Accuracy, completeness, and consistency of emotional memories

TOM SMEETS, INGRID CANDEL, AND
HARALD MERCKELBACH
Maastricht University

The American Journal of Psychology
Winter 2004, Vol. 117, No. 4, pp. 595-610

Accuracy, completeness, and consistency of emotional memories

TOM SMEETS, INGRID CANDEL, AND HARALD MERCKELBACH
Maastricht University

Judges and lawyers often consider inconsistent testimonies to be inaccurate. We addressed this assumption by asking undergraduate students on 2 occasions to write detailed accounts of violent movie fragments they had seen. These accounts were evaluated in terms of accuracy, completeness, and consistency. Experiment 1 showed that accounts tended to be accurate. Moreover, first accounts were marginally more complete than second accounts. The number of inconsistencies between the 2 accounts was not significantly related to their accuracy. Experiment 2 sought to replicate these findings using a more emotionally upsetting movie fragment. Results were highly similar to those of Experiment 1 in that accounts tended to be accurate but incomplete. Inconsistencies were not significantly related to the accuracy of participants' accounts. In line with previous research, we found that accounts of emotional events can be highly accurate but tend to be incomplete. More importantly, inconsistencies cannot be seen as valid predictors of testimonial inaccuracy.

Judges and juries often have to rely on eyewitness testimonies in determining the guilt or innocence of suspects. When they are interviewed on several occasions, it is not unusual for eyewitnesses to provide inconsistent details in their accounts of the event they are interviewed about. Nonetheless, when judges, juries, and lawyers cannot verify the accuracy of testimonies against objective facts, they often focus on such inconsistencies. In general, judges, juries, and lawyers tend to believe that inconsistency reflects inaccuracy (Fisher & Cutler, 1995). Inconsistencies involve both contradictions between two or more interviews and accurate details that are recalled on one occasion but omitted on another (Brewer, Potter, Fisher, Bond, & Luszcz, 1999). In either case, triers of fact consider an eyewitness to be inaccurate. Note that this heuristic assumes that there is an intimate association between broadly defined consistency and accuracy. However, there is little evidence to support this heuristic. Fisher and Cutler (1995) showed participants a video of a robbery and then asked them to describe the offender. Several days later, participants were asked once more to describe the offender. These authors found modest

AMERICAN JOURNAL OF PSYCHOLOGY
Winter 2004, Vol. 117, No. 4, pp. 595–609

correlations between consistency and accuracy, with *r*s ranging from .01 to .37. Apparently, inconsistencies in a witness' testimony are not predictive of global accuracy, so jurors should not assume that inconsistencies necessarily reflect inaccurate testimonies. These results are in line with those of Brewer et al. (1999). In their study, participants watched a videotape of a crime, after which they completed a questionnaire containing items about the event. Approximately 2 weeks later, witnesses completed the same questionnaire once more. The correlation between overall consistency and accuracy was .30. The authors concluded that one "cannot claim that consistency is a strong predictor of overall accuracy" (Brewer et al., 1999, p. 311).

The few studies that focused on the consistency–accuracy link defined accuracy in terms of the number of correct details. However, this is only one index of accuracy. Another index is provided by the number of distortions or commission errors. A distortion is a major change in the details of an existing element (e.g., "a red Volkswagen" instead of "a green Volkswagen"), and a commission error is the introduction of an entirely novel element (Gudjonsson & Clare, 1995). Even if an account is consistent and accurate in the sense that it contains no distortions or commissions, it might still be incomplete. Typically, the completeness of an account is undermined by omission errors, that is, information people tend to leave out when reconstructing an event. Contemporary research has shown that witnesses often fail to report important details of highly emotional events. A good example is provided by Bidrose and Goodman (2000), who compared the allegations of four victims of sexual abuse with forensic evidence contained in photographs and audiotaped records of the abuse. Although victims displayed high levels of testimonial accuracy (78.9%), their omission levels were also high (39%), indicating incompleteness.

So far, no study has looked at how completeness of eyewitness accounts relates to their accuracy and consistency. The aim of the current studies was to examine relationships between accuracy, completeness, and consistency of eyewitness accounts. In Experiment 1, participants watched a videotaped crime of an unsuccessful car theft. After 25 min and after 3 weeks, participants were instructed to recall as much information as possible on people and events shown in the video fragment. Because this video fragment may not have elicited strong emotions, Experiment 2 used a more upsetting movie fragment. After 5 min and after a delay of several weeks, participants were asked to give detailed descriptions of people and events they had seen in the video fragment.

EXPERIMENT 1

METHOD

Participants

A total of 44 Maastricht University undergraduates (34 women, 10 men) with a mean age of 20 years (SD = 2.15, range 18–29) were recruited among first-, second-, and third-year psychology students. All participants were tested individually. They were paid or received course credit in return for their participation.

Movie fragment

A fairly realistic videotape fragment of an unsuccessful car theft specifically developed for the purpose of this study served as the to-be-remembered event. Duration of the fragment was approximately 2 min. It showed a disguised thief who, late in the evening, tries to steal a car parked in back of a house, next to another car. Having succeeded in opening the car, the thief enters the car. Seconds later, garden lights turn on, and a man opens the front door of the house and walks outside with a shotgun in his hand. The thief tries to escape but is shot and falls to the ground.

Procedure

Before they saw the video, participants rated their tension level on a 100-mm visual analog scale (VAS; anchors: 0 = *not at all;* 100 = *extremely*). Next, they watched the movie fragment. After this, participants rated the impact of the movie fragment and once again indicated their tension level on 100-mm VASs (anchors: 0 = *not at all;* 100 = *extremely*). As part of a 30-min filler task, they then completed several personality questionnaires, which will not be considered here. Finally, they were asked to provide a detailed written account of the incident displayed in the movie fragment. The instruction was as follows:

> Imagine that the incident you saw in the movie was a real life incident. You are an eyewitness and you have to make a statement at the police station. Write down the incident, that is all events and their order, all persons and their characteristics, and all central and peripheral details. All information might be important to the police. Please, write down everything in as much detail as possible.

After about 3 weeks (m = 21 days, SD = 0.63), participants returned to the lab and were again asked to provide a written account of the movie fragment. The instructions were exactly the same as those used during the first test occasion.

Scoring

During a pilot study, the movie fragment was shown to 7 people, who were instructed to describe it from a police perspective. We made a list of all elements mentioned by the participants. From this list, only items that were mentioned by at least four of the seven participants were selected. The final list contained 43 critical details for the 2-min fragment (a list of these characteristics is available upon request). An independent judge used this list of critical elements to

determine for each participant account the number of correctly reported details, commissions, distortions, and omissions. Thus, the maximum number of correctly recalled details was 43. An index of accuracy was calculated by dividing the number of correctly recalled details by the sum of the number of correctly reported details plus the number of commissions plus the number of distortions. For example, if a witness recalled seven items correctly and made one commission error and one distortion, his or her accuracy rate would be $7/(7 + 1 + 1)$, or .78. Thus, higher scores indicate higher levels of accuracy. Completeness was defined as 43 minus the amount of omissions, divided by 43. For example, if a witness failed to report 15 details (i.e., omission errors), his or her completeness rate would be $(43 - 15)/43$, or .65. Again, higher scores indicate higher levels of completeness. Finally, a measure of inconsistency was constructed. To this end, participants' first and second accounts were compared with each other. Using a broad definition of inconsistencies (Brewer et al., 1999), for each participant we summed the number of direct discrepancies between the two accounts and total number of additions and omissions in the second account relative to the first one. Higher scores indicate higher levels of inconsistency.

Reliability of this scoring procedure was evaluated by having a second observer scoring commissions, omissions, distortions, and inconsistencies in the accounts of 10 randomly chosen participants. Interrater agreement was .98 for commissions, .86 for omissions, .92 for distortions, and .87 for number of inconsistencies.

RESULTS AND DISCUSSION

Manipulation checks

Mean impact rating of the movie fragment was 56.92, $SD = 23.3$. Participants indicated higher levels of tension after exposure to the movie fragment than before the fragment, $t(43) = 2.49$, $p = .017$, with means of 25.3, $SD = 21.1$, and 20.7, $SD = 17.8$, respectively.

Memory performance

Accuracy. Descriptive statistics for written accounts on both test occasions are shown in Table 1. Bonferroni corrected paired t tests ($\alpha = 0.01$) were conducted on number of correctly reported details, commissions, and distortions and on accuracy rates of written accounts obtained during the first and second test occasion. Number of correct details decreased from the first to the second test occasion, $t(43) = 4.18$, $p < .001$. Participants' testimonies were rather accurate, with average accuracy rates of .85 on both accounts, $t(43) = -0.13$, $p > .10$, implying that the testimonies were largely free of commission and distortion errors. Commissions and distortions did not significantly change over time, both ts < 0.96, both ps $> .10$. On the first test occasion, participants produced on the average 3.16 commissions and distortions, of which 2.16 (68%) were mentioned

Table 1. Number of correctly reported details, commissions, distortions, and omissions and accuracy and completeness rates for accounts obtained during test occasions 1 and 2 (standard deviations and range of scores in parentheses)

	Test occasion 1 ($N = 44$)	Test occasion 2 ($N = 44$)
Correctly reported details* (maximum = 43)	20.64 (3.87; 12–28)	18.82 (3.72; 11–28)
Commissions	0.55 (0.59; 0–2)	0.57 (0.76; 0–3)
Distortions (maximum = 43)	3.16 (1.48; 1–6)	2.91 (2.11; 0–10)
Omissions* (maximum = 43)	19.20 (3.71; 12–27)	21.27 (4.08; 14–30)
Accuracy rate	.85 (.07; .67–.96)	.85 (.09; .55–1.00)
Completeness rate*	.55 (.09; .37–.72)	.51 (.09; .30–.67)

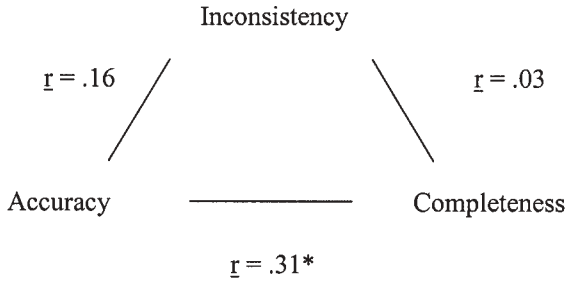
*$p < .01$.

during the second test occasion. Significant Pearson correlations were found between numbers of correctly reported details, $r = .71$; numbers of commissions, $r = .49$; numbers of distorted details, $r = .59$; and accuracy rates, $r = .67$ of both accounts, all ps $< .001$.

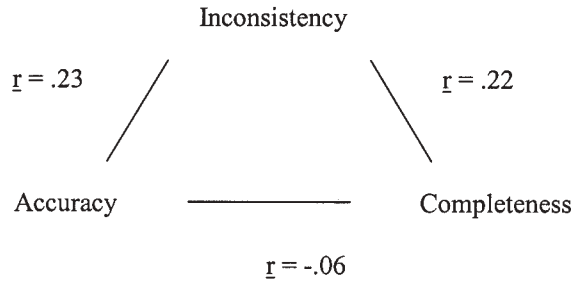
Completeness. Participants made more omission errors on the second than on the first test occasion, $t(43) = -4.02$, $p < .001$, and, as a result, completeness rates decreased over time, $t(43) = 4.02$, $p < .001$. Significant correlations were found between omissions, $r = .62$, and completeness rates, $r = .62$, in both accounts, $p < .001$.

Consistency. Mean number of inconsistencies was 7.55, $SD = 3.03$, range 2–14, of which only 0.80, $SD = 1.00$, range 0–3, pertained to discrepancies between both accounts. Pearson correlations were computed between number of inconsistencies and accuracy and completeness rates of both test occasions. As can be seen in Figure 1, the number of inconsistencies was not related to the accuracy of either the first or the second account: $r = .16$ and $r = .23$, respectively, both ps $> .10$. Additionally, no significant correlations were observed between inconsistencies and completeness on both accounts, $r = .03$ and $r = .22$ for the first and second account, respectively; both ps $> .10$. Moreover, for the first but not the second test occasion a significant but moderate correlation was found between accuracy and completeness, $r = .31$, $p < .05$ and $r = -.06$, $p > .10$, respectively.

The main results of Experiment 1 can be summarized as follows. Overall, accuracy levels were high on both test occasions, with low levels of commission and distortion errors. Furthermore, both accounts tended to be incomplete in that omission levels were high. Additionally, Experiment 1 revealed moderate levels of inconsistencies between the two ac-



*p < .05.



*p < .05.

Figure 1. Experiment 1 Pearson product–moment correlations between accuracy and completeness of the first (upper panel) and second (lower panel) account and inconsistencies between accounts

counts. These inconsistencies were not significantly related to the accuracy of either account, so inconsistencies were not valid predictors of inaccuracies.

EXPERIMENT 2

The stimulus material used in Experiment 1 did not elicit strong emotions. Therefore, one could argue that had we used more provocative

material, we would have found robust links between accuracy, completeness, and consistency. In a second study, we sought to replicate the findings of Experiment 1, relying on a video fragment that was expected to elicit high levels of emotion.

METHOD

Participants

Fifty-two undergraduate psychology students (46 women, 6 men) from Maastricht University with a mean age of 20.3 years (*SD* = 3.22, range 17–33 years) participated in this study. All participants were tested individually. Three to four weeks after the first interview, participants received a questionnaire and a cover letter that invited them to write down the movie fragment they saw in as much detail as possible. Forty-one students (35 women, 6 men) completed and returned this questionnaire (response rate 79%). The mean age of this sample was 20.5 years (*SD* = 3.52, range 17–33). Participants received course credit in return for their participation.

Movie fragment

Participants watched an extremely violent fragment of the movie *American History X,* directed by Tony Kaye (1998). In this fragment, three Black men are stealing a car. The owner, a fascist, shoots two of them and horribly maltreats the third. The fragment is upsetting and contains remarkable details (e.g., tattoos). The duration of the fragment is about 2 min.

Procedure

Participants were invited to come to our lab and were asked to rate their tension and anxiety level on a 100-mm VAS (anchors: 0 = *not at all;* 100 = *extremely*). Subsequently, they watched the movie fragment. Afterwards, participants rated the impact of the movie fragment and again indicated their tension and anxiety level on a 100-mm VAS (anchors: 0 = *not at all;* 100 = *extremely*). Finally, after 5 min they were asked to write a detailed account of the incident portrayed in the movie fragment. The instructions were identical to those given in Experiment 1. We also asked participants whether they had seen the video fragment before. After 3 to 4 weeks (*m* = 26 days, *SD* = 4.5), participants received a questionnaire with a cover letter reminding them of the movie fragment they had seen. Again, they were asked to write down the event and the main characters in as much detail as possible.

Scoring

A detailed scoring guide was developed to evaluate participants' written accounts. To this end, the first and second author independently summarized and selected offender, victim, and event characteristics (e.g., race, hair, tattoo, shooting, breaking glass, arrival of a police car). Only the characteristics that both authors judged to be relevant from a criminological perspective and were clearly observable in the fragment were included in the answer guide. There were 31 such characteristics in the 2-min fragment (a list of these characteristics is avail-

able upon request). An independent and blind judge used this guide to compute the number of correctly reported details, commissions, distortions, and omissions for written accounts obtained during the first and second test occasion. The maximum number of correctly recalled details was, of course, 31. Inconsistencies (i.e., contradictions and omissions or additions) between both testimonies were also computed by this judge. Following the approach of Experiment 1, accuracy, completeness, and consistency rates were calculated.

Scoring reliability was assessed by having a second independent rater assess number of commissions, distortions, omissions, and inconsistencies in 10 randomly chosen written accounts. Across these 10 accounts, interrater agreement was .77 for number of commissions, .73 for number of distortions, .99 for number of omissions, and .74 for number of inconsistencies.

RESULTS AND DISCUSSION

Manipulation checks

Participants gave higher anxiety ratings after exposure to the movie fragment than before the fragment, $t(40) = 4.00$, $p < .001$, with means of 20.71, $SD = 20.68$, and 7.85, $SD = 9.75$, respectively. Much the same was true for tension ratings, $t(40) = 2.76$, $p < .01$, with means of 38.05, $SD = 23.34$, and 27.05, $SD = 20.02$, respectively. Mean rating of the impact of the movie fragment was 79.10, $SD = 13.94$, which is high given the fact that the maximum score is 100. In any case, impact ratings were higher than those obtained in Experiment 1, $t(83) = 5.27$, $p < .001$. Of the 41 participants on both test occasions, 18 were familiar with the movie from which the stimulus fragment was taken. Participants who had seen the movie before did not differ significantly with regard to the accuracy of either account from those who had not seen the movie, both ts < 1.54, both ps $> .10$. Likewise, no significant differences were found for completeness of either account, both ts < 1.0, or for consistency between both accounts, $t(39) < 1.0$.

Memory performance

Accuracy. Descriptive statistics for written accounts on both test occasions are shown in Table 2. Bonferroni corrected paired t tests ($\alpha = 0.01$) were conducted on number of correctly reported details, commissions, and distortions and on accuracy rates of written accounts obtained during the first and second test occasions. Whereas the number of correct details decreased from the first to the second test occasion, $t(40) = 5.72$, $p < .001$, the number of commissions and distortions did not change significantly over time, both ts < 2.1, both ps $> .01$. Overall, participants' testimonies were accurate, with average accuracy rates of .88 for the first account and second account. In other words, few commission and distortion errors were made.

Table 2. Number of correctly reported details, commissions, distortions, and omissions and accuracy and completeness rates for accounts obtained during test occasions 1 and 2 (standard deviations and range of scores in parentheses)

	Test occasion 1 ($N = 41$)	Test occasion 2 ($N = 41$)
Correctly reported details* (maximum = 31)	16.34 (3.99; 7–25)	13.85 (3.51; 5–21)
Commissions	0.56 (0.78; 0–2)	0.71 (1.03; 0–4)
Distortions (maximum = 31)	1.54 (1.00; 0–4)	1.10 (1.02; 0–3)
Omissions* (maximum = 31)	13.12 (3.90; 6–17)	16.05 (3.43; 9–25)
Accuracy rate	.88 (0.08; .69–1.00)	.88 (0.09; .69–1.00)
Completeness rate*	.71 (0.06; .57–.84)	.66 (0.05; .55–.78)

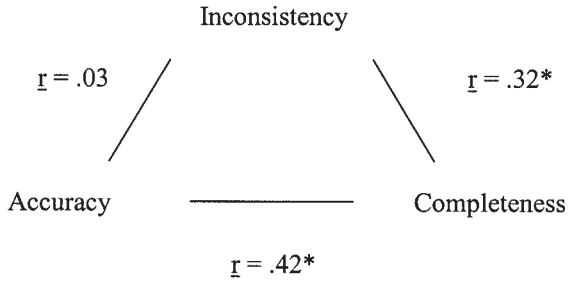
*$p < .01$.

Significant Pearson correlations were found between both accounts in the numbers of correctly reported details, $r = .73$, $p < .001$, and accuracy rates, $r = .47$, $p < .01$. For numbers of commissions and distortions, these correlations remained nonsignificant, $r = .27$, $p = .08$ and $r = .09$, $p = .56$, respectively.

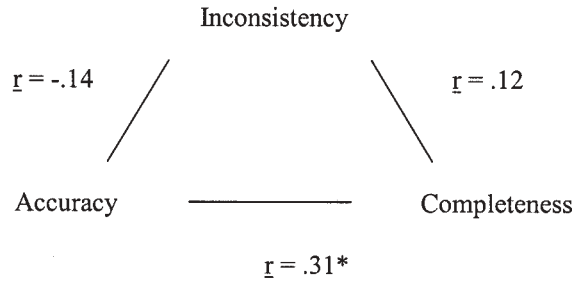
Completeness. Witnesses made more omission errors on the second than on the first test occasion, $t(40) = -6.56$, $p < .001$. As a consequence, the completeness rate decreased over time, $t(40) = 6.68$, $p < .001$. Significant correlations were found between omissions, $r = .70$, and completeness rates, $r = .70$, in both accounts, $p < .001$.

Consistency. Mean number of inconsistencies was 6.13, $SD = 2.41$, range 2–13. Pearson correlations were computed between accuracy rates, completeness rates, and inconsistencies in both test occasions. As can be seen in Figure 2, inconsistencies were not related to the accuracy rate of either the first or the second account, with $r = .03$ and $r = -.14$, respectively, both ps $> .10$. However, for the first but not the second account, a significant but moderate correlation emerged between inconsistencies and completeness, $r = .32$, $p < .05$. Moreover, for both test occasions significant but moderate correlations were found between accuracy and completeness, $r = .42$ and $r = .31$, both ps $< .05$.

Although Experiment 2 used an emotionally provocative event, the results from Experiment 1 were replicated in that participants gave accurate but incomplete accounts of the to-be-remembered event. As was the case in Experiment 1, inconsistencies were rare and were not related to the accuracy of either account.



*p < .05.



*p < .05.

Figure 2. Experiment 2 Pearson product–moment correlations between accuracy and completeness of the first (upper panel) and second (lower panel) account and inconsistencies between accounts

GENERAL DISCUSSION

In two studies, accuracy, completeness, and consistency of memory accounts were evaluated. The main results can be summarized as follows. To begin with, both studies showed that, overall, witness reports were highly accurate. That is, on both test occasions within each study, accounts

were largely free of distortions and commission errors. Second, after a delay of 3 weeks (Experiment 1) or 3–4 weeks (Experiment 2), participants became less complete in their narrative accounts. Third, both studies showed that inconsistencies were not related to the accuracy of either the first or the second account.

The finding that participants were accurate is in line with field studies of Yuille and Cutshall (1986), Bidrose and Goodman (2000), and Peterson, Moores, and White (2001). In the Yuille and Cutshall study, 13 witnesses who observed a shooting incident were interviewed twice about the event. The first interview took place within 2 days after the event, and the second interview was conducted 4–5 months after the event. Accuracy levels ranged from 73% to 88%. As was the case in the current study, Yuille and Cutshall noted that there was almost no change in accuracy over time. Of course, studies documenting stable accuracy rates at high levels contradict older work suggesting that accuracy deteriorates over time (e.g., Barclay & Wellman, 1986). Bidrose and Goodman evaluated testimonies of four female sexual abuse victims against forensic evidence contained in photographs and audiotaped records of the abuse. Again, accuracy rates were found to be high (i.e., 80%). High accuracy rates were also obtained by Peterson et al. These authors repeatedly interviewed children ($N = 96$) about their treatment at a hospital emergency facility. Even at 2-year follow-up, children were remarkably accurate, and confabulations (i.e., commissions) were uncommon. That confabulatory errors in eyewitness accounts are rare is well in line with Dunning and Stern's (1992) and Fisher and Cutler's (1995) findings that commission errors typically account for less than 10% of total memory output.

Despite their high accuracy levels, participants' accounts in the current studies were far from complete. This result is consistent with that of Bidrose and Goodman (2000), who reported that the four victims failed to describe 39% of the sexual acts. It is true that accounts obtained in Experiment 1 were less complete than those obtained in Experiment 2. However, this difference is an artificial byproduct of the way in which completeness was defined. In Experiment 1, participants' accounts had to contain descriptions of 43 separate items in order to be considered fully complete accounts of the event, whereas in Experiment 2 a fully complete account consisted of only 31 items. In both cases, one can argue that conservative measures of completeness were used, that is, measures that reflect upper-bound estimates of completeness of real-life eyewitness accounts.

Participants in our studies became less complete over time, a finding that fits with the traditional Ebbinghaus (1885/1954) view that memory performance deteriorates at a decreasing rate as a function of retention interval length. Interestingly, some studies on repeated eyewitness mem-

ory noted that recall might become more complete with repeated interviews (e.g., Bornstein, Liebel, & Scarberry, 1998; Turtle & Yuille, 1994), a phenomenon called hypermnesia. Although the hypermnesia phenomenon appears to contradict the typical forgetting effect obtained in the current study, Wheeler and Roediger (1992) found evidence that short intervals (i.e., minutes) may lead to hypermnesia, whereas long intervals (i.e., weeks) promote forgetting. Because long intervals are typical for forensic settings, it is intuitively plausible to assume that with repeated interrogations, omissions will increase and completeness of eyewitness accounts will decrease.

The received view among lawyers and laypeople is that inconsistent testimonies reflect inaccurate testimonies (e.g., Brewer et al., 1999; Herlihy, Scragg, & Turner, 2002). Our findings concur with those of Brewer et al. (1999) and Fisher and Cutler (1995) in showing that this popular view lacks firm empirical basis. Thus, two testimonies of the same eyewitness may be both accurate and inconsistent. In the current studies, inconsistent testimonies generally took the form of a correct detail reported during the first test occasion but omitted during the second. This is different from the type of inconsistency that involves contradictions. Contradictions imply that at least in one respect, one testimony must be inaccurate. Thus, by definition, contradictions are related to inaccuracy. Contradictions are caused by commissions or distortions. Our data show that these are extremely rare. Our data also indicate that omissions account for most inconsistencies. Therefore, the typical forgetting effect is the driving force behind inconsistencies. Note that laypeople usually rely on a broad definition of inconsistencies in that they do not differentiate between contradictions and details that are omitted on one occasion. Thus, for example, Berman and Cutler (1996) had their participants view videotaped testimony of an eyewitness and found that any type of inconsistency undermined the perceived credibility of the eyewitness.

As far as we know, only Yuille and Cutshall (1986) looked at the relationship between accuracy and completeness. In their study, the accuracy of reported details was not related to the proportion of correct details, with Pearson correlations ranging from .05 to .23. Our results replicate Yuille and Cutshall's findings in that they show that the link between accuracy and completeness is modest at best (with *r*s ranging from −.06 to .42).

Some comments on the methodological limitations of the present studies are in order. First, our studies relied on samples of university undergraduates whose memory, verbal skills, and motivation may not be representative of the general population. Second, one cannot rule out the possibility that participants had more information in memory than they

provided in their accounts. It is possible that a tendency to answer conservatively, mentioning only details of which one is sure, was responsible for the vast amounts of omissions. Moreover, the information that was omitted may not have been forgotten at all but instead might be temporarily inaccessible (called "transience" by Schacter, 1999) or might never have been encoded in the first place. Also, we acknowledge that simulated videotaped crimes may fall short of many actual eyewitness experiences in eliciting strong emotions. Van der Kolk and Fisler (1995) noted that even watching a movie depicting actual executions did not precipitate posttraumatic stress symptoms in normal college students. On the other hand, not all eyewitness accounts pertain to traumatic events, and even when they do, it is questionable whether witnesses' traumatic memories follow entirely different laws (Shobe & Kihlstrom, 1997; Merckelbach, Dekkers, Wessel, & Roefs, 2003). Moreover, in their recent study, Ihlebæk, Løve, Eilertsen, and Magnussen (2003, p. 325) compared witness accounts of real-life and videotaped events and concluded that "laboratory experiments may overestimate the memory performance of eyewitnesses, but . . . they are otherwise able to simulate essential aspects of the memory performances in naturalistic contexts."

Another limitation of our studies is that we did not systematically look at recall of central versus peripheral details. It has repeatedly been shown (e.g., Burke, Heuer, & Reisberg, 1992; Christianson, 1992; Wessel, van der Kooy, & Merckelbach, 2000) that memory of central details of an emotional event differs from memory of peripheral details. Indeed, a cursory look at our data showed that memory of central details was better than memory of noncentral details. Future research addressing the links between accuracy, completeness, and consistency of emotional memories therefore should take the dimension of central versus peripheral details into account. A final point is that our research relied on free recall. Of course, memory reports obtained during a police interrogation or trial would more resemble cued recall or recognition. Clearly, this point warrants further study.

In sum, then, the results of the current study indicate that accuracy, completeness, and inconsistency are three loosely coupled qualities of eyewitness testimony. The implication of this is that, in contrast to prevailing notions, one cannot rely on inconsistencies to evaluate the accuracy or completeness of testimonies. Our data also show that the more problematic feature of eyewitness testimony is not its lack of accuracy but rather its lack of completeness. Plainly, incomplete testimonies will elicit attempts by the police to make eyewitnesses more complete in their accounts. Such attempts may have the potential to corrupt eyewitness testimony.

Notes

The authors would like to thank two anonymous reviewers for their helpful comments.

Correspondence about this article should be addressed to Tom Smeets, Department of Experimental Psychology, Maastricht University, P.O. Box 616, 6200 MD, Maastricht, The Netherlands (e-mail: tom.smeets@psychology.unimaas.nl).
Received for publication October 30, 2003; revision received March 24, 2004.

References

Barclay, C. R., & Wellman, H. M. (1986). Accuracies and inaccuracies in autobiographical memories. *Journal of Memory and Language, 25,* 93–103.

Berman, G. L., & Cutler, B. L. (1996). Effects of inconsistencies in eyewitness testimony on mock-juror decision making. *Journal of Applied Psychology, 81,* 170–177.

Bidrose, S., & Goodman, G. S. (2000). Testimony and evidence: A scientific case study of memory for child sexual abuse. *Applied Cognitive Psychology, 14,* 197–213.

Bornstein, B. H., Liebel, L. M., & Scarberry, N. C. (1998). Repeated testing in eyewitness memory: A means to improve recall of a negative emotional event. *Applied Cognitive Psychology, 12,* 119–131.

Brewer, N., Potter, R., Fisher, R. P., Bond, N., & Luszcz, A. (1999). Beliefs and data on the relationship between consistency and accuracy of eyewitness testimony. *Applied Cognitive Psychology, 13,* 297–313.

Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory & Cognition, 20,* 277–290.

Christianson, S.-Å. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin, 112,* 284–309.

Dunning, D., & Stern, L. B. (1992). Examining the generality of eyewitness hypermnesia: A close look at time delay and question type. *Applied Cognitive Psychology, 6,* 643–657.

Ebbinghaus, H. (1954). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover. (Original work published 1885)

Fisher, R. P., & Cutler, B. L. (1995). The relation between consistency and accuracy of eyewitness testimony. In G. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Law and criminal justice: International developments in research and practice.* Berlin: De Gruyter.

Gudjonsson, G. H., & Clare, I. C. H. (1995). The relationship between confabulation and intellectual ability, memory, interrogative suggestibility and acquiescence. *Personality and Individual Differences, 19,* 333–338.

Herlihy, J., Scragg, P., & Turner, S. (2002). Discrepancies in autobiographical memories: Implications for the assessment of asylum seekers. Repeated interviews study. *British Medical Journal, 324,* 324–327.

Ihlebæk, C., Løve, T., Eilertsen, D. E., & Magnussen, S. (2003). Memory for a staged criminal event witnessed live and on video. *Memory, 11,* 319–327.

Kaye, T. (1998). *American History X.* New Line Productions.

Merckelbach, H., Dekkers, T., Wessel, I., & Roefs, A. (2003). Amnesia, flashbacks, nightmares, and dissociation in aging concentration camp survivors. *Behaviour Research and Therapy, 41,* 351–360.

Peterson, C., Moores, L., & White, G. (2001). Recounting the same events again and again: Children's consistency across multiple interviews. *Applied Cognitive Psychology, 15,* 353–371.

Schacter, D. L. (1999). The seven sins of memory. Insights from psychology and cognitive neuroscience. *American Psychologist, 54,* 182–203.

Shobe, K. K., & Kihlstrom, J. F. (1997). Is traumatic memory special? *Current Directions in Psychological Science, 6,* 70–74.

Turtle, J. W., & Yuille, J. C. (1994). Lost but not forgotten details: Repeated eyewitness recall leads to reminiscence but not hypermnesia. *Journal of Applied Psychology, 79,* 260–271.

Van der Kolk, B. A., & Fisler, R. (1995). Dissociation and fragmentary nature of traumatic memories: Overview and exploratory study. *Journal of Traumatic Stress, 8,* 505–525.

Wessel, I., van der Kooy, P., & Merckelbach, H. (2000). Differential recall of central and peripheral details of emotional slides is not a stable phenomenon. *Memory, 8,* 95–109.

Wheeler, M. A., & Roediger, H. L. III. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science, 3,* 240–245.

Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology, 71,* 291–301.