15.46) and 31.44 (SD = 13.94), respectively: $t(36) = 2.91$, $P < 0.01$. As part of study 2, a subsample of 12 high and 12 low fantasy prone participants completed the CEQ a second time. The test–retest correlation for their CEQ scores was 0.82 ($P < 0.01$) indicating that high and low CEQ scores had remained fairly stable across time.

6.2. Procedure

Participants were given written instructions that asked them to think about an incident that really happened to them and in which they suffered financial, emotional, and/or physical harm as a result of another person's actions. Participants were then asked to write down what happened, the type of damage that occurred, and who they held accountable for the damage. Next, participants were asked to think about a fictitious event in which they sustained damage. The instructions stressed that the event should be plausible, but in no way related to anything they had ever experienced. Again, they were asked to provide written specifications about the nature of the incident, the resulting damage, and the person who was responsible. Within each group, half of the participants started with the true event and then wrote about the fictitious event, while the other half had the reversed order. For both true and false events, instructions stressed that stories should be written in such a manner that they would be convincing to triers of fact in a tort damage procedure.

Stories were presented in a random order to two independent raters (a teacher and a psychologist) who were blind as to the CEQ status of participants and the truth status (i.e., true or false) of the stories. Raters read each story and rated its theme on a 5-point scale (anchors: 0 = *the event involves a minor social incident*; 4 = *the event involves severe physical harm*). They also rated the emotionality and plausibility of each story, using the 10-point scales that were also employed in study 1 (anchors: 1 = *not at all emotional/plausible*; 10 = *extremely emotional/plausible*).

Two other independent and blind judges (a forensic neuropsychologist and a law sociologist) rated each story in terms of the nine CBCA criteria that were also used in study 1. Both judges were familiar with the literature on the CBCA. As well, they were provided with a written definition of each CBCA criterion.

6.3. Results and discussion

Unlike study 1, there were clear differences in verbosity between high and low fantasy prone participants. A 2 (groups: high vs low fantasy prone participants) × 2 (truth status: true vs false) ANOVA with the last factor being a repeated measure yielded a main effect of groups [$F(1, 36) = 4.40$, $P < 0.05$], indicating that high fantasy prones used more words to describe their stories ($M = 174$; SD = 12.4) than did low fantasy prone individuals ($M = 136$; SD = 13.1). All other effects of this ANOVA remained non-significant [all F's < 2.0, all P's > 0.16].

Raters' evaluations of thematic content, emotionality, and plausibility of true stories correlated significantly with each other, with r's varying between 0.29 (emotionality) and 0.83 (thematic content). Much the same was true for raters' content, emotionality, and plausibility ratings of false stories. Here, r's ranged from 0.24 (emotionality) to 0.79 (content). There were not only significant correlations between raters, but in a few instances there were also significant associations between dimensions, suggesting that they were not always independent of one another. For example, the first rater's evaluations of the thematic content of true stories were significantly related to his emotionality ratings of these stories ($r = 0.35$, $P = 0.02$). Nevertheless, ratings of both judges were averaged separately for each dimension (see Table 2).

A set of three 2 (groups) × 2 (truth status) repeated measurement ANOVA's were performed on the averaged content, emotionality, and plausibility ratings. For content evaluations, only a significant main effect of truth status was found [$F(1, 36) = 9.41$, $P < 0.01$], indicating that overall, aversive content of false stories was rated as more serious than that of true stories. For emotionality ratings, the ANOVA did not yield significant main or interaction effects [all F's < 2.19, all P's > 0.11]. For plausibility, only a significant main effect of truth status was obtained [$F(1, 36) = 5.40$, $P < 0.01$], indicating that the plausibility of true stories was rated as higher than that of false stories.

Cronbach's alpha's for the set of 9 CBCA criteria applied by two judges to true and false stories ranged from 0.40 to 0.86. Raters' summed CBCA ratings for true stories correlated significantly with each [$r = 0.55$, $P < 0.01$], as did their summed ratings for fabricated stories [$r = 0.53$, $P < 0.01$]. As in study 1, summed CBCA ratings were averaged across judges. Mean CBCA ratings for true and false stories of high and low fantasy prone participants are shown in Table 2. A 2 (group) × 2 (truth status) repeated measurement analysis yielded only a main effect of group [$F(1, 36) = 4.17$, $P < 0.04$], indicating that overall, stories of high fantasy prone persons were

Table 2
Mean content (range 1–5), emotionality (range: 1–10), plausibility (range: 1–10), and CBCA (range: 0–27) ratings (averaged over two raters) for true and false stories of high ($n = 20$) and low ($n = 18$) fantasy prone individuals. Standard deviations are given between parentheses

	High Fantasy Prone	Low Fantasy Prone
Content		
True	1.83 (1.10)	1.36 (1.33)
False	2.52 (1.10)	2.36 (1.30)
Emotionality		
True	7.10 (0.84)	6.39 (0.90)
False	6.98 (0.84)	6.94 (1.32)
Plausibility		
True	7.53 (0.97)	7.58 (1.35)
False	7.30 (0.88)	6.67 (1.48)
*CBCA**		
True	9.40 (3.11)	7.55 (2.48)
False	10.10 (2.96)	8.44 (3.82)

*$P < 0.05$ for main effect of group.

richer in CBCA elements than those of controls. This significant group difference disappeared when number of words was entered as a covariate into the ANOVA [$F(1, 35) = 1.35$, $P = 0.25$].

The most parsimonious explanation for the pattern found in study 2 is that our instructions to fabricate a tragic story lead both high and low fantasy prone participants to select a more dramatic content for fabricated than for true stories. This was not accompanied by a shift in emotionality. Thus, notwithstanding their more dramatic content, it was not the case that fabricated stories were rated as more emotional than true stories. Possibly, this undermined the plausibility of fabricated stories. As the ANCOVA showed, the fact that both fabricated and true stories of high fantasy prone participants were richer in CBCA elements than those of low fantasy prone individuals was related to the verbosity of high fantasy prone individuals.

We anticipated that by including true and false accounts on a within-subject basis, study 2 would detect larger differences between stories of high and low fantasy prone individuals than did study 1. In fact, quite the opposite happened. Whereas study 1 found stories of high and low fantasy prone persons to differ in emotionality, plausibility, and CBCA richness, study 2 only found a significant difference for CBCA elements. We do not think that the prospect of having to write down two stories undermined participants' motivation to come up with a good story. The fact that stories obtained in study 2 contained a similar number of words as the stories obtained in study 1 is difficult to reconcile with such a motivational interpretation. Instead, we suspect that the lack of emotionality and plausibility effects and the significant, but modest CBCA effects in study 2 are due to subtle carry-over phenomena. That is, the sheer instruction to produce a true and a false story (or vice versa) might lead participants to use uniform standards for stories such that constraints operate on what is written down. This may tone down fantasy prones' stories, making it difficult to detect main effects of or interaction effects with fantasy proneness. Such a phenomenon would be difficult to control for with counterbalancing measures. Perhaps, then, a between-subject design in which some participants are invited to describe a true story, while others are instructed to describe

a false event would have generated stronger and more pervasive effects of fantasy proneness on storytelling. We emphasize this point because some authors (Dahle, 1997) have speculated that the CBCA is better able to determine the truth status of an eyewitness account when baseline data are collected by asking the eyewitness to fabricate an additional account. A CBCA analysis would then involve a within-subject comparison between an eyewitness' pertinent account and his/her fabricated story. However, the current data suggest that at least in the hands of untrained specialists, the CBCA does not allow one to differentiate between true and false stories.

7. General discussion

On theoretical grounds, Pezdek and Taylor (2000) argued that fabricated stories differ from true stories in the amount of schematic and episodic descriptions they contain. More specifically, these authors assume that fabricated stories are dominated by schematic, gist-like information, while true stories contain both schematic information and purely episodic details. This view is not unlike the ideas that formed the basis of the CBCA (Undeutsch, 1982) and other tools (e.g., reality monitoring criteria; e.g., Sporer, 1997) that intend to differentiate between true and fabricated stories. Little empirical research has examined the influence that storytellers' personality characteristics might have on the ability of these tools to discriminate between true and false stories. One potentially relevant characteristic is fantasy proneness. Arguably, fantasy prone people may be expected to be able to fabricate stories that not only contain schema-like information, but also pseudo-episodic details. There is yet another reason why fantasy proneness is a relevant trait in this context. A number of studies have shown that there is a substantial overlap between fantasy proneness and dissociative symptoms (see for reviews Merckelbach, Devilly, & Rasin, 2002; Merckelbach & Muris, 2001). These symptoms are often found in patients who have recovered traumatic memories (e.g., McNally et al., 2000) or patients with a diagnosis of Dissociative Identity Disorder (DID; formerly known as multiple personality disorder; Merckelbach et al., 2002). In these cases, there might be urgent forensic reasons to apply tools like the CBCA.

The current studies evaluated whether the trustworthiness of negative stories varies as a function of storytellers' fantasy proneness. In line with previous studies (e.g., Merckelbach, Muris, et al., 2000; Merckelbach, Rassin, et al., 2000), undergraduates selected for their high score on a fantasy proneness scale also displayed more dissociative symptoms than those selected for their low score on this scale. Both study 1 and study 2 provide evidence for the view that individuals high on fantasy proneness are better at telling stories than people low on fantasy proneness. In study 1, fabricated stories of high fantasy prone participants were found to be richer in CBCA elements than were the fabricated stories of low fantasy prone students. In study 2, both true and fabricated stories of high fantasy prone students received higher CBCA ratings than those of low fantasy prone students. Accordingly, our findings underline Ruby and Brigham's (1997, p. 723) speculation that "a person who is good at telling stories would likely to be judged by the CBCA as more truthful than someone who is not good at storytelling." Our results also accord well with the finding that fantasy prone people are good at simulating certain symptoms (e.g., Merckelbach & Rasquin, 2001).

Admittedly, the differences in CBCA richness between high and low fantasy prones' stories were modest. Furthermore, our raters had not undergone extensive training in using the CBCA. Thus,

one could argue that with appropriate training, they might have evaluated the stories of high fantasy prone and control individuals in a different way than they did now. This point is underlined by the relatively modest Cronbach's alphas that we found for CBCA ratings of some judges and the relatively low inter-correlations between CBCA ratings of different judges. On the other hand, one could reason that our results underestimate the effects of fantasy proneness on storytelling. After all, our samples consisted of relatively homogeneous groups of undergraduate students who are probably versed in writing down stories and who rarely score in the extreme upper ends of the fantasy proneness distribution.

Another limitation of the current studies pertains to the global nature of the emotionality and plausibility ratings that we obtained. The precise stylistic features (e.g., self-references, emotion words) that fantasy prone people use when making up their stories warrant further study. This would require a more fine-grained linguistic analysis of the type developed by, for example, Newman, Pennebaker, Berry, and Richards (2003) in their study on false stories.

In sum, then, the present studies found evidence for the view that persons high on fantasy proneness are better at telling true and false stories than are people low on fantasy proneness. The practical implication of this is that forensic experts who rely on the CBCA are well advised to include measures of fantasy proneness in their evaluations.

References

Bennett, W. L., & Feldman, M. S. (1981). *Reconstructing reality in the courtroom*. London: Tavistock.

Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease, 174*, 727–735.

Brick, S., & Chu, J. A. (1991). The simulation of multiple personalities: A case report. *Psychotherapy, 28*, 267–272.

Candel, I., & Merckelbach, H. (2003). Fantasy proneness and thought suppression as predictors of the medical student syndrome. *Personality and Individual Differences, 35*, 519–524.

Dahle, K. (1997). Wege zu einem linguistischen Wahrheitstest [Approaches to a linguistic truth-test]. *Diagnostica, 43*, 3–26.

DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review, 1*, 346–357.

Dinwiddie, S. H., North, C. S., & Yutzy, S. H. (1993). Multiple personality disorder: Scientific and medicolegal issues. *Bulletin of the American Academy of Psychiatry and the Law, 21*, 69–79.

Foa, E. B., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress, 6*, 459–473.

Launay, G., & Slade, P. D. (1981). The measurement of hallucinatory predisposition in male and female prisoners. *Personality and Individual Differences, 2*, 221–234.

Lees-Haley, P. R. (1989). Malingering post-traumatic stress disorder on the MMPI. *Forensic Reports, 2*, 89–91.

McNally, R. J., Clancy, S. A., Pitman, R. K., & Schacter, D. L. (2000). Personality profiles, dissociation, and absorption in women reporting repressed, recovered or continuous memories of childhood sexual abuse. *Journal of Consulting and Clinical Psychology, 68*, 1033–1037.

Merckelbach, H., Devilly, G., & Rasin, Q. (2002). Alters in dissociative identity disorder: Metaphors or genuine entities? *Clinical Psychology Review, 22*, 481–497.

Merckelbach, H., Horselenberg, R., & Muris, P. (2001). The creative experiences questionnaire (CEQ): A brief self-report measure of fantasy proneness. *Personality and Individual Differences, 31*, 987–995.

Merckelbach, H., & Muris, P. (2001). The causal link between self-reported trauma and dissociation: A critical review. *Behaviour Research and Therapy, 39*, 245–254.

Merckelbach, H., Muris, P., Horselenberg, R., & Stougie, S. (2000). Dissociative experiences, response bias, and fantasy proneness in college students. *Personality and Individual Differences, 26*, 961–967.

Merckelbach, H., & Rasquin, S. (2001). Psychobiological research on DID: A methodological note. *Acta Neuropsychiatrica, 160*, 327–340.

Merckelbach, H., Rassin, E., & Muris, P. (2000). Dissociation, schizotypy, and fantasy proneness in undergraduate students. *Journal of Nervous and Mental Disease, 188*, 428–431.

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.

Myers, S. A. (1983). The Wilson–Barber inventory of childhood memories and imaginings: Children's form and norms for 1337 children and adolescents. *Journal of Mental Imagery, 7*, 83–94.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*, 665–675.

Penington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory-structure on judgment. *Journal of Experimental Psychology: Learning Memory and Cognition, 14*, 521–533.

Pezdek, K., & Taylor, J. (2000). Discriminating between accounts of true and false events. In D. F. Bjorklund (Ed.), *False-memory creation in children and adults* (pp. 69–91). Mahwah NJ: Erlbaum.

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.

Porter, S., Yuille, J. C., & Lehman, D. R. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: Implications for the recovered memory debate. *Law and Human Behavior, 23*, 517–537.

Purisch, A. D., & Sbordone, R. J. (1997). Forensic neuropsychology: Clinical issues and practice. In A. MacNeill Horton, D. Wedding, & J. Webster (Eds.), *The neuropsychology handbook: Treatment issues and special populations* (pp. 309–355). New York: Springer.

Ruby, C. L., & Brigham, J. C. (1997). The usefulness of the criteria-based content analysis technique in distinguishing between truthful and fabricated allegations. *Psychology, Public Policy, and Law, 3*, 705–737.

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.

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

Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences ("absorption"), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology, 83*, 268–277.

Undeutsch, U. (1982). Statement reality analysis. In A. Trankell (Ed.), *Reconstructing the past: The role of psychologists in criminal trials* (pp. 27–56). Stockholm Sweden: Norstedt.

Wilson, S. C., & Barber, T. X. (1983). The fantasy-prone personality: Implications for understanding imagery, hypnosis, and parapsychological phenomena. In A. A. Sheikh (Ed.), *Imagery: Current theory, research, and application* (pp. 340–387). New York: Wiley.

Wright, D. B., & Loftus, E. F. (1999). Measuring dissociation: Comparison of alternative forms of the dissociative experiences scale. *American Journal of Psychology, 112*, 497–519.