

## Telling a Story or Telling it Straight: The Effects of Entertaining Versus Accurate Retellings on Memory

NICOLE M. DUDUKOVIC<sup>1\*</sup>, ELIZABETH J. MARSH<sup>2</sup>  
and BARBARA TVERSKY<sup>1</sup>

<sup>1</sup>*Stanford University, USA*

<sup>2</sup>*Washington University in St Louis, USA*

### SUMMARY

People retell events for different reasons. Sometimes they try to be accurate, other times entertaining. What characterizes retellings from different perspectives? How does retelling perspective affect later recall of events? In the current research, participants retold a story either three times or not at all. By instruction, retellings were either entertaining or accurate. Compared to accurate retellings, entertaining retellings contained more affect, but fewer sensory references. On a subsequent memory test, participants who retold with an accuracy goal recalled the greatest number of story events, and their recall protocols were the most accurate and detailed, and least exaggerated. However, recognition memory did not differ across groups, suggesting that differences in retrieval structures (necessary for recall but not recognition) were key to understanding later differences in memory. Compared to telling it straight, the creative process of telling a story leads to qualitative and quantitative changes in later recall. Copyright © 2004 John Wiley & Sons, Ltd.

In everyday life, when people witness an interesting event or hear an amusing story, they often share it with others, telling the story in their own words. How one retells an event depends on the audience and the purpose of retelling. For example, when testifying in court or supplying evidence to a police officer, people usually try to be as accurate as possible. However, when relating an anecdote to friends, people often focus on entertaining their audience rather than on accuracy. In this case, they may make the story more interesting by omitting certain details and exaggerating and embellishing others. The act of retelling is a creative, constructive process, and the final product depends on the perspective the reteller adopts. For better or for worse, this perspective can affect what the reteller later remembers or misremembers (Tversky & Marsh, 2000).

Most laboratory memory studies encourage veridical rehearsal of information. Participants are typically asked to study lists of words or read stories, recall them, and then recall them again. Such rehearsal or retrieval practice generally aids memory (e.g. see Roediger & Gynn, 1996; Roediger, Wheeler, & Rajaram, 1993). However, this kind of detailed, accurate rehearsal may not be natural for participants. Even in a standard laboratory

\*Correspondence to: Nicole M. Dudukovic, Department of Psychology, Bldg. 420, Stanford University, Stanford, CA 94305-2130, USA. E-mail: nicoled@stanford.edu

Contract/grant sponsor: Stanford University.

Contract/grant sponsor: National Institutes of Mental Health.

paradigm, participants introduced more errors into rehearsal when strict accuracy was not emphasized (Gauld & Stephenson, 1967). In real life, the accuracy goal is rarely emphasized. The consequence is that people often feel free to bend the truth; diary studies reveal that people are quite willing to admit to telling lies in everyday conversation (e.g. DePaulo & Kasky, 1998; DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Turner, Edgley, & Olmstead, 1975). In one study, subjects were recorded while interacting with a new person for 10 min, and immediately afterwards indicated the truth-value of each of their statements. Sixty per cent of subjects admitted to lying during the conversation (Feldman, Forrest, & Happ, 2002). Lies do not necessarily constitute retellings of autobiographical events; however, similar results were found when autobiographical retellings were examined. Marsh and Tversky (in press) found that participants who kept diaries of naturally-occurring retellings of real events reported large amounts of distortion: 61% of the stories they told about their own lives were distorted in at least one way by their own admission. However, participants judged only 42% of their retellings as inaccurate, indicating that distortion may be so common in everyday conversation that it is considered acceptable and often not regarded as inaccurate. Thus, while laboratory studies that emphasize veridical retellings may have valuable implications for eyewitness testimony and other real-life situations that focus on accuracy, they do not capture the everyday, spontaneous nature of the retelling process.

The few studies that have manipulated retelling instructions in the laboratory have found that people's stories do shift depending on such factors as the goal and audience of retellings. For example, participants in one study memorized a dialogue and then were instructed to retell the dialogue to entertain versus to be accurate (Wade & Clark, 1993). Participants produced verbatim quotations only when they had been instructed to be accurate. Wade and Clark argued that in real conversations, quotations are used as a stylistic device to make stories more entertaining and are often inaccurate. Social context is another factor that influences the type of information retold (Hyman, 1994). Hyman manipulated whether participants retold a story to a peer or to an experimenter. Peer-directed retellings included more personal evaluations of a story, while experimenter-directed retellings tended to be more consistent with the original narrative style. Together, these studies suggest that people alter the stories they tell depending on their audience and their communication goals.

The way in which people retell events has consequences for later behaviour. For example, subjects who described a character to a hostile audience (that disliked the target character) later rated the character less positively (Sedikides, 1990). That is, a communication goal introduced at retelling led to changed person impressions. One possibility is that the biased retelling changed what subjects remembered about the target character. In support of this, Tversky and Marsh (2000) showed that the perspective used in retelling events affects both the amount of original information recalled and the type of errors made in final recall and recognition. For example, participants who crafted a letter of complaint about one of two story characters later recalled more negative story details about the discussed character. More interestingly, they also misattributed perspective-relevant activities to the discussed character, even though such errors were rarely made during retellings. Together, these studies suggest that the goals that guide retellings affect later memory.

One way that retellings affect memory is via creation of a schema that is later used to guide reconstruction of events. It is well known that schemas guide both what is encoded and what is retrieved. For example, in the classic Bransford and Johnson (1972) study,

people failed to understand (encode) an abstract passage without being told that it was about washing clothes. Similarly, Anderson and Pichert (1978) demonstrated that schemas also guide retrieval. Their participants read a story about two boys playing in a house from one of two perspectives: a home-buyer or a burglar. At test, the provision of the second perspective yielded recall of additional details. That is, a subject who encoded the paragraph from the perspective of a home-buyer was later able to use the burglar schema to retrieve additional details.

As with schemas present at encoding, schemas are used during retelling to re-organize events, to select some details and omit others. The schema provides a top-down structure for events, and thus allows for elaborations in the retelling to bridge across and interpret events. When later attempting to recall the original event, after retelling, the subject may be biased by prior perspective. The result is as if a schema were also present at retrieval: better memory for schema-consistent information but worse memory for schema-inconsistent information.

But what happens when the relevant schema for retelling is less clear? If the retelling does not map onto a pre-established schema or script, will it still have consequences for memory? Work on retellings has focused on goals and audiences for which there are clear scripts and schemas. A large number of stories are told, however, simply to entertain. What constitutes the 'entertainment script'? As described already, the one study examining entertainment retellings (Wade & Clark, 1993) focused on how the entertainment goal led to fewer verbatim quotations. Any listener knows, however, that a good story is much more than a lack of verbatim quotations. One aspect of the current research was to more fully understand what constitutes an entertaining story. How does an entertaining story differ from an accurate rehearsal, both quantitatively and qualitatively? The second goal was to examine the effect of telling an entertaining story on memory for the original event. That is, do entertaining stories lead to entertaining free recall protocols? Are there benefits from an entertaining retelling, even though these are less veridical than the accurate rehearsals? The third goal was to examine the connection between retellings and recall, with the aim of better understanding the mechanisms by which retellings affect memory.

The present study addressed these questions by looking at memory for a humorous short story in three different experimental groups: a group who retold a story under instructions to be accurate, a group who retold it under instructions to be entertaining, and a control group who did not retell the story. Of interest were the retellings themselves, and performance on a series of post-retelling memory measures including recall, character impressions, verbatim memory, and recognition memory for story events.

## METHOD

### Participants

Sixty-three Stanford undergraduates participated in the experiment. They were recruited through campus advertisements for a study on comprehension of verbal materials. Twenty-one participants were randomly assigned to each of the three conditions: entertainment, accuracy, and no-talk. Participants in the entertainment and accuracy conditions were either paid \$20.00 for their participation in all three experimental sessions or given 2 hr of experimental course credit. Participants in the no-talk condition were either paid \$15.00 for their participation in two experimental sessions or given 1.5 hr of experimental course credit.

## Materials

The story was told from the first-person perspective of a bartender who had only been working for 3 wk. The story described a particularly hectic night for the bartender. It contained both interesting events, such as a pregnant woman going into labour in the bar and an intoxicated woman spilling wine all over another woman's white suit, and boring events, such as a man asking the bartender for a napkin and the bartender preparing martinis. The story was written in an amusing manner in order to insure that it was a series of events that people would want to retell in real life.<sup>1</sup>

The beliefs questionnaire asked participants to rate the story characters on 10 attributes, such as rudeness, on a 1 to 5 scale. Higher values corresponded to more extreme beliefs. There were two orders of this questionnaire, counter-balanced across the participants.

The actions test consisted of 24 items, which tested memory for story events. One third of the items were correct, meaning that they were events that had occurred in the original story, such as one of the men in the bar asking for an ashtray. Another third of the items were incorrect but plausible given the context of the story, such as a character called Ms Make-up putting on more lipstick. The remaining items were incorrect and considered implausible given the context of the story, such as someone bringing a dog into the bar. Participants were asked whether or not each event had occurred in the original story, and then they rated their confidence for each response on a 5-point scale. There were also two orders of this test, counter-balanced across the participants.

The verbatim test contained 15 items, which tested memory for the exact wording of the story. Each item consisted of three possible answer choices, one of which was the original wording from the story. There was only one version of this test.

## Procedure

Participants in the accuracy and entertainment conditions attended three experimental sessions, which were spaced 2 days apart. Participants in the no-talk condition attended two experimental sessions, which were spaced 4 days apart. Participants were tested individually.

At the first session, all participants were informed that they might be videotaped. Each participant was then given a typed copy of the story. They were instructed to read the story carefully two times through and to inform the experimenter when they were finished. After reading the story, each participant was given a distractor task, consisting of a series of puzzles and games, and was told to complete as much as possible in 8 min. Following the distractor task, the participants in the no-talk condition were given permission to leave. The participants in the other two conditions read either the entertainment or the accuracy retelling instructions. The instructions for the entertainment condition were as follows:

I would like you to retell, in an amusing manner, the story that you read earlier. You will be videotaped while you tell it, and you will later be rated on how entertaining your retelling was. It is not at all important that you tell the story exactly as it was written, and you do not need to use the same words that appeared in the original story. You may want to imagine that your audience is a group of friends who have never heard the story before and whom you would like to amuse. Remember that you will be rated on how entertaining you are.

<sup>1</sup>A complete copy of the story may be obtained from any of the authors.

The instructions for the accuracy condition were as follows:

I would like you to retell, in a precise manner, the story that you read earlier. You will be videotaped while you tell it, and you will later be rated on how accurate your retelling was. It is very important that you tell the story exactly as it was written, and when possible, use the same words that appeared in the original story. You may want to imagine that your audience is a policeman or lawyer who needs an accurate account. Remember that you will be rated on how accurate you are.

After reading the instructions, the participants were told that the experimenter would leave the room while they retold the story to a VHS camera. They were instructed to let the experimenter know when they were finished. Participants were allowed as much time as they needed for retelling.

Only the participants in the entertainment and accuracy conditions participated in the second session, which was held 2 days after the first session. Retelling instructions were repeated, and participants retold the story to the video camera without the experimenter being present.

Four days after the first session (2 days after the second session), all participants returned to the laboratory for the final session. Participants in the entertainment and accuracy conditions were given the same retelling instructions and retold the story for the third and last time to the video camera. All participants then worked on the distractor task for 8 min. Following the distractor task, all participants read the following free recall instructions:

I would now like you to recall the original story that you read as best you can and type it on the computer as you remember it. Report exactly what happened in the story you read, and when possible, use the original wording from the story. Try to get the facts and events correct, and avoid guessing. When you are finished, please let the experimenter know.

The participants then typed the story in a Word document on a Macintosh computer. Following free recall, the participants were informed that the experimenter was interested in their opinions about the characters, and they completed the beliefs questionnaire. They then completed the actions test followed by the verbatim test.

The first session took approximately 30 min, the second session took 15 min, and the third session took 1 hr. At the end of the final session, all participants were thanked and debriefed.

## RESULTS

### Retellings

#### *Judgments of accuracy and entertainment*

A manipulation check was performed on the retellings to ensure that participants were following instructions and to verify that participants in the entertainment condition were in fact more entertaining than participants in the accuracy condition. A separate scorer who was blind to experimental condition watched the videotapes of the retellings from the third session. The scorer rated the global entertainment value of each of these retellings on a scale from 1 to 5, with 1 being the least entertaining and 5 being the most entertaining. On

average, the retellings of participants in the entertainment condition were rated as significantly more entertaining ( $M = 3.48$ ,  $SD = 1.12$ ) than were the retellings of accuracy participants ( $M = 2.05$ ,  $SD = 0.86$ ) [ $t(40) = 4.62$ ,  $p = 0.000$ ]. The scorer also rated the global accuracy of the final session retellings on a scale from 1 to 5, with 1 being the least accurate and 5 being the most accurate. The retellings of accuracy participants were rated as significantly more accurate ( $M = 4.24$ ,  $SD = 0.77$ ) than the retellings from the entertainment participants ( $M = 3.24$ ,  $SD = 1.04$ ) [ $t(40) = 3.53$ ,  $p = 0.001$ ]. Thus, the instruction manipulation was validated, with entertainment participants retelling the story in a more entertaining manner and accuracy participants retelling the story more accurately.

### Use of language

Transcripts of the retellings from each participant's third experimental session were analysed using a text analysis program developed by Pennebaker and Francis (see Pennebaker & Francis, 1996) in order to quantify qualitative differences between the two types of retellings. The Linguistic Inquiry and Word Count (LIWC) program counts words in several linguistic, emotional, and cognitive categories. Each of the LIWC categories is composed of dictionary words that define the scale. For example, the tentative scale includes words such as 'maybe,' 'perhaps,' and 'guess,' which indicate hesitancy or tentativeness. Of particular interest in the present study were the use of emotion words, sensory words, verb tenses, the presence of disfluencies, and indicators of tentativeness versus certainty.

Table 1 shows the results from the accuracy and entertainment retellings along with the  $p$ -values for the differences. Word count is reported in words, but all other variables are reported as percentages of total length of the retelling protocol. Items in Table 1 are not strictly significant unless  $p$  has been adjusted for multiple comparisons to less than 0.00625. However, to capture the bigger picture we will describe here all results even if significant only at the unadjusted significance level (0.05).

Accuracy and entertainment retellings did not differ significantly in terms of number of words, and there was large within-condition variability on this dimension. Accuracy

Table 1. LIWC results for accuracy and entertainment retellings

	Accuracy	Entertainment	$p$ -Value
Linguistic factors			
Word count	1069.5 (296.4)	1163.1 (616.0)	0.5
Verb tense			
Past	7.5 (3.1)	6.8 (3.0)	0.4
Present	10.0 (3.3)	13.0 (4.1)	0.013
Affective or emotional processes			
Affect	2.5 (0.9)	3.2 (1.2)	0.040
Sensory and perceptual processes			
Senses	4.3 (1.2)	2.6 (0.8)	0.000*
Cognitive processes			
Certainty	0.7 (0.4)	1.2 (0.7)	0.025
Tentativeness	1.7 (0.6)	2.5 (0.7)	0.000*
Additional variables			
Disfluencies	3.0 (2.4)	1.3 (0.9)	0.005*

Standard deviations are in parentheses.  $p$ -values marked with \* are significant after adjusting for multiple comparisons ( $p < 0.00625$ ).

retellings ranged from 547 to 1538 words, while entertainment retellings ranged from 387 to 3166 words. Thus, length of the retelling seems to vary by individual and be independent of experimental condition.

Entertainment retellings tended to include a greater percentage of present tense verbs than accuracy retellings [ $t(40) = 2.59$ ]. The difference involving use of past tense verbs was not significant,  $t < 1.01$ . Use of future tense verbs was not compared, as future tense verbs were rarely used in the original story, and this was reflected in the retellings.

Entertainment retellings also contained a greater percentage of emotion words than did accuracy retellings [ $t(40) = 2.12$ ], and more words indicating certainty [ $t(40) = 2.34$ ]. Interestingly, entertainment retellings also contained more words indicating tentativeness [ $t(40) = 3.90$ ]. While the entertainment retellings included a greater percentage of references to affect, the accuracy retellings included more sensory words, and this difference was highly significant [ $t(40) = 5.04$ ]. Accuracy retellings also contained relatively more disfluencies, such as ‘uh’ and ‘um’ [ $t(40) = 2.98$ ].

As described already, the accuracy and entertainment goals led to differential language use in the retellings (see Table 1). Of interest was whether language use also predicted the ratings of accuracy and entertainment made by the condition-blind judge (the manipulation checks described in the first paragraph of this section). Is there a ‘language of storytelling’? If so, we would expect entertainment ratings to be related to language usage. To examine this, the LIWC results were correlated with the accuracy and the entertainment ratings for each of the two retelling conditions. The results are shown in Table 2.

For the Accuracy condition, only two variables predicted accuracy ratings: retellings with less past tense and more present tense were rated as more accurate ( $r$ 's =  $-0.52$ ,  $+0.46$ ). This finding is explainable by the fact that the original story actually contained a

Table 2. For each of the two retelling conditions, Table 2 reports the correlations between the LIWC language counts and the global ratings of the retellings’ accuracy and entertainment

	Accuracy rating	Entertainment rating
Accuracy retelling		
Accuracy rating		-0.17
Affect	0.07	-0.04
Certainty	-0.08	-0.29
Disfluencies	-0.01	-0.36
Entertainment rating	-0.17	
Past tense	-0.52*	-0.01
Present tense	0.46*	0.15
Senses	0.34	0.18
Tentativeness	0.29	0.02
Entertainment retelling		
Accuracy rating		-0.40
Affect	-0.23	0.62**
Certainty	-0.18	0.37
Disfluencies	0.29	-0.47*
Entertainment rating	-0.40	
Past tense	0.12	-0.46*
Present tense	0.04	0.32
Senses	0.33	-0.02
Tentativeness	-0.23	0.10

\*Correlation is significant at the 0.05 level (two-tailed).

\*\*Correlation is significant at the 0.01 level (two-tailed).



greater proportion of present than past tense. Not a single variable predicted the entertaining ratings of accuracy retellings, perhaps due to restriction of range (that is, retellings from the accuracy condition simply were not rated as very entertaining).

The more interesting data are for retellings with an entertainment goal. Although language did not predict accuracy ratings, retellings in the entertainment condition that were rated as more entertaining tended to be rated as less accurate ( $r = -0.40$ ,  $p < 0.10$ ). Language was important, however, for predicting entertainment ratings. These retellings were rated as more entertaining when they contained more affect words ( $r = +0.62$ ), less use of the past tense ( $r = -0.46$ ), and fewer disfluencies ( $r = +0.47$ ). In addition, there was an interesting trend: retellings were considered most entertaining when they were told with certainty ( $r = +0.37$ ). These data suggest that a good story is not necessarily accurate, but is told with certainty, does not contain disfluencies, and is not told in the past tense. Thus, the results suggest that a 'language of storytelling' was visible in the most entertaining retellings, and that accurate stories were less associated with this particular type of language.

Overall, the ratings of accuracy and entertainment value as well as the word percentages reported in Table 1 indicate that the two types of retellings contained qualitatively different information. Namely, an entertainment goal resulted in retellings that were indeed more entertaining, and contained more references to emotions and involved more evaluations (as suggested by use of words from the certainty and tentativeness scales). These retellings tended to use more present tense, suggesting the story was being told 'in the moment.' In contrast, an accuracy goal led to retellings that were rated as more accurate and that contained more references to sensory details. These retellings included more disfluencies, perhaps as a retelling aimed at accuracy is less natural than is telling a story.

### *Content*

The two retellings were told differently—but did they contain different amounts or different pieces of information? To answer questions about content, the original story was broken down into 105 event units for coding purposes. Two coders blind to experimental condition scored each of the transcribed retelling protocols from the third experimental session. Each coder counted the number of event units included and then rated each event on the amount of accuracy/detail included and the amount of exaggeration. Both sets of ratings were done on a scale from 1 to 5. For the accuracy/detail scale, a rating of 5 signified an event that was reported in a completely accurate manner with no details missing and no details added, while a rating of 1 signified a reported event that was barely recognizable because it was so inaccurate. For the exaggeration scale, a rating of 1 was given to an event that was reported exactly like the one in the original story with no exaggeration, while a rating of 5 was given to an event that was grossly exaggerated with information added to improve its entertainment value. The coders also noted the number of intrusions contained in each retelling protocol, where an intrusion was defined as an event that did not appear in any form in the original story.

Both coders were given specific coding instructions that included examples of statements and how they should be rated. The correlations between the two experimenters were 0.98 for the number of events included, 0.71 for the accuracy/detail measure, 0.76 for the exaggeration measure, and 0.60 for the number of intrusions. All data reported are those of the primary coder.



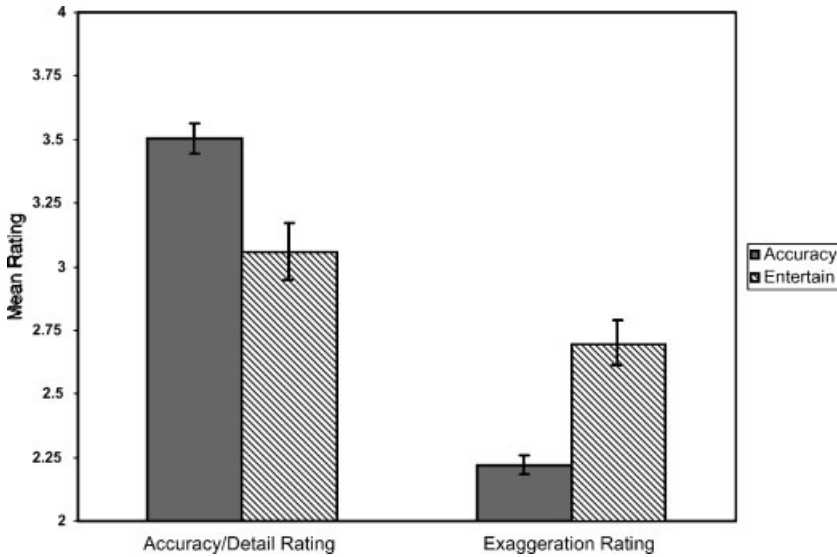


Figure 1. Mean accuracy/detail ratings and exaggeration ratings for retellings as a function of retelling condition. Higher values indicate greater accuracy and exaggeration

Accuracy retellings did in fact contain a greater proportion of story events ( $M = 0.55$ ,  $SD = 0.10$ ) than did entertainment retellings ( $M = 0.44$ ,  $SD = 0.14$ ) [ $t(40) = 2.89$ ,  $p < 0.01$ ]. As shown in Figure 1, across the two retelling conditions, retold events differed in their average ratings of accuracy and detail [ $t(40) = 3.49$ ,  $p = 0.001$ ] and exaggeration [ $t(40) = 4.65$ ,  $p = 0.000$ ]. As expected, on average, items retold by the accuracy group were rated as more accurate and detailed ( $M = 3.50$ ,  $SD = 0.27$ ) than those retold by the entertainment group ( $M = 3.06$ ,  $SD = 0.52$ ). Similarly on average, items recalled by the entertainment group were rated as more exaggerated ( $M = 2.70$ ,  $SD = 0.43$ ) than those recalled by the accuracy group ( $M = 2.22$ ,  $SD = 0.18$ ). This item-by-item evaluation of accuracy and entertainment yields a pattern of data that nicely parallels the global impressions of coders (the manipulation check).

While retellings from accuracy participants contained more correct story information, the reverse was true for incorrect information. Entertainment participants made more intrusions when retelling the story than did accuracy participants [ $t(40) = 2.53$ ,  $p < 0.02$ ]. On average, entertainment participants made a large number of intrusions ( $M = 12.6$ ,  $SD = 16.2$ ) whereas accuracy participants made few ( $M = 3.4$ ,  $SD = 3.3$ ).

Did accuracy and entertainment participants differ in *which* items they included in their retellings? The story was not constructed to contain accurate versus entertaining events, and, as such, few events were exclusively associated with one of the two retelling conditions. Figure 2 plots the 105 story events on the x-axis. For each of these story events, was it included more often in the accuracy or entertainment retellings? The y-axis shows the following computed value: the number of accuracy participants including a given story event in their retelling, minus the number of entertainment participants including the same story event in their retelling. Thus, positive values indicate that more accuracy participants included the item in their retellings, whereas negative values indicate that more entertainment participants included the item in their retellings. The absence of a bar indicates that equal numbers of accuracy and entertainment participants

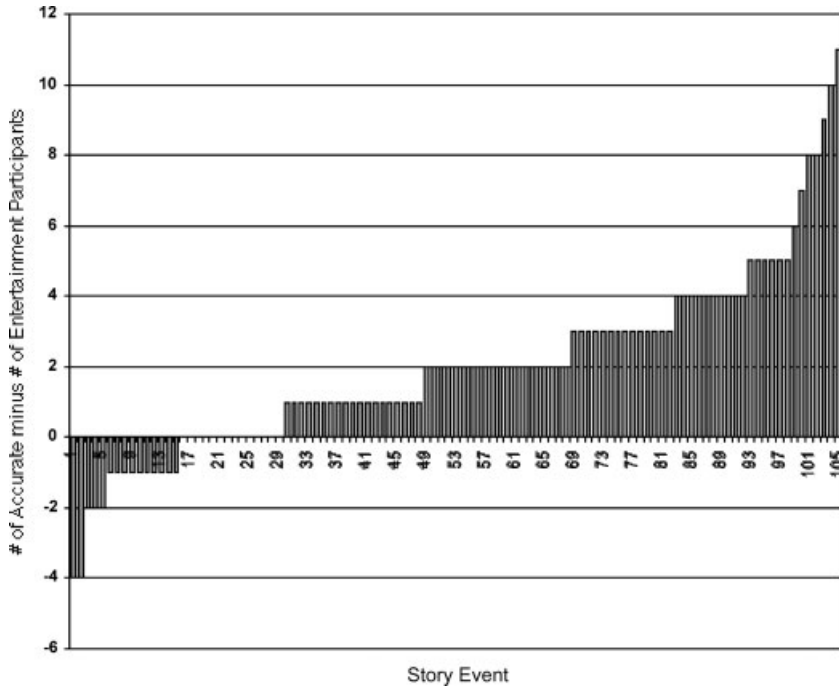


Figure 2. Individual story event inclusion in retellings as a function of retelling condition. The y-axis depicts in how many more accuracy than entertainment retellings a given story event occurred

included the event in their retellings. The relative heights of the bars indicate the relative discrepancy between the two groups in retelling frequency; tall bars mean there was more of a discrepancy between the two groups in the number of participants including the event in their retellings.

The first major point is that rarely did more entertainment participants (than accuracy) use a story event. That is, of the 105 story events, only 15 were rehearsed by more entertainment than accuracy participants (as depicted visually by the 15 left-most bars in the graph). Fourteen events were rehearsed by equal numbers of entertainment and accuracy participants (as indicated by no bars on the graph). The remaining 76 events were rehearsed by more accuracy than entertainment participants. It is not that entertainment participants described a different (particularly entertaining) set of events in their retellings. Rather, it is that accuracy participants tended to include story events omitted by entertainment participants. It is not, however, that these events were never included in entertainment retellings. On average, for these 76 events, there was only a difference of three participants between accuracy and entertainment retellings. So, it is not that events tended to be associated with the accuracy and not the entertainment perspective. Rather, across a large number of story events, more accuracy participants included them in their retellings, leading to a greater proportion of story events included in accuracy retellings.

The same conclusion was reached via a different analysis. To test the hypothesis that rehearsal may have varied by event type, two separate raters classified the story events into three categories: action, conversation, and description. Twenty events depicted action, 56 involved conversation, and 11 were purely description. The remaining 18 events were left out of the analysis because the raters did not agree on a classification.

A 2 (retelling)  $\times$  3 (event type) analysis of variance (ANOVA) revealed a main effect of event type [ $F(2, 80) = 70.88, p = 0.000$ ]. Overwhelmingly, a greater proportion of descriptions were included in the retellings ( $M = 0.73, SD = 0.26$ ) than actions ( $M = 0.55, SD = 0.35$ ) or conversations ( $M = 0.47, SD = 0.29$ ). There was also a marginally significant effect of retelling condition [ $F(1, 40) = 3.60, p = 0.065$ ], which is not surprising since accuracy participants retold the greatest number of events overall. More importantly, the interaction between retelling condition and event type did not even approach significance ( $F < 1$ ). Therefore, participants demonstrated a similar pattern of rehearsal of the different event types regardless of retelling condition.

### *Summary of retelling analyses*

Compared to standard rehearsals told for an accuracy goal, entertainment retellings were both qualitatively and quantitatively different. Entertainment retellings received higher global ratings of entertainment value than did accuracy retellings. Entertainment ratings were also rated as less accurate, both globally and on an item-by-item basis. In addition, they contained less story information, although this deficit was not linked to any particular items. Entertainment retellings were also linked to a pattern of language use, the 'language of storytelling'. As such, they were more fluent, told in the moment, and contained emotion words. Some of this language use was highly related to the retellings' being perceived as entertaining by judges.

### **Free recall**

Only one participant failed to finish the free recall in the time allotted, and his recall data were not included in the analyses. Thus, these analyses are based on 62 participants.

### *Judgments of accuracy and entertainment*

A scorer blind to experimental condition assigned global ratings of accuracy and entertainment to each free recall protocol. The scorer read each protocol and then rated the overall accuracy on a scale from 1 to 5, with 1 being the least accurate and 5 being the most accurate. The scorer also rated the entertainment value of each recall protocol on a scale from 1 to 5, with 1 being the least entertaining and 5 being the most entertaining. Thus, these data parallel the global ratings collected for the third retelling.

An ANOVA was used to compare the global accuracy ratings across the three conditions. The main effect of condition was significant; participants in the three conditions were not rated as equally accurate [ $F(2, 59) = 7.02, p < 0.005$ ]. Recall protocols from the accuracy condition were rated as more accurate ( $M = 3.4, SD = 0.68$ ) than the recall protocols from the entertainment condition ( $M = 2.8, SD = 0.81$ ) and the no-talk condition ( $M = 2.6, SD = 0.83$ ). To test for predicted differences between conditions, two *t*-tests were conducted at the  $p = 0.05$  level, with the error rate adjusted to  $p = 0.025$  for two contrasts. Recall protocols from the accuracy condition were rated as significantly more accurate than protocols from the entertainment condition [ $t(40) = 2.68, p < 0.025$ ]. However, recall protocols from the entertainment and no-talk conditions did not differ in their global accuracy ratings ( $t < 1.1$ ).

Global entertainment ratings for the free recall protocols were also compared using an ANOVA. The main effect of condition was not significant ( $F < 1$ ) and none of the groups differed significantly on this measure. Thus, unlike the retellings, for which the entertainment group was rated as significantly more entertaining overall, all free recall protocols

were rated as equally entertaining regardless of experimental condition ( $M = 3.2$ ,  $SD = 0.83$ ). Since the retellings were rated by watching the videotapes while the recall protocols were rated by reading the actual protocols, it is possible that the difference in entertainment ratings found in the retellings can be attributed to the animation of the entertainment participants as they retold the story. Gestures, facial expressions, and voice inflection can obviously not be captured on paper.

#### *Use of language*

All free recall protocols were analysed using the LIWC program to investigate whether the language used in the retellings would impact the language used in free recall. On the whole, the answer seemed to be no. The only significant difference was in length. An ANOVA was used to compare the number of words used in the free recall across the three conditions (accuracy, entertainment, and no-talk). The main effect of condition was significant; on average, participants in the three conditions did not use the same number of words in their free recall [ $F(2, 59) = 5.26$ ,  $p < 0.01$ ]. The most words were used in the accuracy condition ( $M = 833$ ,  $SD = 154$ ), followed by the entertainment condition ( $M = 725$ ,  $SD = 218$ ), and the fewest words used by participants in the no-talk condition ( $M = 636$ ,  $SD = 208$ ). The remaining LIWC analyses did not show any significant differences for the variables of interest when the error rate was adjusted for multiple comparisons; thus, the data are not displayed here. The absence of a 'language of storytelling' in the entertainment recall protocols may explain the lack of difference in entertainment value found between the two conditions for free recall.

#### *Content*

The coding procedure for the recall data was identical to that used for the retellings. The correlations between the two experimenters were 0.97 for the number of events included, 0.89 for the accuracy/detail measure, 0.92 for the exaggeration measure, and 0.70 for the number of intrusions. All data reported are those of the primary coder.

An ANOVA was used to compare the proportion of events recalled across the three conditions. The main effect of condition was significant; participants in the three conditions did not recall equal numbers of story events [ $F(2, 59) = 5.09$ ,  $p < 0.01$ ]. As expected, the proportion of events recalled was largest in the accuracy condition ( $M = 0.58$ ,  $SD = 0.10$ ) and smaller in the entertainment ( $M = 0.50$ ,  $SD = 0.12$ ) and no-talk conditions ( $M = 0.47$ ,  $SD = 0.12$ ). This finding supports the prediction that participants in the accuracy condition would recall the greatest number of events. Looking at predicted differences between conditions, participants in the accuracy condition recalled significantly more than did participants in the entertainment condition [ $t(40) = 2.35$ ,  $p < 0.025$ ]. However, participants in the entertainment and no-talk conditions did not differ in amount recalled,  $t < 1$ .

Did participants differ in how detailed their recall of each item was? That is, given that items were recalled, were there average differences across conditions in ratings of accuracy and exaggeration? As with the retellings, each recalled item (up to 105) was rated for its accuracy and exaggeration on separate 5-point scales. An ANOVA was used to compare the mean accuracy/detail ratings for the events included in the free recall across the three conditions. There was a main effect of condition [ $F(2, 59) = 8.93$ ,  $p = 0.000$ ]. As shown in the left-hand side of Figure 3, the mean accuracy/detail rating was highest (indicating the greatest amount of accuracy) in the accuracy condition ( $M = 3.74$ ,  $SD = 0.19$ ) and lower in the entertainment ( $M = 3.60$ ,  $SD = 0.16$ ) and no-talk conditions

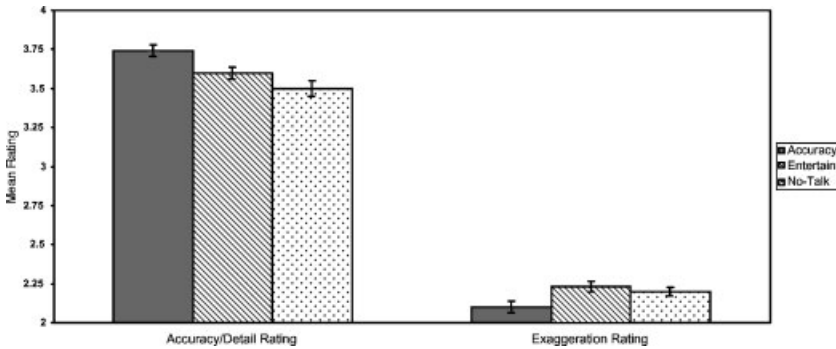


Figure 3. Mean accuracy/detail ratings and exaggeration ratings for free recall as a function of original retelling condition. Higher values indicate greater accuracy and exaggeration

( $M = 3.50$ ,  $SD = 0.21$ ). As predicted, with the error rate adjusted for two contrasts, the events recalled by accuracy participants were rated as significantly more accurate/detailed than events recalled by entertainment participants [ $t(40) = 2.69$ ,  $p = 0.01$ ]. Again, there was no significant difference between the entertainment and no-talk conditions,  $t < 1.70$ . These averages of item-by-item ratings again parallel the global impressions formed by the raters.

An ANOVA was used to compare the mean ratings of exaggeration for the events included in the free recall across the three conditions. As shown in the right-hand side of Figure 3, the mean exaggeration rating was highest (indicating the most exaggeration) in the entertainment ( $M = 2.23$ ,  $SD = 0.16$ ) and no-talk conditions ( $M = 2.20$ ,  $SD = 0.12$ ), and lowest in the accuracy condition ( $M = 2.10$ ,  $SD = 0.17$ ). This difference was significant, [ $F(2, 59) = 3.79$ ,  $p < 0.05$ ]. Therefore, even though the global ratings of entertainment were the same across conditions, the free recall protocols did differ in their average amount of item exaggeration. As predicted, with the error rate adjusted for two contrasts, events recalled by entertainment participants were rated as significantly more exaggerated than events recalled by accuracy participants [ $t(40) = 2.44$ ,  $p = 0.02$ ]. Interestingly, the mean exaggeration ratings for the entertainment and no-talk conditions were not significantly different,  $t < 1$ .

The groups differed in amount of correct story information recalled; of interest was whether lower recall of the story would be associated with an increased number of errors about the story. An ANOVA was used to compare the mean number of intrusions included in the free recall across the three conditions. Although the entertainment group ( $M = 3.5$ ,  $SD = 3.1$ ) and the no-talk group ( $M = 3.6$ ,  $SD = 3.1$ ) made more intrusions than the accuracy group ( $M = 3.2$ ,  $SD = 3.4$ ), the variability was extremely high, and there were no significant differences across conditions ( $F < 1$ ).

Finally, we were interested in whether the difference in amount recalled across groups could be attributed to a particular type of information. We again examined recall as a function of story item type: action, conversation, or description. A 3 (retelling)  $\times$  3 (event type) ANOVA revealed a main effect of event type [ $F(2, 118) = 66.88$ ,  $p = 0.000$ ]. As in the retellings, a greater proportion of descriptions ( $M = 0.66$ ,  $SD = 0.27$ ) were included in the free recall than actions ( $M = 0.53$ ,  $SD = 0.34$ ) or conversations ( $M = 0.48$ ,  $SD = 0.29$ ). As reported for the full 105 events, there was also a main effect of condition [ $F(2, 59) = 5.10$ ,  $p = 0.009$ ]. However, the critical interaction between retelling condition

and event type was not significant [ $F(4, 118) = 1.56, p > 0.10$ ]. Entertainment retellings did not lead to a difficulty specific to one type of item in the free recall.

### Summary of recall analyses

Accuracy participants recalled the most story events, although this advantage was not linked to a particular type of story information. Rather, across the board, accuracy subjects' recall protocols were rated as more accurate and detailed (both globally and on an item-by-item basis). Recall by entertainment participants was surprisingly like that by participants who never rehearsed the story at all. Recall was not marked by differences in language use across conditions.

### Connections between retellings and recall

Characteristics of the retellings were indeed related to characteristics of the later recall protocols. The observant reader has already noted similarities in results between retelling and recall (e.g. the higher ratings of accuracy in both retelling and recall for accuracy participants). To examine this formally, we correlated three retelling characteristics (average accuracy/detail rating, average exaggeration rating, and proportion of story items included) with the same three characteristics of final recall. The correlations are shown in Table 3. As expected, more accurate retellings were linked to more accurate recall protocols (both in terms of ratings and amount recalled). Negative correlations occurred only when correlating exaggeration and accuracy variables. Furthermore, retellings that contained more story items were associated with final recall protocols that were rated on average as less exaggerated.

The number of story items included in retelling and recall were highly correlated, and this held for both retelling conditions. We did two additional analyses to investigate the relationship between an item's inclusion in retelling and its inclusion in final recall. For each participant, we computed the proportion of items included in retelling that were also included in final recall. Similarly, we computed the proportion of items *not* retold that were included in final recall. While the probability of recall was always high if an item had been retold, this conditional probability was significantly higher in the accuracy condition

Table 3. Correlations between characteristics of the retellings (accuracy/detail and exaggeration ratings, proportion of story items included) and characteristics of the final recall protocols

Characteristics of retellings	Characteristics of Final Recall		
	Accuracy/detail rating	Exaggeration rating	Prp. of story items
<b>Accuracy retelling</b>			
Accuracy/detail rating	0.57**	-0.69**	0.64**
Exaggeration rating	-0.35	0.70**	-0.44*
Prp. of story items included	0.60**	-0.53*	0.94**
<b>Entertainment retelling</b>			
Accuracy/detail rating	0.61**	-0.49*	0.35
Exaggeration rating	-0.32	0.51*	-0.04
Prp. of story items included	0.74**	-0.53*	0.89**

\*Correlation is significant at the 0.05 level (two-tailed).

\*\*Correlation is significant at the 0.01 level (two-tailed).



( $M = 0.90$ ,  $SD = 0.05$ ) than in the entertainment condition ( $M = 0.84$ ,  $SD = 0.06$ ) [ $t(40) = 3.44$ ,  $p = 0.001$ ]. The reverse pattern held for the proportion of items *not* retold that were included in final recall. While the chance of recalling an un-retold item was always low, the conditional probability was significantly higher in the entertainment condition ( $M = 0.22$ ,  $SD = 0.10$ ) than in the accuracy condition ( $M = 0.16$ ,  $SD = 0.06$ ) [ $t(40) = 2.19$ ,  $p < 0.05$ ].

### Beliefs questionnaire

Two participants failed to complete the questionnaire within the time allotted for the experimental session; thus the data reported here are from 61 participants. For each participant, a belief score was calculated by averaging over the 10 belief ratings. Scores ranged from 1 to 5, and higher scores are indicative of stronger, more extreme beliefs. The strength of the beliefs held by participants in all three conditions was approximately equal. An ANOVA on the mean scores across the three conditions did not yield significance ( $F < 1$ ). Thus, participants in the entertainment condition ( $M = 3.65$ ,  $SD = 0.47$ ) did not hold more exaggerated beliefs about the story characters than did participants in the accuracy ( $M = 3.82$ ,  $SD = 0.43$ ) or no-talk conditions ( $M = 3.68$ ,  $SD = 0.45$ ).

### Recognition memory

#### *Actions test*

Two participants failed to complete the actions test within the experimental session, and their data were not included in the analysis. For the remaining 61 participants, an ANOVA was used to compare the mean proportion of hits across the three conditions. The mean proportion of hits ( $M = 0.74$ ,  $SD = 0.16$ ) was very close across conditions, and there was no significant effect of condition ( $F < 1$ ). Mean proportion of false alarms to both plausible ( $M = 0.35$ ,  $SD = 0.18$ ) and implausible items ( $M = 0.06$ ,  $SD = 0.07$ ) also did not differ significantly across the three conditions ( $F$ 's  $< 1$ ). ANOVAs were also done on mean confidence ratings across the three conditions for all items, true items, plausible items, and implausible items. None of these results were significant and thus will not be discussed further.

#### *Verbatim test*

One participant failed to complete the verbatim test in the time allotted, and his data are not included here. For the remaining 62 participants, an ANOVA was used to compare the mean proportion of correct responses out of 15 items across the three conditions. Overall performance was quite good ( $M = 0.75$ ,  $SD = 0.14$ ) but there was no effect of condition ( $F = 1$ ).

## DISCUSSION

The current study examined the nature of entertaining and accurate retellings of events, and how the manner of retelling in turn affected memory. As expected, accurate and entertaining retellings differed dramatically. Telling a good story bore little similarity to accurate rehearsal. Stories told to entertain referred to fewer original events, and less detail was provided for events that were included. Compared to accuracy retellings, entertainment



retellings included relatively more emotion words and relatively fewer disfluencies. These differences resulted in judges rating retellings in the entertainment condition as more entertaining and less accurate than those in the accuracy condition. How did the manner of retelling affect memory? When subjects recalled the original story, differences emerged as a function of retelling condition. Recall of story events was greatest in the accuracy condition, both in number of events recalled and in judges' ratings of accuracy. Rehearsing information as accurately as possible did indeed have positive consequences for memory. However, there were surprisingly few differences between the entertainment and no-talk conditions. Telling three lively tales of the bartender's bad night led neither to an advantage in memory (as might be expected due to rehearsal) nor particularly exaggerated or entertaining final recall protocols. There were no differences across conditions on a number of cued memory measures: event recognition, a verbatim memory test, and ratings of character impressions.

We first discuss the retellings; of particular interest was whether or not there is an 'entertainment script'. It was not the case that there was a particular set of items associated with the entertainment perspective (as would be with a home-buyer or burglar or annoyed roommate or other perspective). Although entertainment participants included fewer story events in their retellings, it was not that they honed in on a few key events and omitted other boring events. It is possible that this finding is due to our choice of stimulus materials; the bartender story was designed to be entertaining, so there were fewer boring events to include or omit than one might typically find in real life. However, what appeared to make a story entertaining was the *way* in which it was told, not necessarily what was told. Entertainment participants made their stories more interesting by taking the same basic group of events and exaggerating them and adding their own details (intrusions). The same events can be boring or funny, depending on the skills of the storyteller (the reader may think of Jerry Seinfeld, known for making people laugh hysterically over such trivial events as ordering take-out food). Even within our sample, some people told better stories than others. As rated by judges, entertaining stories were told fluently, had lots of emotion words, were told in the present tense, and used words indicating certainty. These are all linguistic devices to hold a reader's attention. Descriptions of similar events were rated as more entertaining when told using the 'language of storytelling'.

Importantly, these different forms of retelling had different consequences for memory. As in the retellings, participants in the accuracy condition recalled significantly more events in their free recall than participants in the entertainment condition, and they recalled those events more accurately and in greater detail. An accurate retelling functioned like rehearsal in many laboratory situations. Accurate practice of information aids its later retrieval.

A different picture emerged when examining the free recall protocols in the entertainment condition. First, unlike the retellings, these free recall protocols were not rated as overall more entertaining (the global impression) than those from the other conditions. As noted already, much entertainment value is derived not only from the type of language used but also from non-verbal gestures and facial expressions, something present in the videotaped retellings but not in the typed recall protocols. The average exaggeration rating (the item-by-item measure) was higher in the entertainment condition than in the accuracy condition; however, exaggeration ratings were not on average higher in the entertainment than no-talk condition. There were no differences across conditions in the language used in free recall; the 'language of storytelling' did not mark recall in the entertainment condition as it had in the retellings. Finally, participants in the entertainment condition did not make

more intrusions in free recall, even though they had in the retellings. Thus, largely absent in free recall were the qualitative characteristics that had made the entertaining retellings good stories. When asked to tell an accurate recounting of the original story, participants in the entertainment condition were able to monitor for and edit many of the intrusions that they had made during retelling. They were able to switch from using the 'language of storytelling' and attempted to recall the story as it had originally appeared.

However, entertainment retellings led to less complete final recall protocols than those produced by participants in the accuracy condition. This may not seem surprising since accuracy participants rehearsed more story events than did entertainment protocols, and numerous studies have shown that rehearsal enhances free recall (Roediger et al., 1993; see also Bergman & Roediger, 1999). However, amount rehearsed cannot explain the full pattern of results observed. Entertainment participants retold many more events than no-talk participants (who did not rehearse at all)—and these two groups did not differ in amount recalled. Thus, amount of rehearsal is not the only variable predictive of later recall. Rather, we argue that two other factors are important: the quality of the rehearsal and the match between the retrieval structures at retelling and test.

First, the quality of the rehearsal has consequences for memory. As described already, entertainment retellings differed from accuracy rehearsals in a number of ways, including number of intrusions and on average lower ratings of retold events' accuracy and detail. This poorer rehearsal led to lower average ratings of accuracy and detail in free recall, as compared to the accuracy condition. Critically, these accuracy/detail ratings were no higher than in the no-talk condition. So not only did entertainment participants recall about the same number of events as participants in the no-talk condition, the two groups provided equivalent levels of detail for recalled events. Telling a good story may be fun for the listener, but such inaccurate rehearsal does not help participants to later provide detailed accounts. Even within the accuracy condition, there was a strong positive correlation between average accuracy ratings of the retellings and the recall protocols; if a retelling was qualitatively poor (in terms of accuracy), so was the resultant recall protocol.

Second, the match between the retrieval structures at retelling and test is key. Several data points are important for this argument. Given that an event had been included in the retellings, the probability of later recalling it was still significantly higher in the accuracy condition ( $M = 0.90$ ) than in the entertainment condition ( $M = 0.84$ ). Conversely, for events that were not retold, the probability of later recall was higher in the entertainment ( $M = 0.22$ ) than accuracy condition ( $M = 0.16$ ). These data suggest that participants in the entertainment condition had to switch their mode of remembering at final recall; that is, the way entertainment participants recalled in the final session was different from the way they retold their entertaining stories. The organization created during storytelling (that is, during entertainment retellings) was not utilized when attempting to remember original events as accurately as possible. Instead, participants switched organizations, possibly to a more chronologically organized one. This switched perspective at recall aided retrieval of previously unrehearsed items, but hindered recall of events tightly tied to the storytelling perspective. In this way, the results are similar to those found with traditional perspective shifts in prose materials (e.g. the switch from a home-buyer to a burglar; Anderson & Pichert, 1978). When participants were provided with full retrieval cues (e.g. on the story event recognition test), then there were no differences across conditions. Internally-generated retrieval structures are generally much more important on free recall tests than in situations like recognition that minimize the necessity of generating one's own retrieval structure and cues (e.g. Bower, Clark, Lesgold, & Winzenz, 1969; Eich, 1980).

This interpretation is consistent with transfer appropriate processing (TAP; Morris, Bransford, & Franks, 1977; Tversky, 1973). That is, when similar retrieval processes are used at retelling and at final recall, memory benefits—as in the accuracy condition. When there is a mismatch between the retrieval processes used at retelling and final recall, memory suffers—as in the entertainment condition. While we agree that TAP theory explains the resultant memory performance, we note that TAP predicts nothing about the specific retrieval processes in question. That is, TAP specifies nothing about the processes used to create a story in the entertainment condition, or a veridical account in the accuracy condition.

We wish to comment on one additional implication from the finding of equivalent exaggeration and accuracy ratings in the entertainment and no-talk conditions. Apparently, minor exaggerations are a natural result of recollecting a story. Accuracy participants (who were rated as more accurate and less exaggerated in final recall) had practised minimizing exaggerations during their retellings, because of their explicit accuracy goal. Entertainment participants toned down their exaggerations somewhat in final recall (as suggested by the lack of differences across groups in number of intrusions), but they still exaggerated at similar levels as the no-talk group. Exaggeration is a part of recalling a story, unless one has specifically practised not doing so.

The present study suggests that accuracy-focused rehearsal yields better memory for original events than does telling an entertaining story. However, it is important to recognize that even attempts to rehearse accurately result in distorted memory. In the present study, accuracy participants recalled only 58% of the story events. Furthermore, accuracy participants included just as many intrusions in their free recall as participants in the other two conditions, indicating that the memories of accuracy participants were also prone to distortion. This is consistent with other research on memory for more complex stimuli. For example, Bergman and Roediger (1999) found memory distortions and errors in recall of a prose passage even when participants had received strict accuracy instructions. Even when people attempt to be accurate, rehearsal of complex material may be reconstructive or selective, leading to memory errors.

Researchers often criticize Bartlett's (1932) study of remembering because his participants reconstructed the story instead of attempting to repeat the story as accurately as possible (Gauld & Stephenson, 1967; Roediger et al., 1993). However, the process of reconstruction that occurred in Bartlett's study may be more similar to the natural process of retelling events than to rehearsal emphasized in typical laboratory experiments. In real life, we do not simply recount events in order as accurately as possible. We have a goal in mind during retelling that determines what event details are included, how they are described, and how we weave them together. These constructive processes have consequences for memory. If we want to study how memory processes work in real life, we need to continue conducting studies that more closely mimic how memory is used naturalistically.

### ACKNOWLEDGEMENTS

We are grateful to many people for their help with this project. We thank Ewart Thomas and Heidi Sivers for providing many helpful comments and support in the early stages of this research. Kelly Adamson, Christine Arnold, Julia Beeson, Aleksandra Dudukovic, Devin Gabriel, Beatriz Hernandez, and Addie Strumolo were excellent research assistants.

We also thank Eric Bergman and two anonymous reviewers for their insightful reviews and valuable comments. This research was supported by a Stanford University Undergraduate Research Opportunity grant to the first author, and a National Research Service Award from the National Institutes of Mental Health, #1F32MH12567, to the second author.

## REFERENCES

- Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning & Verbal Behavior*, *17*, 1–12.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bergman, E. T., & Roediger, H. L., III. (1999). Can Bartlett's repeated reproduction experiments be replicated? *Memory and Cognition*, *27*, 937–947.
- 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 & Verbal Behavior*, *8*, 323–343.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*, 717–726.
- DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality & Social Psychology*, *74*, 63–79.
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality & Social Psychology*, *70*, 979–995.
- Eich, E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory & Cognition*, *8*, 157–173.
- Feldman, R. S., Forrest, J. A., & Happ, B. R. (2002). Self-presentation and verbal deception: do self-presenters lie more? *Basic and Applied Social Psychology*, *24*, 163–170.
- Gauld, A., & Stephenson, G. M. (1967). Some experiments relating to Bartlett's theory of remembering. *British Journal of Psychology*, *58*, 39–49.
- Hyman, I. E. (1994). Conversation remembering: story recall with a peer versus for an experimenter. *Applied Cognitive Psychology*, *8*, 49–66.
- Marsh, E. J., & Tversky, B. (in press). Spinning the stories of our lives. *Applied Cognitive Psychology*.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, *16*, 519–533.
- Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion*, *10*, 601–626.
- Roediger, H. L., III, & Gynn, M. J. (1996). Retrieval processes. In E. L. Bjork, & R. A. Bjork (Eds.), *Memory* (pp. 197–236). New York: Academic Press.
- Roediger, H. L., III, Wheeler, M. A., & Rajaram, S. (1993). Remembering, knowing, and reconstructing the past. *The Psychology of Learning and Motivation*, *30*, 97–134.
- Sedikides, C. (1990). Effects of fortuitously activated constructs versus activated communication goals on person impressions. *Journal of Personality and Social Psychology*, *58*, 39–408.
- Turner, R. E., Edgley, C., & Olmstead, G. (1975). Information control in conversations: honesty is not always the best policy. *Kansas Journal of Sociology*, *9*, 69–89.
- Tversky, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology*, *5*, 257–287.
- Tversky, B., & Marsh, E. J. (2000). Biased retellings of events yield biased memories. *Cognitive Psychology*, *40*, 1–38.
- Wade, E., & Clark, H. H. (1993). Reproduction and demonstration in quotations. *Journal of Memory and Language*, *32*, 805–819.