

# Retelling Narratives as Fiction or Nonfiction

**Deborah J. O. Hendersen (djoh@stanford.edu)**

Jordan Hall, Bldg. 420, 450 Serra Mall.  
Stanford, CA 94305

**Herb Clark (herb@psych.stanford.edu)**

Jordan Hall, Bldg. 420, 450 Serra Mall.  
Stanford, CA 94305

## Abstract

Participants read and then retold narratives that were labeled either fiction or nonfiction. They showed a robust fiction superiority effect, recalling 20 to 50% more words and details when a narrative was labeled fiction. Fiction narratives also contained more of the language found in the original narrative. The fiction superiority effect held whether the retellings were oral or typed, and with and without an audience. This effect was found for recall but not recognition memory.

**Keywords:** fiction; nonfiction; narrative understanding; recall; recognition

## Introduction

On October 30th, 1938 the *War of the Worlds* radio production scared families into fleeing their homes because they could not tell that the production was a fiction. The next day the New York Tribune highlighted the problem by saying “Hitler managed to scare all of Europe to its knees a month ago, but he at least had an army and an air force to back up his shrieking words...But Mr. Welles scared thousands into demoralization with nothing at all.” The reason Welles’s production could be described as “nothing at all” is because it fails to follow one of Grice’s (1975) conversational maxims, the Maxim of Quality. The Maxim of Quality says: Do not say what you believe to be false; do not say that for which you lack adequate evidence. Fiction openly flouts this maxim, and in doing so raises the question of how we are to understand it.

Having abandoned the Maxim of Quality, it becomes somewhat unclear how two of Grice’s other maxims are to be applied to fiction. First, the Maxim of Quantity which states: Make your contribution as informative as is required for the current purposes of the exchange; do not make your contribution more informative than is required. And Secondly, the Maxim of Relation: Be relevant. But how do you judge a piece of information informative or relevant when you know that it is false? When reading nonfiction, information is deemed informative or relevant at least in part because the reader understands it to be historical or present-day fact outside of the narrative. The reader of fiction, however, must look within the narrative to understand why a piece of information is informative or relevant.

The question this paper asks is does it matter whether a narrative is fiction or nonfiction? While readers of fiction must look within a narrative to understand what is informative or relevant about a particular action, looking within a narrative is hardly unique to fiction. Narratives are narratives (rather than say lists of actions) because the actions included are related to one another. Indeed, much research has focused on how readers understand the internal structure of narratives, and it is now generally accepted that in order to understand narratives readers must be able to understand the cause-and-effect relations that exist between events (e.g., O’Brien & Myers, 1987; Trabasso, Secco, & van den Broek, 1984; Trabasso & van den Broek, 1985; Varnhagen & Goldman, 1986). To do this, readers track characters through time and space and monitor their concerns (Graesser, Millis, & Zwaan, 1997; Zwaan & Radvansky, 1998). In light of this, the question becomes if we hold all of the characters, their actions, the cause-and-effect relations in a given narrative constant, will participants think about the narrative differently simply because they believe it to be fiction or nonfiction?

The labels “fiction” and “nonfiction” are external to narratives, so it is important to recognize that people trying to understand narratives are influenced by background knowledge (e.g., Gerrig & McKoon, 1998; Graesser, Singer, & Trabasso, 1994) and these influences are at many levels including both the lexical and the script levels (e.g. Kambe, Rayner, & Duffy, 2001; Cook & Myers, 2004).

Not only does outside information help readers to understand narratives, but outside influences can alter how a reader approaches a narrative. Lorch, Lorch, and Klusewitz (1993) asked readers about the different kinds of reading tasks they experienced and participants distinguished two types of reading tasks: reading for study purposes and reading for entertainment purposes. Consistent with this division, Navaez, van den Broek, Ruiz (1999) found that readers with a study purpose were more likely to engage in repeating and evaluating the text than were readers with an entertainment purpose. Indeed, Zwann (1994) found that when readers believe they are reading a newspaper article they are concerned about what is true about the world, however they attend to more of the wording and stylistic devices when they read literature.

To see how construing a narrative as fiction or nonfiction would alter the way participants thought about it, we

harkened back to a method used first to great effect by Bartlett (1932). We asked participants to retell the narratives that they had just read, and we made particular effort to put participants in a situation where they either had an audience or anticipated an audience to encourage them to retell the narrative as a narrative, rather than as a memory exam. By examining the retellings, Bartlett was able to demonstrate that participants understand narratives in light of their preexisting knowledge. In this paper, we present evidence for what we will call the *fiction superiority effect*: people recall more information from narratives that are fictional than from narratives that are nonfictional.

## Experiment 1

The idea behind Experiment 1 was to test how participants would retell the same narratives when they were construed as fiction and as nonfiction.

### Method

The participants in all the experiments reported here were Stanford University undergraduates fulfilling a course requirement. There were 23 pairs of participants in this experiment. Each member of the pairs read and retold two of the passages and listened to their partner retell the other two.

We found four narratives that could sensibly be labeled either fiction or nonfiction. These were interviews and presented them to participants under one of two framings:

Fiction:

You will be reading an interview drawn from a contest on NPR where people were asked to imagine and make-believe about a variety of professions they had never held before. These people were then interviewed about their imagined profession, and you are going to be reading one of these interviews. Because these are just people talking you will periodically see things like ‘laughs’ in brackets, because these are just transcripts of people making up information as they go along, so please don’t be thrown by that.

Nonfiction (which represents the true nature of the narratives):

You will be reading an interview drawn from a book called *Gig*, which is a sort of modern-day version of Stud Turkel’s *Working*. Basically, the editors hopped in a van, drove across America, and interviewed real people about their jobs. You are going to be reading one of these interviews. Because these are just people talking you will periodically see things like ‘laughs’ in brackets, because these are just transcripts of real people talking about their lives, so please don’t be thrown by that.

The passages had been edited (by the *Gig* editors) to remove the question and answer format and contained only the words of the interviewee. Participants read a passage, completed a filler task for five minutes, and then retold the passage to a partner.

The four passages were interviews that had been edited by *Gig* to read like spontaneous narratives. They covered four professions (a flight attendant, a medicine woman, a labor-support doula, and a traveling salesman), averaging 1625 words in length. The narratives were picked to be as different as possible, evoking emotions that varied widely. Half the pairs of participants were assigned at random to the fiction condition, and the other half to the nonfiction condition; the order of the four passages was counterbalanced. The retellings were video-taped and transcribed, and the transcripts checked by independent listeners. The original passages were each divided into the smallest possible units of information, or details. For instance, “I never thought I’d be a flight attendant. I was bartending in Denver because, basically, I didn’t know what else I wanted to do” was divided into these four details:

1. did not think she’d be a flight attendant
2. was a bartender
3. in Denver
4. because she didn’t know what else to do.

Two raters blind to condition independently coded the details for each transcript, agreeing over 95% of the time. They met to reconcile the remaining differences.

The language of the each of retellings was also compared to the original passage using an ngram calculator. This generates a frequency count of all the sets of words ( $n$  in length) in each file, and calculates how many of the words occur in both passages by returning the dot product of each file’s distribution normalized by their respective magnitudes.

### Results and Discussion

Labeling the passages as fiction led to longer, more elaborate retellings than labeling them as nonfiction. The fiction retellings contained 1.5 times as many words as the nonfiction retellings. The difference averaged 648 to 434 words per retelling ( $t(90) = 4.50$ ,  $p < 0.0001$ , effect size 0.95, all tests throughout were 2-tailed), and was consistent across passages (flight attendant, 851 to 564; medicine woman, 543 to 307; doula, 638 to 455; and salesman, 561 to 411), see figure one. This difference is the *fiction superiority effect*. The fiction retellings also had 1.5 times as many details as the nonfictional retellings, an average difference of 27.9 to 18 ( $t(90) = 3.28$ ,  $p < 0.01$ , effect size 0.68). This difference, too, was consistent across passages (respectively; 48 to 29, 23 to 14, 21 to 15, and 21 to 14 for the four passages).

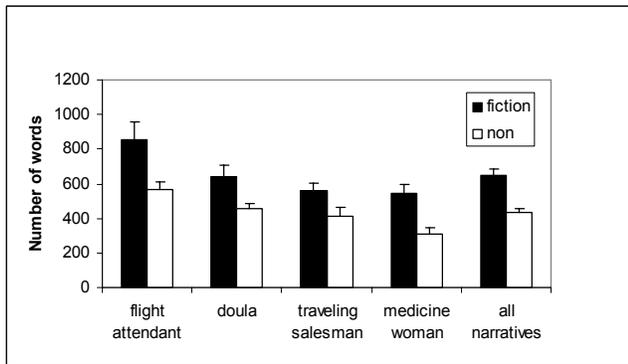


Figure 1: Average number of words retold, error bars indicated 1 SE.

We wondered if the increase in words was because participants in the fiction condition were including extraneous information that not included in the original passage. To check this, we asked whether the increase in details was proportionate to the increase in words in each of the retellings. Indeed, there was no difference in number of words per detail for the fiction vs. nonfiction retellings, 29.4 to 27.3 ( $t(90) = 0.701$ ,  $p = 0.49$ , effect size 0.15). So the fiction superiority effect was reflected not just in the number of words used, but in the amount of information recalled.

We also wondered if there was a pattern to the details included in fiction that were not included in nonfiction. Were, for instance, participants equally much including the causally relevant details, but the fiction condition participants were fleshing out their retellings? To test this possibility we ran Fisher's exact test on all of the details. Of the 557 details we coded for across all retellings, 32 differed significantly by condition (30 appeared more frequently in fiction than non, 2 appeared more frequently in nonfiction). Unfortunately, we are unable to conclude anything from these details as they make up 5.7% of the details we coded for, and as we used a 5% significance cut-off we are faced with the problem repeated measures and we must assume that these differences are due to chance.

The language used in the fiction retellings contained significantly more of the original narrative's language. After excluding outliers more than 2 standard deviations away, the fiction dot-products were significantly longer than nonfiction for single words (Fiction = 0.76, Non = 0.70,  $t(89) = 3.56$ ,  $p < 0.001$ ), word dyads (Fiction = 0.32, Non = 0.26,  $t(84) = 3.26$ ,  $p = 0.005$ ) and word triads (Fiction = 0.10, Non = 0.07,  $t(83) = 2.23$ ,  $p < 0.05$ ). This difference complements the Zwaan (1994) finding that readers of fiction pay attention to more of the surface details of a narrative than readers of nonfiction.

The most interesting and surprising finding of Experiment 1 is the fiction superiority effect. As the increase in length is proportionate to the increase in information included, this effect could reflect either a true increase in memory on

behalf of the fiction participants or it could indicate that fiction participants have a lower criterion for what they include in their retellings. Nonfiction participants may feel an obligation to include only the information that they are certain of in order to preserve Grice's *Maxim of Quality*, while participants in the fiction condition do not feel so constrained. Or the fiction group might have retained more information in order to preserve the feel of the narrative, while the nonfiction group, with equally good memory for the narrative, may simply have summarized the important points based on a higher criterion for reporting. Experiment 2 used recognition memory and signal detection to determine if there was a true memory advantage ( $d'$  in signal-detection terms) for the fiction participants, or merely a lowered criterion ( $c$ ) for including information.

## Experiment 2

The goal of Experiment 2 was to compare the recognition of the details originally coded for in Experiment 1 in passages labeled fiction or nonfiction.

### Method

We created 50 true and 50 false statements for each of the four passages from Experiment 1 using the original list of details. Participants were brought into the lab and randomly assigned to one of the four narratives. Half at random were given the fiction or the nonfiction frames used in Experiment 1. Participants read the narrative, completed a five-minute distracter task (standardizing stimuli for a map experiment), and were tested with a true-false test, using a button-box that collected reaction times as well. There were 96 participants, 24 for each passage.

### Results and Discussion

There were no differences in recognition between fiction and nonfiction passages. The average number of correct answers was almost identical for two conditions, fiction 79.3 vs. nonfiction 80.3 ( $t(94) = -0.71$ ,  $p = 0.48$ , ; effect size -0.14). The average number of correct answers ranged from 75 to 85% for the four passages, so the scores were certainly not at ceiling. The mean reaction times to answer were also nearly identical for the two conditions, fiction 3112 ms vs. nonfiction 3121 ms ( $t(94) = -0.058$ ,  $p = 0.95$ , ; effect size -0.012). So the time needed to verify details also did not differ for fiction and nonfiction, suggesting that participants found the task equally easy.

The crucial evidence comes from the signal detection analysis. There was no reliable difference in  $d'$  between fiction (1.74) and nonfiction (1.86) ( $t(94) = -1.11$ ,  $p = 0.27$ , ; effect size -0.23). Nor was there a reliable criterion shift ( $c$ ) from fiction (0.21) to nonfiction (0.23) ( $t(94) = -0.30$ ,  $p = 0.765$ , effect size -0.06).

So recognition was very different from recall. Framing a passage as fiction vs. nonfiction had a large influence on recall (the fiction superiority effect), but no influence on recognition—even though the true/false questions in Experiment 2 were a subset of the details counted in

Experiment 1. The point is made clear in the effect sizes. In Experiment 1, the effect size was a respectable 0.95, whereas in Experiment 2, the combined effect sizes of  $d'$  and  $c$  totaled a much smaller -0.29 in the wrong direction. The fiction superiority effect cannot be attributed either to a greater accessibility (in recognition) of details in passages labeled fiction or to a criterion shift in this accessibility.

Recognition, of course, is not the same as recall (Tversky, 1973; Bower & Winzenz, 1970). The details of a narrative may be equally well encoded in fiction and nonfiction (as seen in recognition memory), and yet recall could differ because of the connections among those details (Bower, Clark, Lesgold, & Winzenz, 1969; Mandler, 1967; Puff, 1970; Underwood, 1964). The fiction participants may apply retrieval strategies that exploit these connections more effectively than the nonfiction participants (e.g. Hacker, Dunlosky, & Graesser, 1998). If this is the case, then participants should show the fiction superiority effect even when they are told that a narrative is fiction or nonfiction after they have read it.

### Experiment 3

Experiment 3 was designed to address two questions about recall strategies.

(1) Would the fiction superiority effect remain when the retellings were written instead of oral? The effect might disappear if fiction narrators were more deliberate in their retelling, or if they were less interested in entertaining a responsive partner.

And (2) would the fiction superiority effect remain when participants were told that the narrative was fiction or nonfiction only after they had read it? The effect should disappear if it depends on the original encoding of the narratives as fiction or nonfiction. It should remain if it is a result of retrieval and retelling strategies.

### Method

One hundred and sixty Stanford students participated in this study.

The participants were assigned to one of four conditions in a two-by-two design. In the typed recall conditions they were told the narrative was fiction, or nonfiction, before they read it (as in Studies 1 and 2). In the post recall condition, participants were told that the narrative (referred to as “a passage”) was fiction or nonfiction only after they had read it (using the framings from Experiment 1). Otherwise, their instructions were identical to those in the typed recall condition.

Participants read either the salesman or the medicine woman passage from Experiments 1 and 2 and were then directed to a text editor and asked to retell it. They were told:

We are interested in how narratives are understood, discussed and communicated, and so what we’re doing is bringing participants into the lab and giving half the participants an original passage and half the participants a retelling. You were given the original passage, and so

what I would like you to do for the rest of the time is retell this passage, just as you would to a friend you thought was interested, so I can give that retelling to another participant and have them do something else with it.

Participants were given as much time as they needed to retell the narrative using a keyboard to enter it in the text editor.

Thirteen participants were eliminated for exceeding an a priori standard of two standard deviations from the mean of their condition, leaving 71, and 76 in the typed recall, and post recall conditions.

### Results and Discussion

Switching from oral retellings to typed retellings reduced but did not eliminate the fiction superiority effect. The fiction retellings had more words by a factor of 1.3, 449 to 359, ( $F(1, 70) = 7.89$ , effect size 0.66). This pattern was seen within each of the narratives. The traveling salesman fiction retellings had 1.4 times as many words on average as the nonfiction retellings, 462 to 342 ( $F(1, 34) = 5.071$ ,  $p < 0.005$ , effect size 0.75). The medicine woman fiction retellings were 1.2 times longer than nonfiction retellings, 435 to 356 ( $F(1, 33) = 4.70$ ,  $p < 0.05$ , effect size 0.74). The fiction superiority effect is clearly not dependent upon either oral retellings, or having a responsive audience. Even when given time to review and edit their retellings, participants in the fiction condition retold longer narratives than participants in the nonfiction condition.

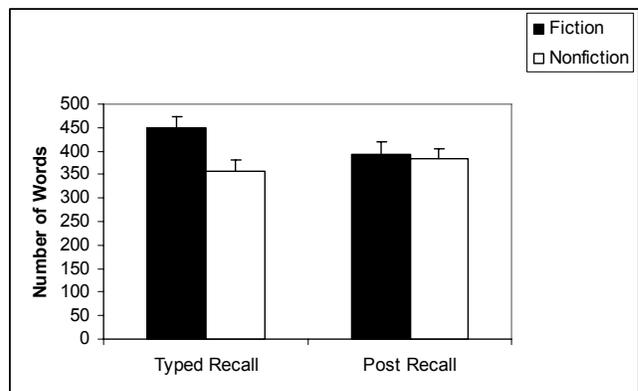


Figure 2: Average number of words retold, error bars show 1 SE.

The fiction superiority effect was eliminated in the post recall condition where participants were told the narrative was fiction or nonfiction *after* they had read the narrative. The fiction retellings were not significantly longer than the nonfiction retellings (fiction: 393, nonfiction: 385,  $F(1, 74) = 0.057$ ,  $p = 0.81$ ). There was not even a consistent pattern between the two passages. For the traveling salesman, the fiction superiority effect was reduced to 1.15, fiction 477 and nonfiction 417. This effect is not reliably larger than 1.0 (477 to 416,  $F(1, 38) = 1.62$ ,  $p = 0.21$ , effect size 0.20). For the medicine woman passage nonfiction

retellings were longer than fiction ones, 351 to 293 ( $F(1, 34) = 2.95, p = 0.095$ ). Across both passages the fiction superiority effect was larger for the typed recall condition than the post recall condition 1.3 to 1.02 ( $F(1, 143) = 3.85, p = 0.05$ ), see figure two. These findings argue that the fiction superiority effect is not a product of retrieval or retelling strategies. Instead, this effect appears to occur when participants are encoding the narrative.

### General Discussion

The fiction superiority effect is an advantage in recall for narratives that are construed to be fiction rather than nonfiction. The effect occurred whether the retellings were oral and with a responsive audience or written and without an audience. The effect occurred in recall, but not in recognition, and appears to be a product of how participants encode the narrative.

These findings suggest that not only do we approach narratives differently depending on the task at hand (i.e. study or entertainment) but that we approach narratives differently based on their genre. Readers of fiction are not only more sensitive to surface level features, such as the language used by the author, but they retain more of the content of the passage as well. This retention advantage is seen only recall, not recognition, which implies that the difference in fiction and nonfiction resides in how the reader structures the information in memory. Structure, of course, is a key component to narratives, with causal structure being vital to the reader's understanding of a narrative.

One possible basis for the fiction superiority effect is that details serve different functions in fiction and nonfiction. Authors of nonfiction are constrained by the Maxim of Quality and the historical facts they are relating. When readers of nonfiction are faced with a detail, they understand why it is informative and relevant: because it was true historically. If the author says it was raining, then it must in fact have been raining. Authors of fiction, in contrast, have abandoned the Maxim of Quality and are free to invent, manipulate, and elaborate details in ways that promote their narrative. When readers of fiction are told it was raining, they must ask why the author thought rain was informative or relevant, and they may assume that this detail motivates other features of the narrative. In *The Great Gatsby*, almost every major event is highlighted with details of the weather that signal the emotional state of the characters. Similarly, in *Romeo and Juliet*, the fight between Tybalt and Mercutio takes place on a spectacularly hot day. That is, readers of fiction may look for relevance by expecting to find thematic connections among the details, whereas readers of nonfiction are less pressed for explanations of relevance and should not. If readers of fiction are creating these interconnections between details to create relevance, then the fiction superiority effect would only appear in recall, where thematic connections would facilitate retrieval, but not in recognition, where they would not.

People regularly label fiction differently from nonfiction. Libraries contain separate sections on fiction and nonfiction.

Films are labeled as fiction or documentaries. Conversationalists mark some of their stories as ones that "really happened to me" and others as not. The fiction superiority effect is evidence that the fictional status of a narrative affects the recall of its content. Thus, such labels not only indicate whether the reported events are invented or real, but also trigger different processing strategies on the part of the reader.

### Acknowledgments

We would like to acknowledge the very helpful comments of Michael Ramscar, Dan Yarlett, Elsie Wang, and all of the members of SLUGS.

### References

- Bartlett, F. C. *Remembering: a study in experimental and social psychology*. (1932). 317 pp. Oxford, England: Macmillan.
- Bower, G.H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized world lists. *Journal of Verbal Learning and Verbal Behavior*, 8, 323-343.
- Bower, G. H., & Winzenz, D. (1970). Comparison of associative learning strategies. *Psychonomic Science*, 20(2), 119-120.
- Cook, A. E., & Myers, J. L. (2004). Processing discourse roles in scripted narratives: The influences of context and world knowledge. *Journal of Memory & Language*, 50(3), 268-288.
- Gerrig, R. J., & McKoon, G. (1998). The readiness is all: The functionality of memory-based text processing. *Discourse Processes*, 26(2-3), 67-86.
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, 48, 163-189.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological review*, 101(3), 371-395.
- Grice, H. P. 1975. "Logic and conversation." in Cole, Peter, and J. L. Morgan (Eds). 1975. *Syntax and semantics: Speech acts*. Volume 3. New York: Academic.
- Hacker, D. J., Dunlosky, J., & Graesser, A. C. (1998). Metacognition in educational theory and practice (Mahwah, NJ: Lawrence Erlbaum), 1998.
- Kambe, G., Rayner, K., & Duffy, S. A. (2001). Global context effects on processing lexically ambiguous words: Evidence from eye fixations. *Memory & Cognition*, 29(2), 363-372.
- Lorch, R. F., Lorch, E. P., & Klusewitz, M. A. (1993). College students' conditional knowledge about reading. *Journal of educational psychology*, 85(2), 239-252.
- Mandler, G. (1967). Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 327-372). New York: Academic Press.

- Narvaez, D., van den Broek, P., & Ruiz, A. B. (1999). The influence of reading purpose on inference generation and comprehension in reading. *Journal of Educational Psychology, 91*(3), 488-496
- O'Brien, E. J., & Myers, J. L. (1987). The role of causal connections in the retrieval of text. *Memory & Cognition, 15*(5), 419-427.
- Puff, C. R. (1970). Role of clustering in free recall. *Journal of Experimental Psychology, 86*, 384-386.
- Trabasso, T., Secco, T., & van den Broek, P. (1984). Causal cohesion and story coherence. In H. Mandle, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83-111). Hillsdale, NJ; Erlbaum.
- Trabasso, T., van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24*, 610-630.
- Tversky, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology, 5*, 275-287.
- Underwood. B. J. (1964). The representativeness of rote verbal learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 47-78). New York: Academic Press.
- Varnhagen, Connie K; Goldman, Susan R. Improving comprehension: Causal relations instruction for learning handicapped learners. *Reading Teacher. Vol 39*(9) May 1986, 896-904.
- Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*(4), 920-933.
- Zwaan, Rolf A; Radvansky, Gabriel A. Situation models in language comprehension and memory. *Psychological Bulletin. Vol 123*(2) Mar 1998, 162-185.