

How Eyewitnesses Talk about Events: Implications for Memory

ELIZABETH J. MARSH^{1*}, BARBARA TVERSKY²
and MICHAEL HUTSON¹

¹*Duke University, USA*
²*Stanford University, USA*

SUMMARY

Eyewitnesses to traumatic events typically talk about them, and they may do so for different reasons. Of interest was whether qualitatively different retellings would lead to differences in later memory. All participants watched a violent film scene; one third talked about their emotional reactions to the film (as one might do when talking to a friend), one third described the events of the film (as the police might request), and one third did unrelated tasks. Following a delay, all participants were tested on their memories for the clip. Talking about emotions led to better memory for one's emotions, but also led to subjectivity and a greater proportion of major errors in free recall. Differences were minimized on tests providing more retrieval cues, suggesting that retellings' consequences for memory are greater when retellers have to generate their own retrieval structures. Copyright © 2005 John Wiley & Sons, Ltd.

Eyewitnesses to crimes are likely to tell others about what they experienced. This is almost guaranteed if the crime is reported to the police; the police are likely to repeatedly interview the witnesses to gather as many clues as possible. If the case reaches the legal system, then lawyers will further question the witnesses before they eventually testify in court. But witnesses are likely to retell the crime even outside of the context of the legal system, for many reasons. The witnesses may have intrusive memories of the event (Berntsen, 2001; Brewin, Christodoulides, & Hutchinson, 1996), and retell the memories with friends, family, or health care professionals in an attempt to understand and come to terms with the events. Retellings to bear witness to events are likely to be different from retellings to cope with memorial consequences of witnessing. Of interest is whether factual retellings differ from emotional ones, and whether the different ways that eyewitnesses may talk about violent events have consequences for memory.

It is well known that eyewitnesses are susceptible to post-event misinformation (see Ayers & Reder, 1998, for a review). That is, erroneous information encountered since viewing the original event affects eyewitnesses' performance on later memory tests.

*Correspondence to: Dr. E. J. Marsh, Psychology & Brain Sciences, Duke University, 9 Flowers Drive, Campus Box 90086, Durham, NC 27708-0086, USA. E-mail: emarsh@psych.duke.edu

Contract/grant sponsor: Interval Research Corporation.

Misinformation may come from many sources; these include presuppositions in questions about the original event (e.g. Loftus & Palmer, 1974) or in a narrative description of the events (e.g. McCloskey & Zaragoza, 1985; Tversky & Tuchin, 1989). Misinformation can come from a defense lawyer (Dodd & Bradshaw, 1980), from other witnesses of the same crime (e.g. Gabbert, Memon, & Allan, 2003), or from the witness himself or herself (Roediger, Jacoby, & McDermott, 1996; Zaragoza, Payment, Ackil, Drivdahl, & Beck, 2001).

The current research investigates a different mechanism that could influence eyewitness accounts, namely the consequences of the way in which eyewitnesses describe the crime when retelling it to others. Like the misinformation effect, this is a post-encoding effect. However, unlike the misinformation effect, no specific error is suggested to witnesses. Nor do witnesses necessarily generate errors in their retellings. Instead, the act of retelling sets up a schema for events that guides later retrieval and performance on memory tasks. As events are forgotten, participants rely on the schema, leading to schema-consistent intrusions and schema-inconsistent omissions. Such effects of retellings have been documented in other domains; for example, students who wrote biased letters about a story character later remembered more perspective-consistent information, and made perspective-consistent errors—even though the perspective was not present at encoding nor explicitly reinstated at retrieval (Tversky & Marsh, 2000).

A search of the literature revealed a single published paper manipulating eyewitness retellings. In that study, retelling instructions that encouraged rehearsal of (previously suggested) misinformation increased errors on a later test (Lane, Mather, Villa, & Morita, 2001). Again, however, we are more interested in effects of retelling or rehearsal *without* necessarily combining it with misinformation. The closest phenomenon in the eyewitness domain is the verbal overshadowing effect, whereby witnesses who described a perpetrator were later worse at selecting the target from a photo line-up (Schooler & Engstler-Schooler, 1990). Paralleling work on retellings, a recent meta-analysis suggests that elaborative descriptions are more likely to yield the verbal overshadowing effect than are descriptions following standard free recall instructions (Meissner & Brigham, 2001).

Although the verbal overshadowing effect is an example of description affecting memory, it does not answer the question of how different kinds of rehearsal may differentially affect memory. Eyewitnesses are likely to talk about events differently depending on such factors as the audience and goal of retelling (e.g. see Hyman, 1994; Marsh & Tversky, 2004; Tversky & Marsh, 2000; Wade & Clark, 1993). For example, recounting events to an authority figure or in response to a request for detail and accuracy would likely minimize inclusion of personal reactions to an event and would instead focus on perceptual and temporal details. In contrast, an account directed to a friend or therapist is likely to be qualitatively different, focusing instead on emotional reactions and fears about the witnessed event rather than the facts. This analysis is supported by Hyman's (1994) finding that when students retold events to peers, they were more likely to include their own evaluations than when they retold to an experimenter.

What are the likely consequences of emotional retellings for memory for original events? Previous research supports two possibilities. One possibility, which we call the *Emotional Retelling as Deep Encoding hypothesis*, is that emotional retellings should have positive consequences for memory. The other possibility, which we call the *Emotional Retelling as Selective hypothesis*, is that emotional retellings will have selective effects on emotional aspects of memory. Both of these hypotheses will be described in detail.

Data from several different literatures converge to predict that emotional retellings will aid memory, the Emotional Retelling as Deep Encoding hypothesis. When telling with an emotional focus, the teller is likely to retrieve the most important events, and to relate them to oneself. Both behaviours are likely to have positive consequences for memory. First, many positive effects of rehearsal have been obtained in laboratory studies. Rehearsal is a retrieval test, and testing generally aids memory (e.g. Glover, 1989). In some circumstances, testing leads to hypermnesia, or recall of previously unrecalled information (Payne, 1987). An emotional focus also involves talking about one's reactions to events, and relating events to oneself is a particularly deep form of processing (*the self-reference effect*; Rogers, Kuiper, & Kirker, 1977; Symons & Johnson, 1997). Processing information in relation to oneself increases elaboration and provides organization, two factors known to aid memory (indeed, these factors are the bases of many mnemonics). In short, the data on rehearsal, hypermnesia, and the self-reference effect converge to predict that emotional retellings will be particularly powerful rehearsals when the to-be-remembered events are emotional in nature.

Another possibility is that retellings focused on emotions may lead to selective memories in the same way that other types of biased retellings affect memory (e.g. Dudukovic, Marsh, & Tversky, 2004; Tversky & Marsh, 2000). This is the Emotional Retelling as Selective hypothesis. Support for this possibility comes from work by Johnson and colleagues. In one series of studies, people participated in mini-events, such as drinking a cup of coffee, and rehearsed either the perceptual or the emotional aspects of those events. Focusing on one's thoughts and emotional reactions to the mini-event led to a decrease in the self-rated perceptual and contextual qualities of the memory trace (Suengas & Johnson, 1988). Similarly, participants who participated in a play later rehearsed the perceptual or the emotional content of the play, with consequences for final free recall. Participants who had talked about the emotional content of the play recalled less play information, especially less objective information. The affective retelling focus also led to more inferences and elaborations in final recall (Hashtroudi, Johnson, Vnek, & Ferguson, 1994). In these studies, however, the to-be-remembered events were non-emotional in nature. Even if emotional retellings have negative consequences for non-emotional events, this does not mean that emotional retellings will have a negative impact upon memory for emotional events. It is quite possible that emotional retellings will only be associated with negative consequences if to-be-remembered events are non-emotional.

The current research examined the consequences of emotional versus factual retellings within the eyewitness domain—a domain in which to-be-remembered events are often emotional, and a domain in which accuracy is desired. Participants viewed a sequence of violent scenes from a gangster movie that portrayed the violent murders of a family. After viewing the film, one third of participants were instructed to describe the events of the scene as accurately as possible (Factual focus condition), one third were instructed to describe their emotional reactions to film events (Affective focus condition), and one third did an unrelated task and never discussed the film clip (No-Talk control condition). After a delay filled with irrelevant tasks, participants completed a number of memory tasks: they described the film perpetrators, wrote everything they could recall from the clip, recalled their emotional reactions to the film, and answered a series of cued-recall and picture recognition questions. Of interest is how retelling perspective affects subsequent memory for the witnessed events.

METHOD

Participants

Potential participants were recruited in response to fliers. Prior to the experiment, they filled out a short questionnaire via e-mail. They indicated their age, native language, and which of a list of five movies they had previously seen. Invitations to participate in the study were issued only to respondents who had not seen the critical movie, who were at least 18 years of age, and who were native speakers of English. Of eligible respondents, 77 students participated in the experiment in exchange for monetary compensation. Each participant was tested individually in a single 1.25 hrs session. One participant withdrew consent while viewing the film clip, and another participant deviated from experimental protocol. Thus, the data from 75 participants will be presented here, from 42 females and 33 males.

Materials

Participants watched a 7 min segment from the film *The Professional*. Selected scenes portrayed a violent sequence of mob-style murders. The clip was chosen to fit three criteria. First, the clip was disturbing to watch; pre-testing with Stanford graduate students indicated that watching the selected scenes significantly lowered their self-rated well-being. Second, the clip contained not only graphic violence but also showed emotionally-engaging events such as the near-stabbing of a small child, a psychotic criminal on drugs, and the repeated shooting of a dead body. Third, although the clip was disturbing, it came from a major motion picture released with an 'R' rating, and thus we felt comfortable showing it to our participants as it seemed similar to the type of film often viewed by our participant population. Consistent with this, we found that many potential participants had indeed seen the film and were thus ineligible for the study.

Procedure

Study phase

All participants signed an informed consent form that noted the possibility of watching a violent film. The 7 min film clip was then shown; the experimenter waited outside the lab room while the film was being shown. After 7 min had elapsed, the experimenter returned to the room and immediately gave the participant two questionnaires. The participant first rated how each of 32 adjectives described his or her current mood. The participant also indicated whether or not he or she had previously seen or heard of the film.

Retelling phase

Participants were assigned to one of three conditions: the Factual focus condition, the Affective focus condition, or the No-Talk control condition. The 25 participants in the No-Talk condition did not talk about the film; they did an unrelated engaging task for 7 min that should have prevented reminiscing about the film. Participants in the two experimental conditions were instructed to talk to a video camera about the film for up to 7 min; the 7 min upper-limit was determined in pre-testing and was chosen to allow participants to finish but to avoid leaving extra time that participants might feel they had to fill. Participants spoke to a video camera as opposed to an experimenter as even subtle non-verbal signals can change retellings (Pasupathi, Stallworth, & Murdoch, 1998). Factual

focus participants were instructed: 'I'd like you to talk about what happened in the film. Try to talk about the events in the film in the order in which they occurred and in as much detail as you can, not leaving anything out, so that somebody who has not seen the film can imagine exactly what happened.' Affective focus participants were instructed: 'I'd like you to talk about your emotional reactions to the film. Try to talk about your thoughts and feelings about the film in the order in which they occurred and in as much detail as you can, not leaving anything out, so that somebody who has not seen the film can imagine exactly how you felt.' The remainder of the experiment was the same for all participants, regardless of experimental condition.

Memory phase

Following a filled 25 min delay, each participant completed five memory tasks: a perpetrator recall task, free recall of the film, an emotion memory test, cued-recall questions, and a picture-recognition test. Each of these tasks will be described below. All of the tasks but the free recall task were paper-and-pencil tasks; free recall was typed into the computer.

First, participants were allowed up to 10 min to describe the film's perpetrators. For each perpetrator recalled, the participant gave him a verbal label for later reference, and then described his age, race, hairstyle and colour, body build, clothes, and any distinguishing characteristics.

Second, each participant recalled as much of the film clip as possible; they were given up to 15 min for this task. Participants were instructed to describe the film clip in as much detail as possible, including the action, movement, and dialogue of the clip. They were told to describe the events of the clip in the order in which they happened, being sure to include who did what and where they did it. Thus, recall instructions were fact-oriented and contained no mention of emotions.

Third, each participant recollected how he/she felt after watching the film clip, making responses to the same 32 adjectives (which were presented in a different order than they had appeared on the post-film test). Participants were instructed to think back to how they felt immediately after watching the film clip, and to rate how they thought they felt after watching the clip. Care was taken to ensure that participants understood the retrospective nature of the task, and did not mistake the task as a rating of current mood.

Fourth, each participant answered 23 open-ended questions. The questions asked about specific events of the film, in chronological order. Participants were queried about a variety of film details, including perceptual characteristics (e.g. 'What color were the pills in the pill case?'), actions (e.g. 'What did the father do right before the perpetrators broke into the apartment?'), location (e.g. 'Where was the little boy hiding?'), source attributions (e.g. 'Who stabbed the mattress?') and temporal information (e.g. 'Who was shot first?'). In general, the questions took the participants through the film clip and required them to remember very specific details from the clip. After writing a brief answer for each question, participants rated their confidence in each response using a 5-point scale.

The fifth and final memory test was a 14-item picture recognition test. All questions had a 'whodunit' format; participants were asked to select which of six perpetrators were responsible for a particular action. Participants answered by selecting one of the six perpetrator's pictures (labeled 1-6); these photos were stills taken directly from the video clip that participants had just seen. Some of the picture recognition questions overlapped with the questions asked in the Cued Recall test. Participants again rated their confidence in each answer.

Following the memory tests, all participants were thanked, paid, and debriefed fully. The debriefing procedure involved not only a thorough explanation of the experiment, but also involved an informal assessment of the participants' well-being. Cookies were provided during debriefing in order to ensure all participants left the experiment in a good mood. Participants were also offered the opportunity to watch a section from the comedy *When Harry Met Sally* in order to help them leave the experiment in a good mood.

RESULTS

Retellings

To gain insight into the qualitative differences between emotional and factual retellings of violent events, transcripts of the retellings were analysed using Pennebaker's LIWC (Linguistic Inquiry and Word Count; see Pennebaker & Francis, 1996) program, which records word frequencies for a set number of emotional, cognitive, and linguistic categories. Each category consists of words which the LIWC program searches for and counts; for example, the Insight scale includes 116 words such as 'realize, see, understand' which suggest a person is searching to explain (and gain insight into) a situation. Only categories hypothesized to yield differences between the two retelling conditions were examined. Analysed counts included linguistic factors (overall word count, number of question marks), verb tense (past, present, future), perceptual references (seeing, hearing, motion, time, space, number), emotional references (general affect, positive emotions and feelings, optimism, negative emotions, anxiety, anger, sadness), personal references (I, self, you, other, humans), and references to cognitive processes (cognitive mechanisms, causation, insight).

Overall, the differences between Affective and Factual retellings were striking. An Affective focus led to an emphasis on oneself and one's reactions and feelings; a Factual focus led to a greater emphasis on perceptual information and others (such as the film characters) rather than oneself. The relevant data are shown in Table 1.

The two types of retellings did not differ significantly on such linguistic factors as number of words and question marks, or tense of retelling.

However, Factual retellings included more references to other people (as opposed to oneself), $t(48) = 9.6, p = 0.000$, and to descriptions focused on the perceptual qualities of the remembered event. Factual retellings contained more references to auditory [$t(48) = 4.26, p = 0.000$], spatial [$t(48) = 5.71, p = 0.000$], and numerical information [$t(48) = 6.33, p = 0.000$]. There were trends for Factual retellings to contain more of other types of perceptual information (e.g. temporal and motion), but these differences were not significant when the overall error rate was adjusted to 0.0019 (due to the number of comparisons being made).

Affective retellings were qualitatively different from Factual retellings. Affective retellings contained more language (such as verbs like *think, believe, wonder*) suggesting that participants were attempting to understand the situation, $t(48) = 7.87, p = 0.000$. Affective retellings also included more emotion words [$t(48) = 6.14, p = 0.000$], both positive [$t(48) = 5.06, p = 0.000$] and negative [$t(48) = 4.04, p = 0.000$]. Finally, Affective retellings included more personal references to oneself [$t(48) = 7.79, p = 0.000$] rather than to others.

Table 1. LIWC word proportions for Factual and Affective retellings

	Factual	Affective	<i>p</i> -value
Linguistic factors			
Word count	708.6	591.3	<i>p</i> > 0.1
Question marks	1.6	3.2	<i>p</i> > 0.2
Verb tense			
Past	6.4	9.1	<i>p</i> = 0.01
Present	8.2	7.2	<i>p</i> > 0.2
Future	0.2	0.4	<i>p</i> = 0.003
Perceptual references			
Seeing	1	1.3	<i>p</i> > 0.1
Hearing	1.1	0.5	<i>p</i> = 0.000*
Motion	1.9	1.4	<i>p</i> < 0.03
Time	3.5	2.7	<i>p</i> < 0.04
Space	3.5	1.9	<i>p</i> = 0.000*
Number	1.5	0.6	<i>p</i> = 0.000*
Emotional references			
Affect	3.3	5.8	<i>p</i> = 0.000*
Positive emotions	1.3	2.5	<i>p</i> = 0.000*
Positive feelings	0.5	0.7	<i>p</i> < 0.07
Optimism	0.1	0.3	<i>p</i> < 0.03
Negative emotions	2.0	3.3	<i>p</i> = 0.000*
Anxiety	0.5	1.3	<i>p</i> = 0.000*
Anger	1.0	1.5	<i>p</i> < 0.08
Sadness	0.16	0.1	<i>p</i> < 0.03
Personal references			
I	1.6	6.0	<i>p</i> = 0.000*
Self	1.8	6.1	<i>p</i> = 0.000*
You	0.3	0.9	<i>p</i> = 0.004
Humans	4.3	2.3	<i>p</i> = 0.000*
Other	8.7	3.4	<i>p</i> = 0.000*
Cognitive processes			
Cognitive mechanisms	4.3	8.1	<i>p</i> = 0.000*
Causation	0.7	1.6	<i>p</i> = 0.000*
Insight	1.2	2.8	<i>p</i> = 0.000*

**p*-values are significant when the error rate is adjusted for multiple comparisons (*p* < 0.0019).

In short, an Affective focus led to the use of emotion words whereas a Factual focus led to the use of perceptual terms. Now we turn to effects of focus of retelling on later memory for events.

Memory tests

Perpetrator descriptions

Participants in all three conditions recalled an equal number of perpetrators ($M = 4.5$, $F < 1$), and they did not differ significantly in the number of errors they made when describing the characters ($M = 3.4$, $F < 1$).

Free recall

Recall transcripts were coded for mention of key film events, detail of recalled events, errors in recall, use of quotations, and references to thoughts and feelings. Free recall data

Table 2. Average characteristics of free recall as a function of prior retelling condition

	Factual	Affective	No Talk
Proportion events	0.64 ^a	0.65 ^a	0.57 ^b
Average detail rating	2.4	2.5	2.5
Number of errors	3.9	3.7	3.8
Proportion of errors called major	0.23 ^a	0.45 ^b	0.33 ^{ab}
Number of words	442 ^a	517 ^b	388 ^c
Cognitive words (LIWC)	4.3 ^a	5.6 ^b	4.6 ^a
References to self (LIWC)	0.91 ^a	2.4 ^b	2.2 ^b
Subjectivity score	2.7 ^a	4.4 ^b	3.6 ^{ab}
% Participants quoting	36% ^a	64% ^b	36% ^a

Note: For each dependent measure, items with different superscripts are significantly different from one another; items with the same superscript are not significantly different from one another.

are summarized in Table 2. One coder coded all transcripts; two additional coders each coded half of the transcripts. The third author resolved discrepancies between coders.

To code number of movie events recalled, three research assistants independently listed film events until they converged upon a master list of 30 events. A different set of research assistants actually coded the protocols. For each of the 30 events, they recorded whether or not the event was included in the recall protocol, and rated the detail of that recall using a 3-point scale. In addition to coding the presence/absence of the 30 events, coders noted any errors in the protocol; errors included both major and minor distortions of the 30 key events, distortions of other less key events, and intrusions of non-film events. Coders also recorded references to thoughts and feelings, and use of quotations, to measure how personally involved participants were when recalling the film clip. The subjective composite score, based on Hashtroudi et al. (1994), was a count of personal opinions, thoughts, emotions, and ideas in the free recall protocols. The quotation count was the number of attempted quotations, regardless of the accuracy of those quotations.

We first examined the number of film events recalled across conditions. Although there was no difference between the proportion of events recalled after Factual ($M = 0.64$) versus Affective focus ($M = 0.65$), retelling did lead to greater recall ($M = 0.65$) than in the No Talk condition ($M = 0.57$), $t(73) = 2.02$, $SEM = 0.03$, $p < 0.05$. Given that events were recalled, however, there were no differences across conditions in detail ratings.

There were, however, other differences in the accuracy of free recall as a function of retelling condition. Although the raw number of errors did not differ across conditions ($M = 3.8$), the nature of those errors did. Errors were classified as minor (e.g. colour of clothing was incorrectly reported) versus major (e.g. the wrong person was reported as having been shot), and converted to proportions. The important result is that the proportion of errors considered major differed as a function of retelling condition, $F(2, 71) = 5.31$, $MSE = 0.06$, $p < 0.01$. Prior Factual focus reduced the proportion of subsequent major errors ($M = 0.23$) compared to Affective focus ($M = 0.45$), $t(48) = 3.14$, $SEM = 0.07$, $p < 0.01$. In addition, there was a trend for Affective focus to increase the proportion of errors considered 'major' ($M = 0.45$) over the baseline No-Talk condition ($M = 0.33$), $t(47) = 1.63$, $p = 0.11$.

LIWC analyses of free recall protocols yielded far fewer differences than the retellings. Unlike retellings, there were no significant differences across conditions in frequency of perceptual or emotional references. This makes sense, as all groups had a factual focus for

free recall. The similarity of free recall protocols is further evidence that people can speak flexibly about the same events. However, even though participants were able to follow instructions at recall, differences in free recall remained that were consistent with prior retelling perspective. Perhaps most interesting was that participants who retold with an Affective focus later wrote longer protocols when recalling the film clip—even though they did not recall a greater number of film events nor did they recall events in greater detail. Given this, what additional content would have led to the longer protocols in the Affective condition? The LIWC analysis suggests that people in the Affective condition included more subjective content in their recall protocols. Participants in the Affective condition made more references to themselves when recalling the film events than participants in the Factual condition [$t(47) = 3.11$, $SEM = 0.48$, $p = 0.003$], and they used more words suggestive of cognitive mechanisms and insight than Factual focus participants [$t(47) = 2.39$, $SEM = 0.55$, $p < 0.03$].

Differences in the number of thoughts and feelings in coded recall protocols more directly supported the hypothesis that affective retellings led to more subjectivity in later recall. Following Hashtroudi et al. (1994), we counted the number of opinions, evaluations, and other instances of personal commentary intruded into recalls of the film (using the same procedure for resolving discrepancies). Subjective intrusions were more common when participants had had an Affective focus ($M = 4.4$) than when they had had a Factual focus ($M = 2.7$), $t(48) = 1.72$, $SEM = 0.98$, $p < 0.05$ (one-tailed).

A second measure of personal involvement is the use of quotations. That is, the act of quoting, it has been asserted, is an indication of involvement in the story.¹ Compare participant 8's description using a quote: 'the Italian guy said, 'No, come on, he's dead, he's a piece of meat, it doesn't matter' to participant 11's 'one of his followers stops him and points out the father is already dead, and that he is being excessively angry'. Quotations indicate involvement in storytelling and suggest a desire to tell a good story. Accordingly, a larger percentage of participants attempted to quote dialogue following an Affective focus (64%) than following a Factual focus (36%) [$\chi^2(1) = 3.92$, $p < 0.05$] or no retelling (36%) [$\chi^2(1) = 3.92$, $p < 0.05$]. Note that these data represent attempted quotations and do not speak to the accuracy of those quotes.

In short, free recall showed effects of both retelling and of retelling perspective. Simply retelling led to greater recall of movie events than did not talking at all. Errors occurred in all conditions, but the greatest proportion of major errors occurred after Affective retellings. Affective retellings were longer on average, because participants intruded more of their own personal opinions into free recall.

Memory for emotion

Participants' ratings of how well individual emotion words described their mood were combined to form composite negative and positive emotion variables. Participants' ratings of how 'scared', 'afraid', 'upset', 'distressed', 'jittery', 'nervous', 'ashamed', 'guilty', 'irritated' and 'hostile' they felt were averaged into a Negative score for Time 1 (immediately after film-watching) and a Negative score at Time 2 (remembered emotion after a delay). Participants' ratings of 'enthusiasm', 'interest', 'determination', 'excitement', 'interest', 'alertness', 'activeness', 'strength', 'pride', and 'attention' were averaged into Time 1 and Time 2 Positive scores. Of interest was participants' ability to remember how they felt at Time 1.

¹We thank Herb Clark for suggesting this analysis.

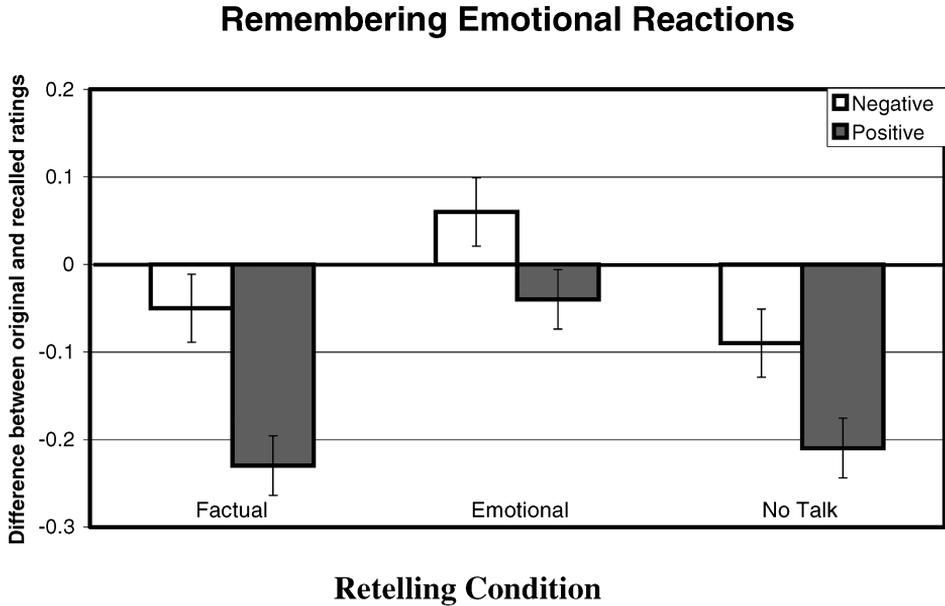


Figure 1. Mean change scores in remembered emotion as a function of retelling condition. A zero score represents accurate remembering; scores below zero represent underestimates of emotion actually experienced

A 3 (retelling condition) \times 2 (valence) ANOVA was computed on difference scores (ratings at Time 2 minus Time 1). Thus, a score of zero meant that participants remembered their prior emotions perfectly, whereas a positive score indicated an overestimate of prior emotions and a negative score indicated an underestimate of prior emotions.

As shown in Figure 1, participants remembered their negative emotions better than their positive emotions, $F(1, 72) = 7.16$, $MSE = 0.09$, $p < 0.01$. Participants were very accurate at remembering the negative feelings they had experienced immediately after film-watching. However, they were more likely to underestimate the amount of positive emotion experienced immediately after film-viewing.

Participants in the Affective condition performed better on the emotion memory task than did participants in the other two conditions, $F(2, 72) = 3.59$, $MSE = 0.11$, $p < 0.04$. Although the interaction between retelling condition and emotional valence did not reach significance ($F < 1$), an examination of Figure 1 suggests that all participants were good at remembering negative emotions and that the main difference between conditions was in remembering positive emotions.

Cued recall

Number of open-ended questions answered correctly did not vary as a function of retelling condition ($M = 11.23$, $F < 1$). Performance was neither at floor nor at ceiling, suggesting that the test was of an appropriate difficulty level to detect differences between conditions if they had existed.

Picture recognition

On average, participants correctly answered 9.2 out of 14 picture recognition questions, performing well above chance levels. Participants in the three retelling conditions did not

differ in the number of questions they answered ($F < 1$). Across participants, confidence was rated lower for incorrect answers ($M = 2.38$) than correct ones ($M = 4.06$), but confidence did not vary across experimental conditions.

DISCUSSION

Participants watched a disturbingly violent film clip. Afterwards, they were divided into three groups: the Affective group told how they felt while watching the film events; the Factual group told what happened from beginning to end; the control No Talk group did unrelated activities. There were dramatic differences in the retellings as a consequence of focus. Factual retellings were longer than emotional ones, and contained more perceptual detail, including spatial, temporal, and activity components of the events. By contrast, the emotional accounts contained far more expressions of affect and feelings as well as references to self. Clearly, people can easily talk about violent events in different ways.

Of central interest are the consequences of such biased retellings for later memory of the events. For cue-driven memory tests, such as picture recognition and cued recall, retelling focus had little effect. Affective retellings led to better memory in one case, namely when participants were asked to remember how they felt after watching the clip. Focus of retelling had larger effects on free recall, where memory retrieval factors are more influential. On the whole, participants followed the factual focus dictated at final recall; LIWC analyses revealed far fewer differences across recall conditions than in retelling, emphasizing that all groups were trying to take a factual focus for final recall as instructed. Even so, when retellers had focused their retellings on emotions as opposed to the sequence of events, they made relatively more major errors in free recall. This result is particularly important when considered in the context of eyewitnesses. That is, when providing evidence, it is more important that eyewitnesses avoid major errors (e.g. misremembering who did what) than minor errors (e.g. the colour of clothing). Affective focus in retellings also yielded more subjective commentary in final recall, even though all participants were strictly instructed to focus on the facts of the event. This result also has particular importance in the eyewitness domain, where it would be distracting to find subjective commentary in the reporting of a real criminal event.

Overall there were both positive and negative consequences of rehearsal: participants in the retelling conditions did recall more film events than control No-Talk participants, but it was not No-Talk participants who made the most errors. Rather, participants in the Affective condition made a greater proportion of major errors, and were more likely to intrude their own opinions into free recall. A general rule such as 'rehearsal helps' or 'rehearsal hinders' is insufficient to capture this pattern of data. It is the quality of rehearsal that determines memory accuracy. Nor was one type of retelling always best—the match between retelling focus and memory test focus was key. Factual retellings led to better performance on a free recall task that emphasized facts, but emotional retellings led to better performance on the emotion memory test.

This pattern is consistent with the Emotional Retelling as Selective hypothesis. The data are inconsistent with the Emotional Retelling as Deep Encoding hypothesis, which posited that emotional retellings would yield an overall mnemonic advantage, primarily due to the recoding of events in relation to oneself. The data were more complicated, with emotional retelling leading to a disadvantage on recall, an advantage on the emotion memory test, and no difference on the cued recall and recognition tests.

The data conform to the Emotional Retelling as Selective hypothesis, which posited that emotional retellings could create an organizing schema. The theme or schema that organizes retellings can later serve to bias memory retrieval (Tversky & Marsh, 2000). Thus retellings will have consequences for later memory to the extent that later memory tasks require participants to generate their own retrieval structures. This analysis accounts for both the memory effects observed in final free recall, and also the null effects in recognition memory (see Dudukovic et al., 2004, and Hashtroudi et al., 1994, for similar results). Recognition and cued recall tests provide retrieval cues, and constrain participants' responses, minimizing differences between conditions (Bower, Clark, Lesgold, & Winzenz, 1969; Eich, 1980).

Although the effects of retelling focus were eliminated on the recognition tests, recall is more typical of memory in practice—and thus it is striking that it is recall where retelling focus was influential. In the real world, interrogators rarely have enough knowledge to create a test that takes witnesses through the to-be-remembered events in chronological order and with exact cues (if they had that information, they wouldn't need to conduct an investigation). Rather, eyewitnesses will typically generate their own retrieval cues. And if eyewitnesses to a traumatic event retell with an emotional focus, this may compromise later testimony. An additional consequence of emotional retellings is that they may reduce the impact of later recall on listeners. Those with an emotional focus included more subjective commentary in their recall, a fact that could compromise their credulity (Johnson & Suengas, 1989). A story that contains more subjective information may be less credible to a jury, regardless of its overall accuracy level.

Eyewitnesses do not need to be exposed to misleading post-event information in order for their memories to change over time. The very act of talking about witnessed events leads to changes in memory, and the focus of retelling directs those changes. The effects reported here are, if anything, underestimates, given that participants only told their stories once. In laboratory paradigms, retelling effects increase with the number of retellings (Bergman & Roediger, 1999). Thus, the biasing effects of retellings are likely to be larger in real life, where witnesses will retell their stories many times, without the benefit of retrieval cues.

ACKNOWLEDGEMENTS

We are grateful to many people for their parts in this project. In particular, we thank Elena Krasnoperova for her help with organizing the study, pre-testing materials, and collecting data. Fiona Gabbert and Robert Belli provided many useful comments on an earlier version of this manuscript. We thank James Pennebaker for using the LIWC program to analyse the narratives. Laurent Perrodin provided technical assistance. Julie Bauer Morrison and Dawn McMurray also helped to collect the data. Coding was done by Drew Kervick, Kristina Olson, Brandon Sessoms, Emily Lowe, and Nancy Yang. Preparation of this manuscript was facilitated by funds from the Interval Research Corporation.

REFERENCES

- Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: predictions from an activation-based memory model. *Psychonomic Bulletin & Review*, 5, 1–21.

- Bergman, E. T., & Roediger, H. L., III. (1999). Can Bartlett's repeated reproduction experiments be replicated? *Memory Cognition*, *27*, 937–947.
- Berntsen, D. (2001). Involuntary memories of emotional events: do memories of traumas and extremely happy events differ? *Applied Cognitive Psychology*, *15*, S135–S158. DOI: 10.1002/acp.838
- Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior*, *8*, 323–343.
- Brewin, C. R., Christodoulides, J., & Hutchinson, G. (1996). Intrusive thoughts and intrusive memories in a nonclinical sample. *Cognition and Emotion*, *10*, 107–112.
- Dodd, D. H., & Bradshaw, J. M. (1980). Leading questions and memory: pragmatic constraints. *Journal of Verbal Learning & Verbal Behavior*, *19*, 695–704.
- Dudukovic, N. M., Marsh, E. J., & Tversky, B. (2004). Telling a story or telling it straight: the effects of entertaining versus accurate retellings on memory. *Applied Cognitive Psychology*, *18*, 125–143. DOI: 10.1002/acp.953
- Eich, E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory & Cognition*, *8*, 157–173.
- Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology*, *17*, 533–543. DOI: 10.1002/acp.885
- Glover, J. A. (1989). The 'testing' phenomenon: not gone but nearly forgotten. *Journal of Educational Psychology*, *81*, 392–399.
- Hashtroudi, S., Johnson, M. K., Vnek, N., & Ferguson, S. A. (1994). Aging and the effects of affective and factual focus on source monitoring and recall. *Psychology and Aging*, *9*, 160–170.
- Hyman, I. E. (1994). Conversational remembering: story recall with a peer versus for an experimenter. *Applied Cognitive Psychology*, *8*, 49–66.
- Johnson, M. K., & Suengas, A. G. (1989). Reality monitoring judgments of other people's memories. *Bulletin of the Psychonomic Society*, *27*, 107–110.
- Lane, S. M., Mather, M., Villa, D., & Morita, S. K. (2001). How events are reviewed matters: effects of varied focus on eyewitness suggestibility. *Memory & Cognition*, *29*, 940–947.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: an example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585–589.
- Marsh, E. J., & Tversky, B. (2004). Spinning the stories of our lives. *Applied Cognitive Psychology*, *18*, 491–503. DOI: 10.1002/acp.1001
- McCloskey, M., & Zaragoza, M. S. (1985). Misleading postevent information and memory for events: arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, *114*, 1–16.
- Meissner, C. A., & Brigham, J. C. (2001). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology*, *15*, 603–616. DOI: 10.1002/acp.728
- Pasupathi, M., Stallworth, L. M., & Murdoch, K. (1998). How what we tell becomes what we know: listener effects on speakers' long-term memory for events. *Discourse Processes*, *26*, 1–25.
- Payne, D. L. (1987). Hypermnnesia and reminiscence in recall: a historical and empirical review. *Psychological Bulletin*, *101*, 5–27.
- Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion*, *10*, 601–626.
- Roediger, H. L., III, Jacoby, J. D., & McDermott, K. B. (1996). Misinformation effects in recall: creating false memories through repeated retrieval. *Journal of Memory and Language*, *35*, 300–318.
- Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*, 677–688.
- Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: some things are better left unsaid. *Cognitive Psychology*, *22*, 36–71.
- Suengas, A. G., & Johnson, M. K. (1988). Qualitative effects of rehearsal on memories for perceived and imagined complex events. *Journal of Experimental Psychology: General*, *117*, 377–389.
- Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: a meta-analysis. *Psychological Bulletin*, *121*, 371–394.

- Tversky, B., & Marsh, E. J. (2000). Biased retellings of events yield biased memories. *Cognitive Psychology, 40*, 1–38.
- Tversky, B., & Tuchin, M. (1989). A reconciliation of the evidence on eyewitness testimony: comments on McCloskey and Zaragoza. *Journal of Experimental Psychology: General, 118*, 86–91.
- Wade, E., & Clark, H. H. (1993). Reproduction and demonstration in quotations. *Journal of Memory and Language, 32*, 805–819.
- Zaragoza, M. S., Payment, K. E., Ackil, J. K., Drivdahl, S. B., & Beck, M. (2001). Interviewing witnesses: forced confabulation and confirmatory feedback increase false memories. *Psychological Science, 12*, 473–477.