

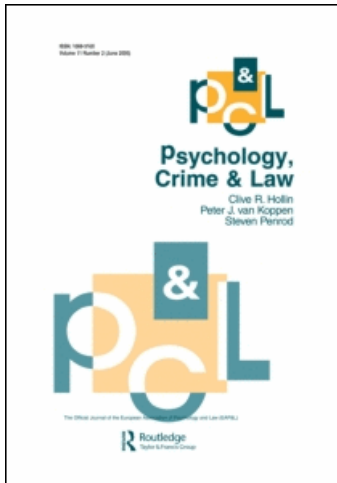
This article was downloaded by: [University of Maastricht]

On: 10 June 2010

Access details: Access Details: [subscription number 781062704]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Psychology, Crime & Law

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713647155>

Differentiating between children's true and false memories using reality monitoring criteria

Henry Otgaar^a; Ingrid Candel^a; Amina Memon^b; Jehanne Almerigogna^b

^a Maastricht University, Maastricht, the Netherlands ^b School of Psychology, University of Aberdeen, Aberdeen, UK

First published on: 10 June 2010

To cite this Article Otgaar, Henry , Candel, Ingrid , Memon, Amina and Almerigogna, Jehanne(2010) 'Differentiating between children's true and false memories using reality monitoring criteria', Psychology, Crime & Law,, First published on: 10 June 2010 (iFirst)

To link to this Article: DOI: 10.1080/10683160902926166

URL: <http://dx.doi.org/10.1080/10683160902926166>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Differentiating between children's true and false memories using reality monitoring criteria

Henry Otgaar^{a*}, Ingrid Candel^a, Amina Memon^b and Jehanne Almerigogna^b

^a*Maastricht University, Maastricht, the Netherlands;* ^b*School of Psychology, University of Aberdeen, Aberdeen, UK*

(Received 20 November 2008; final version received 24 March 2009)

The present study examined the efficacy of the reality monitoring (RM) criteria in differentiating between children's true and false memories. Two independent judges rated 190 transcripts of children's true and false memory reports along the RM criteria. Results showed that, overall, the RM criteria failed to distinguish between children's accounts of true and false events. However, when examining each RM criterion separately, we found that more visual details were present in children's true memories than in their false memories. Results also showed that more RM criteria were present at the second interview than at the first interview.

Keywords: children; false memories; reality monitoring framework

Introduction

Whether children's accounts of events are based on fabrication (e.g. false memories) or true experience has spawned widespread scientific interest (e.g. Strömwall, Bengtsson, Leander, & Granhag, 2004; Vrij, 2008). In legal settings in which the veracity of children's statements is often doubted (see Ceci & Bruck, 1993), determining whether these statements refer to authentic or invented events can be of crucial weight. The purpose of the current study was to examine whether children's true and false memories can be discriminated using the reality monitoring (RM) approach (Johnson & Raye, 1981).

In the past decades, a variety of verbal deception techniques has been developed to assess the authenticity of statements. A frequently used technique in forensic settings is the Criteria-Based Content Analysis (CBCA; Kohnken, Schimossek, Aschermann, & Hofer, 1995; Raskin & Esplin, 1991; Vrij & Mann, 2006). Inspired by children's alleged claims of sexual abuse, the CBCA is based on the Undeutsch hypothesis, which posits that accounts of experienced events about sexual abuse qualitatively and quantitatively differ from accounts of invented events (Steller, 1989; Undeutsch, 1989). Hence, the CBCA concentrates on the presence or absence of 19 criteria (e.g. logical structure, details, corrections) supposed to discriminate between truthful and deceived accounts.

The CBCA, however, has been criticized (see Vrij, 2005). That is, studies have shown that CBCA scores are affected by event familiarity (Blandon-Gitlin, Pezdek, Rogers, & Brodie, 2005; Pezdek et al., 2004), coaching (Vrij, Akehurst, Soukara, &

*Corresponding author. Email: Henry.Otgaar@maastrichtuniversity.nl

Bull, 2004) and that false positives (i.e. wrongfully asserting that an account is true) are extremely likely to occur (Ruby & Brigham, 1997). Furthermore, the CBCA is criticized because it lacks any theoretical grounding (Sporer, 1997).

A more promising instrument to distinguish truthful and fabricated statements is the RM approach. In contrast to the CBCA, the RM method contains solid theoretical underpinnings (Johnson & Raye, 1981; Vrij, 2008). According to this approach, memory characteristics of experienced events qualitatively and quantitatively differ from memory characteristics of fabricated events. Specifically, memories originating from true experiences would include more perceptual information (visual details, sounds, smells, tastes and physical feelings related to the event), contextual information (information regarding when and where the event happened), and affective information (details about feelings or apperceptive reactions to the event) than memories based on fabrication. The latter would consist of more information concerning cognitive operations (details about thoughts, reasoning, and inferences of events) that govern these memories. Over the past decades, researchers have examined whether the RM framework could be used as a tool to detect deception (for an overview see Masip, Sporer, Garrido, & Herrero, 2005; Sporer, 2004). Based on this approach, researchers (e.g. Vrij, Mann, Kirsten, & Fisher, 2007) constructed a list of RM criteria (see Table 1) as a means of assessing the likelihood of a statement being true or false.

However, whether the RM criteria are useful in differentiating between recollections of true events and ‘recollections’ of false events – events that individuals falsely report to have experienced – is to date unclear. According to Vrij (2008, p. 265), ‘the findings to date give only a weak indication that observers can distinguish between true and false memories by employing the RM tool’. Moreover, he (p. 265) argues that ‘research in this area is scarce, and the issue has not yet fully been explored’.

Recent studies focusing on the discriminative power of the RM tool have generally found that it operates above chance level when deciding whether a statement is true or deceptive (see Masip et al., 2005). The standard procedure in these studies is that participants have to describe a true event or have to make up a report about a non-experienced event (i.e. lie). Then, the presence of specific RM criteria is used to examine whether they can discriminate between truth-tellers and liars. Sporer and Sharman (2006), for example, instructed participants to write reports of invented or experienced events. Next, participants had to rate to qualitative details of their own statements using RM criteria. They showed that participants rated statements of experienced events as more qualitatively detailed (more clear, realistic, sensory information, spatial details and temporal information) than invented statements.

Obviously, examining whether the RM tool is valuable in distinguishing true and false memories would be of practical importance. Consider for example studies (e.g. Otgaar, Candel, & Merckelbach, 2008; Otgaar, Candel, Merckelbach, & Wade, 2009; Strange, Sutherland, & Garry, 2006) and legal cases (e.g. ‘McMartin Preschool’, Wee Care Nursery School; Ceci & Bruck, 1993; Garven, Wood, Malpass, & Shaw, 1998) that showed that children are easily led to develop full-blown false memories of non-existing events. Clearly, people involved in legal cases (e.g. interrogators, researchers, parents) might wonder whether children describe a true or a fictitious event. With

this in mind, the present study focused on the potential of the RM technique in distinguishing children's true and false memories.

As far as we know, only two studies to date have examined whether RM criteria are successful in differentiating true and suggested accounts of events. Schooler, Gerhard, and Loftus (1986) were the first to examine the efficiency of the RM approach. In their study, half of the participants viewed a slide in which a yield sign was depicted while the other half did not witness this sign. However, the participants who did not view the sign were led to believe that the original slide contained a yield sign. Results showed that 25% ($n=21$) of these participants endorsed this suggestion. Furthermore, participants were instructed to provide a verbal description about the event. Transcripts of these descriptions were then presented to independent raters who had to decide whether each transcript was based on a true or fabricated experience using RM criteria (e.g. sensory details, cognitive processes). Results showed that transcripts could be accurately split into true or suggested accounts of events (i.e. 62% overall accuracy).

In a related study, Porter, Yuille, and Lehman (1999) led participants to believe they experienced entire false events (e.g. serious animal attack). Specifically, participants were read a true and a false narrative about events they had experienced in their childhood. One of the narratives was invented by the experimenters and described a fictitious event. Next, participants were asked what they could remember about the events. Results indicated that 26% ($n=20$) of the participants assented to the false events. Participants were also instructed to rate their own memories of the events along several dimensions (e.g. confidence, vividness, details, coherence). Interestingly, participants judged their true memories as more vivid (i.e. a RM criterion) than their false memories.

The above-mentioned studies, however, concentrated exclusively on *adult* participants and relied merely on a small subset of the RM criteria. No study up until now has examined whether *children's* true and false recollections can be discriminated using the RM criteria. The present study has the unique opportunity to test the merits of the RM tool by using transcripts of children's true and false memories from studies by Otgaar and colleagues (2008, 2009, under review). In these studies, Otgaar and co-workers implanted several false events (e.g. UFO abduction, receiving a rectal enema) into children's memory using the same procedure. More precisely, children were presented with true and false narratives about events that they supposedly had experienced some years ago. Next, across two interviews, children had to report everything they could remember about the events.

On the basis of the RM approach (Johnson & Raye, 1981), we on the one hand hypothesized that children's true and false memories could be distinguished. Specifically, we hypothesized that children's true memory reports would contain more perceptual, contextual, and affective information than children's false memory reports. In contrast, we predicted that children's false memory statements would include more references to cognitive operations than children's true memory statements.

On the other hand, one could assert that the RM tool is *not* efficient in distinguishing between children's true and false memories. This prediction stems from the underlying nature of true and false memories. According to Pezdek and colleagues (Pezdek, Finger, & Hodge, 1997; Pezdek & Hodge, 1999), false memories are formed because people connect false events with pre-existing schemas (i.e.

organized knowledge structures) and scripts (i.e. typical sequence of events; Fivush, 1997). Thus, when people freely imagine about these events, an increased likelihood exists that details of these pre-existing schemas are transported to these images resulting in highly detailed, coherent and well-integrated false memories. Accordingly, since these memories would be as detailed as true memories, the RM method would not be able to detect any differences between children's true and false recollections. A recent study indeed found that true and false memories were alike in terms of CBCA criteria (Blandon-Gitlin, Pezdek, Lindsay, & Hagen, in press).

We also predicted that at Interview 2, children's true and false memories would include more RM details than at the first interview (T1). The reason for this is two-fold. First, our participants were instructed to think about the true and false events between the two interviews and to report more information during the second interview (T2; see Otgaar et al., 2008, 2009). Hence, children's true and false memory reports could become more elaborate over time resulting in an increase in perceptual, contextual, affective, and cognitive details (see also Bruck, Ceci, & Hembrooke, 2002; Ceci, Huffman, Smith, & Loftus, 1994). Second, children's memory benefits from repeated interviews. That is, children have the tendency to report more details over time (Quas et al., 2007).

Method

True and false memory reports

The current study included 190 transcripts ($n_{T1/false} = 56$, $n_{T1/true} = 54$; $n_{T2/false} = 40$, $n_{T2/true} = 40$) of children ($M_{age} = 9.13$, $SD = 2.16$, range 7–12) from three studies by Otgaar and colleagues (2008, 2009, under review). Although across the studies, 99 and 114 children developed false memories at the first and second interview, respectively, only 94 and 96 transcripts from both interviews were appropriate for this study since some recordings were incomplete or inaudible. In the previous studies, different false events were inserted into children's memory. That is, almost choking on a candy, being abducted by a UFO, being accused of copying off someone else's work, moving to another classroom, receiving a rectal enema, and being stuck with your finger in a mousetrap. Specifically, children were read narratives about a true event and a fictitious event that we supposedly received from their parents. Then, children were encouraged to report everything they remembered about the events across two interviews with one week in between. When children could not remember anything about the events, they were provided with context reinstatement and guided imagery techniques. True events referred mainly to moderate significant events (e.g. going on a holiday, first day at school) and were obtained and confirmed by the children's parents. Children's responses were coded as a false memory when they stated that they remembered the false event and provided additional event-related details.

Design and procedure

This study employed a 2 (Veracity: True vs False) \times 2 (Interview 1 vs 2) between-subject design. We regarded Veracity and Interview 1 and 2 as between-subject factors since our aim was to examine differences in the presence of RM criteria

between true and false memory transcripts and transcripts from Interview 1 and transcripts from Interview 2. Of importance, the majority ($n = 160$) of the transcripts from both interviews were not from the same children. The reason for this is that in our previous studies, some children developed false memories at Interview 1 or Interview 2 while other children falsely remembered the events at both interviews. However, all children in these studies were instructed to think about the true and false events between the two interviews and to come up with more information at Interview 2.

Transcripts of children's true and false memories were coded by two independent raters. Both raters were trained in the use of RM criteria as specified by Vrij and colleagues (2007) and Colwell, Hiscock, and Memon (2002). Specifically, the training was supervised by the third author. The training started with a general introduction into the background literature of Statement Validity Analysis and Reality Monitoring. The raters were also presented with detailed information regarding the coding criteria with examples and definitions to ensure consistency (for a detailed description see Memon, Fraser, Colwell, Odinot, & Mastroberardino, in press). Both raters were blind to the children's experimental condition.

The RM tool distinguishes in total six different types of details. Visual details are defined as any action or thing that was seen in the event (e.g. 'He walked into the room', consists of three visual details: *he*, *walked*, and *room*). Spatial details consist of any information relating to the location and positioning of items (e.g. 'he put it on the table', contains one spatial detail: *on*). Temporal details are defined as any information related to the timing of events (e.g. 'shortly after the man arrived', contains one temporal detail: *shortly after*). Auditory details refer to any speech- or sound-related information (e.g. 'she asked my name', contains one auditory detail). Cognitive operations are defined as suppositions, thoughts, reasoning, and attributions of intent (e.g. 'he didn't make a fuss when he dropped the laptop, he didn't care, it wasn't his business', contains one cognitive operation). Finally, affective details include information about emotion and feelings (e.g. 'it hurt a lot', contains one affective detail). One point was assigned to each detail. A total RM score was computed by summing all details.

As a check on the reliability of the coding system, about one-third ($n = 71$) of all transcripts were coded by both raters. Overall inter-rater reliability using intra-class correlations (ICC) was excellent with $ICC = 0.97$. Inter-rater reliability values for the separate RM criteria varied from 0.79 (auditory) to 0.96 (temporal). Following these analyses, the two coders compared their results and disagreements were discussed.

Results

Response length

Tables 1 and 2 present the means of the RM criteria as a function of true and false memories and time of interview, respectively. We found that age was significantly correlated with total RM scores ($r = 0.22$, $p < 0.01$). Since some researchers have argued that response length of true and false narratives could exert a confounding effect on the discrimination of true and invented accounts (see Masip et al., 2005; Sporer, 2004), we conducted a 2 (Veracity: True vs False) \times 2 (Interview 1 vs 2) between-subjects ANCOVA with age as covariate to examine whether children's true

Table 1. Means and standard deviations (in parentheses) of RM criteria as a function of children's true and false memories.

	RM criteria	True M (SD)	False M (SD)	F	d
Memories ($n = 190$)	Visual	12.26 (6.65)	9.46 (5.52)	10.41*	0.47
	Temporal	1.85 (2.23)	1.89 (2.13)	0.01	0.01
	Spatial	2.05 (1.72)	2.11 (1.78)	0.06	0.04
	Auditory	0.34 (0.85)	0.34 (0.74)	0.00	0.00
	Cognitive	1.46 (1.51)	1.59 (1.75)	0.33	0.08
	Affective	0.14 (0.45)	0.61 (0.67)	32.57*	0.83

* $p < 0.001$

and false accounts significantly differed in terms of the number of words. No significant interaction was observed ($p > 0.05$). However, we found a significant main effect of Interview ($F(1,185) = 7.34$, $p < 0.01$, $r = 0.17$) with statements at Interview 2 ($M = 49.60$, $SD = 28.67$) containing more words than at Interview 1 ($M = 40.30$, $SD = 26.13$). Furthermore, results showed that children's true ($M = 45.43$, $SD = 30.37$) and false statements ($M = 43.03$, $SD = 24.55$) did not significantly vary with respect to the number of words ($p > 0.05$). Therefore, we did not include the response length of children's true and false narratives as a covariate in further analyses.

Total RM scores

To examine whether the RM criteria could discriminate between children's true and false accounts and whether the second interview would contain more RM details than the first interview, a 2 (Veracity: True vs False) \times 2 (Interview 1 vs 2) between-subjects ANCOVA with age as covariate was executed on total RM scores. No significant interaction between children's true and false memories and time of interview emerged ($p > 0.05$). Furthermore, results showed that children's true and false memory reports¹ did not significantly differ on total RM scores ($F(1,185) = 1.71$, NS). As expected, however, we found a significant main effect of time of interview ($F(1,185) = 8.85$, $p < 0.01$, $r = 0.19$) with true and false memory reports

Table 2. Means and standard deviations (in parentheses) of RM criteria as a function of time of interview.

	RM criteria	Interview 1 M (SD)	Interview 2 M (SD)	F	d
Memories ($n = 190$)	Visual	9.99 (6.04)	12.01 (6.38)	7.05**	.39
	Temporal	1.58 (1.89)	2.26 (2.47)	5.86*	.36
	Spatial	1.84 (1.59)	2.43 (1.90)	8.33**	.43
	Auditory	0.24 (0.61)	0.49 (0.98)	4.23*	.30
	Cognitive	1.45 (1.52)	1.63 (1.77)	1.13	.16
	Affective	0.41 (0.67)	0.34 (0.55)	.7	.12

* $p < .01$, ** $p < .001$

containing more RM details at the second ($M = 19.15$, $SD = 10.62$) than at the first interview ($M = 15.51$, $SD = 9.39$)².

RM criteria

A MANCOVA with age as covariate was performed to examine the six RM criteria as a function of children's true and false memories and time of interview. This revealed no significant interaction between children's true and false memories and time of interview for all RM criteria ($p > 0.05$). However, we found a significant main effect of Veracity (true/false; Pillai's Trace = 0.27, multivariate $F(3,186) = 10.93$, $p < 0.01$, partial $\eta^2 = 0.27$) and a significant main effect of Interview (Pillai's Trace = 0.07, multivariate $F(3,186) = 2.30$, $p < 0.05$, partial $\eta^2 = 0.07$). We used the more conservative Pillai's Trace statistic to prevent concerns of sphericity (for more information see Howell, 2002, pp. 519–524). To investigate which RM criterion contributed to the main effects of Memories and Interview, we conducted six follow-up univariate ANCOVAs using Bonferroni correction.

First, we found that children's true memory reports had significantly more visual details than children's false memory reports. Surprisingly, results also showed that affective details were more predominant in children's false memory reports than true memory reports. The other RM criteria (i.e. temporal, spatial, auditory, and cognitive), however, did not distinguish between children's true and false accounts of events (see Tables 1 and 2).

A third key finding was that, at the second interview 1 week later, children's true and false recollections included significantly more visual, temporal, spatial, and auditory information than at the first interview (all $F_s > 1$, see Table 2).

Affective details

Our finding that more affective details were present in children's false memories than in true memories might be explained since 76 (79%) of our false memory reports were about negative events (i.e. receiving a rectal enema, almost choking on a candy, being accused of copying off, being stuck with your finger in a mousetrap) whereas 20 (21%) reports were about a neutral event (i.e. moving to another classroom). Therefore, we examined whether our false negative statements boosted the affective details in our transcripts using a 3 (Veracity: True, False_{negative}, False_{neutral}) \times 2 (Interview 1 vs 2) between-subjects ANOVA. This analysis revealed a significant main effect of Veracity ($F(2,184) = 19.36$, $p < 0.001$, $r = 0.06$). *Post hoc* analyses using Bonferroni correction showed that indeed the false negative accounts contained more affective references ($M = 0.70$, $SD = 0.69$) than the false neutral ($M = 0.30$, $SD = 0.47$) and true accounts ($M = 0.14$, $SD = 0.45$). Thus, the reporting of affective details will vary with the nature of the event. To investigate whether the affective details confounded our results, a 2 (Veracity: True vs False) \times 2 (Interview 1 vs 2) between-subjects ANCOVA with affective details and age as covariate was performed on total RM scores. No significant interaction between children's true and false memories and time of interview emerged ($p > 0.05$). Moreover, children's true and false memory reports did not significantly differ from each other ($p > 0.05$) and transcripts at Interview 2 contained more RM details than at Interview 1 ($F(1,184) = 10.58$, $p < 0.001$, $r = 0.21$).

Discussion

The purpose of the current study was to examine whether the RM tool could discriminate between children's true and false memories. The key finding of our study is two-fold. First, we found that the RM tool as a whole was inefficient in discriminating between children's true and false recollections. However, when considering each RM criterion separately, we showed that children's true statements included more visual information than their false statements. Furthermore, we found that children's false memories contained more affective details than their true memories. However, this finding was likely due to the fact that our false memories predominantly were about negative events (e.g. receiving a rectal enema). Second, our study showed that over time the number of RM details increased significantly. Specifically, we found that, except for the affective and cognitive details, memories contained more RM details after a one week delay.

Our finding that the RM criteria failed to discriminate between children's true and false memories is in contrast with what the RM framework (Johnson & Raye, 1981) would predict. According to this framework, memories about true experiences would contain more perceptual, contextual, and affective details than memories about imagined and non-experienced events. Clearly, it seems that although a bulk of recent studies show that the RM criteria are valuable in discriminating between true accounts of events and invented accounts of events (see Masip et al., 2005), a huge mismatch seems to exist between accounts of deliberately invented events and accounts of suggested events. Indeed, Johnson (1988) also acknowledged that when people give a verbal recount of their memories, this causes memories, especially memories about imagined events, to be more concrete than they really are. A result might be that differences between memories about true and imagined events will decrease.

Hence, the results presented here are broadly consistent with the work of Pezdek and colleagues (Pezdek et al., 1997, Pezdek & Hodge, 1999) and others (e.g. Hyman, Husband, & Billings, 1995; Loftus & Bernstein, 2005) who stated that false memories arise in part because of a mixture of knowledge from related true events with information from suggested events. So, for example, when people imagine about a false event that they could have almost choked on a candy, information about choking (i.e. schemas) and details about what typically occurs during choking (i.e. scripts) is activated and transported to these images. This in turn results in highly detailed and coherent false memories which are experienced as authentic as true memories. This may in part explain why, in the current study, false memories do not differ from true memories in terms of total RM criteria. Indeed, Strange, Hayne, and Garry (2008) recently found that 'once children had developed a false memory, [...] adult raters could not detect any difference in the quality of their true and false reports' (p. 598).

Although we found false memories being highly similar to true memories, true memories contained more visual details than false memories. Obviously, this finding is in accordance with the RM framework. However, studies have revealed mixed results with respect to the number of visual details in true and false memory reports. Specifically, some studies (e.g. Vrij et al., 2008) have found that visual details did not yield significant differences between true and invented reports while other studies did detect differences in visual details between true statements and

fabricated statements (e.g. Strömwall & Granhag, 2005). Clearly, this matter demands further investigation.

Unexpectedly, we found that affective details were more prevalent in children's false memories than in true memories. The RM framework would predict the opposite. This finding, however, might have been confounded because children's false memories referred mainly to negative events (e.g. almost choking on candy) which led to many affective details in their false memory statements. Furthermore, children's true memories were predominantly about neutral events. Indeed, results showed that the amount of affective details was higher in the false memory statements about negative events than in the false memory statements about neutral events and the true memory statements. Thus, it appears that false memories about negative events may appear more truthful, at least in terms of affective criteria.

Interestingly, we also showed that children's memory at Interview 2 consisted of more RM details than at the first interview. As already stated, in the implantation paradigm (e.g. Pezdek & Hodge, 1999; Otgaar et al., 2008; Strange et al., 2006) used in our previous studies, participants are often encouraged to report additional details at the second interview. Although one would expect to find children reporting less over time (e.g. Quas & Schaaf, 2002), this result clearly shows that this process boosts RM details for true and false memory transcripts. So, it is likely that children rehearsed the events between the two interviews, which could have made their memory reports more detailed. Furthermore, recent research shows that repeated interviews are beneficial for children's memory (Quas et al., 2007). Since it is probable that children regarded the false events as a true experience, repeated interviewing could have made children's true and false memories more detailed over time.

Furthermore, the finding that both true *and* false memories contain more RM details at the second interview is particularly interesting as it suggests that both true and false memories are equally regarded as referring to an authentic experience. This is consistent with literature showing that in child abuse investigations, children came up with highly convincing and believable accounts (e.g. Schreiber et al., 2006). Furthermore, our result is in accordance with earlier work showing that repeated recall increases RM details over time (e.g. Alonso-Quecuty & Hernández-Fernaud, 1997; Granhag, Strömwall, & Landström, 2006).

However, at a similar time, the result that more RM details were present at the second than at the first interview has to be interpreted with caution. As stated above, we regarded time of interview as a between-subjects factor because the majority of the transcripts at Interview 1 and 2 did not belong to the same children. Therefore, our purpose was to examine whether transcripts obtained at Interview 2 differ from transcripts at Interview 1 in terms of RM criteria. It might be the case that other factors (e.g. individual differences) could have also contributed to this result. Yet, when we analysed a subgroup of children who had false memories at both interviews, we found an even stronger effect of time of interview, with memory reports at Interview 2 containing more RM details than at Interview 1.

Three potential limitations of the current study need to be mentioned. First, true and false memory reports were drawn from three different studies from Otgaar and colleagues (2008, 2009, under review). Therefore, true and false memory reports were not from the same participants, which could have affected our results. So, differences between true and false memory reports in terms of RM criteria could have been influenced by other factors. Only a few studies implemented true and false memory

reports as a within-subjects factor in their design (e.g. Blandon-Gitlin et al., in press; Porter et al., 1999). Second, children's narratives were rather short, which could have made RM analyses difficult. It could be that children did not mention some details (e.g. auditory details) because they did not regard them as important or because these details were not part of the event. Relatedly, research shows that limited reports are a key feature when children have to describe experienced events, such as sexual abuse (Sjöberg & Lindblad, 2002). Third, research shows that false memories are not only the products of memory distortions, but might also be affected by social influences (Ceci & Bruck, 1993). Therefore, examining memory characteristics between true and false memories using the RM tool could be affected by this issue.

With respect to the practical implications of our study, our results are not encouraging for researchers in forensic settings who want to use the RM tool as an index to assess the veracity of children's questionable accounts. Our results clearly demonstrate that children's false memories are qualitatively and quantitatively similar to true memories in terms of RM criteria. Although some studies show that the RM criteria can differentiate between true accounts of events and intentionally invented accounts of events (see Masip et al., 2005), this seems not to be the case for children's false memories. This finding is important since in legal cases it might be quite exceptional that children deliberately fabricate events. In fact, a wealth of research (e.g. Otgaar et al., 2008, under review; Strange et al., 2006) shows that children are more easily led in *unintentionally* developing false memories for events, therefore rendering the RM tool an inefficient instrument. Moreover, well-known legal cases as the 'McMartin Preschool' and 'Wee Care Nursery' trials (Ceci & Bruck, 1993; Garven et al., 1998) show that children truly can believe that their false memories refer to experienced events. The overarching message is that practitioners in the forensic field should not use the RM tool when children's statements were likely the result of suggestive practices. We agree with Memon and colleagues (in press) in that analysing statements with the RM tool should not be undertaken when no information is present about the circumstances that have led to these statements.

Notes

1. Twenty-nine per cent ($n=28$) of the false memory reports were from children who developed false memories after context reinstatement and guided imagery instructions. A 2 (Veracity: True vs False/Guided Imagery) \times 2 (Interview 1 vs 2) ANCOVA with age as covariate showed that true and false memory reports did not differ on total RM scores ($F(1,116)=1.15$, NS). However, also no main effect of time of interview emerged ($F(1,116)=1.03$, NS).
2. When we performed a repeated measure analysis on children who developed false memories at Interview 1 and 2 ($n=15$), we also found that children's true and false memory reports did not significantly differ in terms of total RM scores ($F(1,28)=0.03$, NS). Moreover, results showed that children's memory reports at the second interview ($F(1,28)=13.82$, $p < 0.001$, $r=0.57$; $M=21.63$, $SD=11.85$) contained more RM details than at the first interview ($M=14.47$, $SD=7.09$).

References

- Alonso-Quecuty, M.L., & Hernández-Fernaud, E. (1997). Tócale otra vez Sam: Repitiendo las mentiras [Play it again Sam: Retelling a lie]. *Estudios de Psicología*, 57, 29–37.
- Blandon-Gitlin, I., Pezdek, K., Lindsay, D.S., & Hagen, L. (in press). Criteria-Based Content Analysis of true and suggested accounts of events. *Applied Cognitive Psychology*.

- Blandon-Gitlin, I., Pezdek, K., Rogers, R., & Brodie, L. (2005). The effects of event familiarity on Criteria-Based Content Analysis ratings: An experimental study. *Law and Human Behavior, 29*, 187–197.
- Bruck, M., Ceci, S.J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review, 22*, 520–554.
- Ceci, S.J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403–439.
- Ceci, S.J., Huffman, M.L.C., Smith, E., & Loftus, E.F. (1994). Repeatedly thinking about a non-event: Source misattributions among preschoolers. *Consciousness and Cognition, 3*, 388–407.
- Colwell, K., Hiscock, C.K., & Memon, A. (2002). Interviewing techniques and the assessment of statement credibility. *Applied Cognitive Psychology, 16*, 287–300.
- Fivush, R. (1997). Event memory. In Cowan N. (Ed.), *The development of memory in childhood* (pp. 139–162). Hove, UK: Psychology Press.
- Garven, S., Wood, J., Malpass, R., & Shaw III, J. (1998). More than suggestion: The effect of interviewing techniques from the McMartin Preschool case. *Journal of Applied Psychology, 83*, 347–359.
- Granhag, P.A., Strömwall, L.A., & Landström, S. (2006). Children recalling an event repeatedly: Effects on RM and CBCA scores. *Legal and Criminological Psychology, 11*, 81–98.
- Howell, D.C. (2002). *Statistical methods for psychology* (5th ed.). Boston, MA: Thomson Learning.
- Hyman, I., Husband, T., & Billings, J. (1995). False memories of childhood experiences. *Applied Cognitive Psychology, 9*, 181–197.
- Johnson, M.K. (1988). Reality monitoring. *Journal of Experimental Psychology: General, 117*, 390–394.
- Johnson, M.K., & Raye, C.L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85.
- Kohnken, G., Schimosek, E., Aschermann, E., & Hofer, E. (1995). The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology, 80*, 671–684.
- Loftus, E.F., & Bernstein, D.M. (2005). Rich false memories: The royal road to success. In A.J.F. Healy (Ed.), *Experimental cognitive psychology and its implications*. Decade of Behavior (pp. 101–113). Washington, DC: American Psychological Association.
- Masip, J., Sporer, S., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach. A review of the empirical evidence. *Psychology, Crime & Law, 11*, 99–122.
- Memon, A., Fraser, J., Colwell, K., Odinet, G., & Mastroberardino, S. (in press). Distinguishing truthful from invented accounts using reality monitoring criteria. *Legal and Criminological Psychology*.
- Otgaar, H., Candel, I., & Merckelbach, H. (2008). Children's false memories: Easier to elicit for a negative than for a neutral event. *Acta Psychologica, 128*, 350–354.
- Otgaar, H., Candel, I., Merckelbach, H., & Wade, K.A. (2009). Abducted by a UFO: Prevalence information affects young children's false memories for an implausible event. *Applied Cognitive Psychology, 23*, 115–125.
- Otgaar, H., Candel, I., Scoboria, A., & Merckelbach, H. (under review). Script knowledge affects the development of children's false memories.
- Pezdek, K., Finger, K., & Hodge, D. (1997). Planting false childhood memories: The role of event plausibility. *Psychological Science, 8*, 437–441.
- Pezdek, K., & Hodge, D. (1999). Planting false childhood memories in children: The role of event plausibility. *Child Development, 70*, 887–895.
- Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G.S., Quas, J.A., Saywitz, K.J. et al. (2004). Detecting deception in children: Event familiarity affects Criteria-Based Content Analysis ratings. *Journal of Applied Psychology, 89*, 119–126.
- Porter, S., Yuille, J.C., & Lehman, D.R. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: Implication for the recovered memory debate. *Law and Human Behavior, 23*, 517–537.

- Quas, J.A., Malloy, L.C., Melinder, A., Goodman, G.S., D'Mello, M., & Schaaf, J. (2007). Developmental differences in the effects of repeated interviews and interviewer bias on young children's event memory and false reports. *Developmental Psychology, 43*, 823–837.
- Quas, J.A., & Schaaf, J.M. (2002). Children's memories of experienced and nonexperienced events following repeated interviews. *Journal of Experimental Child Psychology, 83*, 304–338.
- Raskin, D.C., & Esplin, P.W. (1991). Statement validity assessment: Interview procedures and content analysis of children's statements of sexual abuse. *Behavioral Assessment, 12*, 265–291.
- Ruby, C.L., & Brigham, J.C. (1997). The usefulness of the Criteria-Based Content Analysis Technique in distinguishing between truthful and fabricated allegations: A critical review. *Psychology, Public Policy, and Law, 3*, 705–737.
- Schooler, J.W., Gerhard, D., & Loftus, E.F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 171–181.
- Schreiber, N., Bellah, L.D., Martinez, Y., McLaurin, K.A., Strok, R., Garven, S., & Wood, J. (2006). Suggestive interviewing in the McMartin Preschool and Kelly Michaels daycare abuse cases: A case study. *Social Influence, 1*, 16–47.
- Sjöberg, R.L., & Lindblad, F. (2002). Limited disclosure of sexual abuse in children whose experiences were documented by videotape. *American Journal of Psychiatry, 159*, 312–314.
- Sporer, S.L. (1997). The less traveled road to truth: Verbal cues in detection deception in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology, 11*, 373–397.
- Sporer, S.L. (2004). Reality monitoring and deception of detection. In P.A. Granhag, & L. Strömwall (Eds), *Deception detection in forensic contexts* (pp. 64–102). Cambridge: Cambridge University Press.
- Sporer, S.L., & Sharman, S.J. (2006). Should I believe this? Reality monitoring of accounts of self-experienced and invented recent and distant autobiographical events. *Applied Cognitive Psychology, 20*, 837–854.
- Steller, M. (1989). Recent developments in statement analysis. In J.C. Yuille (Ed.), *Credibility assessment* (pp. 135–154). Dordrecht, the Netherlands: Kluwer.
- Strange, D., Hayne, H., & Garry, M. (2008). A photo, a suggestion, a false memory. *Applied Cognitive Psychology, 22*, 587–603.
- Strange, D., Sutherland, R., & Garry, M. (2006). Event plausibility does not determine children's false memories. *Memory, 14*, 937–951.
- Strömwall, L.A., Bengtsson, L., Leander, L., & Granhag, P.A. (2004). Assessing children's statements: The impact of a repeated experience on CBCA and RM ratings. *Applied Cognitive Psychology, 81*, 653–668.
- Strömwall, L.A., & Granhag, P.A. (2005). Children's repeated lies and truths: Effects on adult's judgments and reality monitoring scores. *Psychiatry, Psychology, and Law, 12*, 345–356.
- Undeutsch, U. (1989). The development of statement reality analysis. In J.C. Yuille (Ed.), *Credibility assessment* (pp. 101–121). Dordrecht, the Netherlands: Kluwer.
- Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law, 11*, 3–41.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester: Wiley.
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004). Let me inform you how to tell a convincing story: CBCA and reality monitoring scores as a function of age, coaching, and deception. *Canadian Journal of Behavioral Science, 24*, 239–263.
- Vrij, A., & Mann, S. (2006). Criteria-based content analysis: An empirical test at its underlying process. *Psychology, Crime & Law, 12*, 337–349.
- Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*.
- Vrij, A., Mann, S., Kirsten, S., & Fisher, R.P. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior, 31*, 499–518.