ABSTRACT. The present study investigated the mnemonic effects of using a joke and the influence of the location of the joke within a speech. In Experiment 1, participants heard 2 passages—one beginning with a humorous limerick and the other with a nonhumorous one. In Experiment 2, the limericks were presented at the end of the passages. Across both experiments, humor enhanced memory only for the humorous limericks, $t(44) = 3.22, p < .01; t(35) = 2.59, p < .002$, and did not influence memory for the rest of the passage. These results were discussed in context of the current literature, and future directions for study were outlined.

An elderly husband and wife visit their doctor when they begin forgetting little things. Their doctor tells them that many people find it useful to write themselves little notes. When they get home, the wife says, “Dear, will you please go to the kitchen and get me a dish of ice cream? And maybe write that down so you won’t forget?”

“Nonsense,” says the husband, “I can remember a dish of ice cream.”

“Well,” says the wife, “I’d also like strawberries and whipped cream on it.”

“My memory’s not all that bad,” says the husband. “No problem—a dish of ice cream with strawberries and whipped cream. I don’t need to write it down.” He goes into the kitchen for a while; his wife eventually hears pots and pans banging around. The husband finally emerges from the kitchen and presents his wife with a plate of bacon and eggs.

She looks at the plate and asks, “Hey, where’s the toast I asked for?”

Many guides designed to enhance presentation skills suggest that speakers should start their speeches with a joke in an effort to gain the attention of the audience and to make the presentation more memorable (e.g., Jeary & Cottrell, 2008; Kosslyn, 2007). The present study was designed to assess whether the presence of humor in a speech impacted memory for both the humorous stimulus (i.e., joke) as well as the surrounding content (i.e., speech).

Past research has shown that the cognitive processes involved in humor generally enhance memory for humorous events or material (e.g., Schmidt, 1994; Worthen & Deschamps, 2008). Most commonly, researchers have developed humorous stimuli in the form of sentences, pictures, and lectures (e.g., Kaplan & Pascoe, 1977; Schmidt & Williams, 2001). For instance, Schmidt (1994) assembled a number of humorous sayings and created nonhumorous counterparts (controls) for them. As an example, consider the humorous quote, “The only way to keep your good health is to eat what you don’t want, drink what you don’t like, and do what you’d rather not,” and its nonhumorous revision, “The only way to keep your good health is to eat good food, drink healthy drinks, and do healthy activities” (Schmidt, 1994, p. 954). More complex manipulations have integrated humor into hour-length lectures while manipulating the degree to which the humor related to the topic at hand (Kaplan & Pascoe, 1977). Although it seemed straightforward, the manipulation of the humor was quite difficult because humor is not universal—what is funny to one person may not be funny to the next (Kuiper et al., 2010). Humorous stimuli, therefore,
Researchers interested in the mnemonic effects of humor considered a wide variety of design issues, such as: (a) will participants be informed of the impending memory test (informed/intentional learning vs. uninformed/incidental learning) and (b) should humor be manipulated between-subjects (unmixed list) or within-subjects (mixed list)? For instance, Schmidt (2002) and Takahashi and Inoue (2009) explored these design issues and found that humor was more likely to elicit an effect with incidental learning, but that effects were also seen with intentional tasks. Although early work (e.g., Schmidt, 2002) seemed to suggest that humor effects were only present in mixed-list designs, where funny material could stand out from the context of unfunny material, more recent work has shown mnemonic benefits to humor with both mixed and unmixed lists (Takahashi & Inoue, 2009). That said, the standard design exploring humor and memory employs an incidental learning task with mixed lists (e.g., Schmidt, 1994; Schmidt, 2002).

Researchers have attempted to explain the mnemonic benefit of humor by appealing to a variety of related mechanisms. One explanation is that the surprise felt by encountering humorous, distinctive, or even bizarre material might lead to deeper (more meaningful) processing, greater elaboration or connection with long-term memory, additional rehearsal (or retelling) of the material, or even the storage of additional contextual information with the unique stimulus—each of which is sufficient to improve memory for such information (Schmidt, 2002). Other explanations have suggested that the mnemonic benefit of humor stems from the greater emotional or physiological arousal that accompanies humor.

In particular, the encoding-based explanations suggested that humorous materials were given additional rehearsal, elaboration, or privileged storage during encoding (e.g., Schmidt, 2002; Takahashi & Inoue, 2009). One implication of these explanations was that the extra rehearsal or elaboration for humorous stimuli would likely be at the expense of surrounding material. In terms of the present research question, this theory might suggest the presence of humor in a speech would negatively impact memory for speech itself. Evidence from Kaplan and Pascoe (1977) caused doubt in this explanation as they found that the material following humorous stimuli was remembered no better or worse than material following nonhumorous stimuli. However, one study does not provide a definitive test of a theory.

In contrast, retrieval-based explanations of the humor effect suggested that humorous materials were easier to distinguish from nonhumorous material in memory—that is, they were more distinctive (Schmidt, 2002). One could argue that the prevalence of humor benefits with mixed-list designs clearly support this theory—presumably, the humorous material stood out from the background of nonhumorous material. However, the presence of humor effects with unmixed designs was potentially problematic, unless one viewed the humorous material as standing out from all other information in memory, most of which is not likely to be humorous. Returning to the present question of interest, if humor influenced memory at the point of retrieval through distinctiveness, one would expect the mnemonic benefit to remain solely with the jokes and not with the surrounding speech information. It was difficult to tell from the current literature whether a retrieval-based account would predict a decline in memory for the speech information, as predicted by the encoding-based accounts.

The goal of the present study was to explore whether there was any truth to the old adage that starting a speech with humor would render the speech more memorable. The major explanations seem to predict that, while memory for the joke would be enhanced, the memory for the surrounding material would be either unaffected or impaired by the presence of humor. Findings in this direction would call into question the utility of starting with a joke.

**Experiment 1**

In Experiment 1, participants listened to two different passages of factual information. One passage began with a humorous limerick, whereas the other began with a nonhumorous limerick. To better simulate the conditions under which one might hear a speech, participants were given a brief orientation statement that asked them to imagine they were about to listen to a keynote speaker at a relevant conference before they actually listened to the limerick and passage. Following presentation, participants completed a brief distracting activity and then completed a recall and recognition test for all of the presented information. Based on previous research, we predicted that the humorous limerick would be remembered better than the nonhumorous limerick—the standard humor
benefit. Further, based on the encoding-based explanations of the humor effect, we predicted that memory for the passage might be impaired, especially the content that directly follows the limerick.

**Method**

**Participants.** Forty-five (6 men, 39 women) undergraduates were recruited from an Introduction to Psychology course at Lake Forest College and were compensated with extra credit in the course. Participants were run in groups of one to four people and completed the task at separate computers.

**Materials and Design.** Prior to the experiment, we created 16 limericks with ideas from Internet sources (i.e., Jokes.com, Laughoutloud.net) and modified them to create four types of limericks: funny and taboo, funny and not taboo, not funny and taboo, and not funny and not taboo. For example, a funny-taboo limerick began, “There once was a man named Dave. He had a dead whore in a cave,” whereas its nonfunny-taboo equivalent would start, “There once was a corporate whore named Dave. His investment plan was brave.”

In order to assess whether the categories were valid (humorous limericks were actually humorous; taboo limericks were taboo), 16 first-year students at Lake Forest College rated how funny and how taboo each limerick was on a simple 4 point Likert-type scale (1 = funny/taboo, 4 = not funny/not taboo). From 16 limericks, we selected two funny (mean rating: 3.29, taboo: 3.26) and two not funny (mean rating: 1.17, taboo: 1.06) versions that were matched on humor and taboo ratings (see Table 1). The four limericks were coded 1F (1 Funny), 1NF (1 Not Funny), 2F, and 2NF and appear in Table 1. We found that limericks with taboo or inappropriate language were the only limericks rated as funny, so the second original intended manipulation (presence of possibly taboo words in the limerick) was instead held constant by using the same taboo words in both the funny and the not funny limericks. Due to limerick 1NF’s low rating on the taboo scale, the second line was changed from “Who realized she could work no more” to “Who felt like an old corporate whore” and the fifth line was changed from “Would give her a job at the store” to “Would hire her to cut meat at the store” to add the same taboo features as Limerick 1F. Due to limerick 2NF’s low rating on the taboo scale, the word “shit” was added to the fourth line to create the same taboo effect as Limerick 2F.

Two content passages were created: one passage discussing the Spanish economy and the other outlining the recent popularity of Hollywood celebrities to work as voice actors for large commercial campaigns (see Table 2). The subject matter was decided based on simplicity of the topics and potential relation to the limericks. Each was designed to relate to the subject matter of a different limerick. The “Woman from Moher” limericks (1F; 1NF) flowed with the Spanish economy passage because both discuss issues surrounding job availability, whereas the “Man named Dave” limericks (2F; 2NF) related to the money theme in the Hollywood celebrity endorsements passage. The relationships are purposefully weak so that limerick memory and passage memory could not be confused at test. Both passages were of similar length (73 and 80 words, 4 and 5 sentences, respectively).

Across participants, each limerick was paired equally often with each passage and was placed at the beginning, before the passage content. Four audio recordings were created for the experiment. Each recording contained one limerick and one passage read in the same voice with natural inflection; the same male speaker was used in each recording.

Each sentence in each passage was converted into a series of propositions—singular idea units with truth value—for the purposes of scoring participants’ recalls. Overall, each passage contained 12 idea units. The limericks were also separated into idea units to ensure consistent scoring.

We created a two-alternative, forced-choice recognition test for the limericks and passages. Distractor sentences were created for the recognition test. Small details were changed in each sentence that altered its meaning. For example, “Many unemployed citizens have turned to the Spanish under-market, taking jobs that are ‘off the books’
and are paid ‘under the table,’” was changed to “Many unemployed citizens have turned to the welfare system, refusing to take jobs where they would be paid under the table.”

The independent variable in this study was the humor status of the limerick: funny vs. nonfunny in a within-subjects design with the variable manipulated factorially across participants. The dependent variable, memory for passage and limerick content, was measured through a free recall period and a recognition test. To ensure that the experiment was completely automatized, the experiment was created using E-Prime and conducted on IBM-compatible computers.

**Procedure.** Participants entered the laboratory where they were greeted and seated at a computer and read specific instructions regarding the nature of the experiment on the screen. Before each trial, participants were shown the following orientation statement to give a sense of a real situation (outside the lab): “Imagine that you are at a conference on (international economics or marketing in television). You are sitting down to watch a speaker and you will hear the opening paragraph of his 15-minute speech.” On each trial, participants heard an audio recording of a limerick followed immediately by an informational passage. Each participant saw either 1F/2NF or 2F/1NF; passage order was counterbalanced and the condition selection was preset in a random order. After the passage, they completed a 3-minute digit-tracking distractor task in which random numbers between 1 and 9 appeared on the screen and the participant identified whether the number was odd or even with a button press (‘o’ or ‘e’). Following the distracting activity, each participant completed a free recall task of the content of the limerick and the passage. They were instructed to remember as much from the limerick and the passage as possible and write it on a blank sheet handed out upon entrance to the lab by the lab supervisor. A timer on the computer screen updated every 10 s for 2 min and sounded when the interval was over, at which time the presenter collected the sheet.

Immediately following the free recall test, participants completed a two-alternative, forced-choice recognition test on the computer. On each trial, two sentences were shown on the screen (one from the limerick or passage and one of the distractors detailed above), and the participant was asked to choose the one actually heard and seen in that trial. Participants finished by completing a short rating sheet regarding their answers to three questions on a 5-point Likert-type scale (5 = very much, 1 = not at all) by circling the appropriate number on the page. The three questions were as follows: How well did the limerick relate to the topic of the presentation, did the limerick increase your interest in the paragraph that followed, and did the presence of the limerick distract you from the paragraph that followed?

**Data Analysis.** A rater assessed recall performance by determining which of the propositions in each limerick/passage combination were accurately remembered. The rater used lenient scoring criteria in which participants were awarded credit if they recalled the general idea of the proposition. The rater discussed any uncertain cases with a fellow researcher until consensus was reached and kept a running list of “acceptable phrases” to ensure consistent scoring of propositions throughout the process. For analysis, the mean proportion of propositions correctly recalled was calculated for accurate comparison between limericks and passages and among the different limericks.

**Results and Discussion**

As seen in Figure 1, the mean proportions of the propositions accurately remembered in the free recall portion of the experiment for the funny and nonfunny limericks were .60 (SD = .32) and .43 (SD = .23), respectively. A paired-samples t test showed that the difference between the means was statistically significant, t(44) = 3.22, p < .002, Cohen’s d = 0.61. The funny limericks, therefore, were remembered better than the nonfunny limericks. A separate paired-samples t test showed no significant difference between recall performance

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Passages for Experiments 1 and 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passage 1</td>
<td>As a result of the economic downturn in Spain, small businesses have closed and large companies have laid off workers across the country. Many unemployed citizens have turned to the Spanish under-market, taking jobs that are ‘off the books’ and are paid ‘under the table.’ After a series of bank reforms, Spain appears poised to rise out of the recession. However, the growth of the under-market will likely slow the economic recovery process.</td>
</tr>
<tr>
<td>Passage 2</td>
<td>After the recent changes of the culture in Hollywood, many famous actors have felt no shame in doing large commercial campaigns for big corporations. The majority has gone to voice-over work, but a few actually appear in the commercials. Of course, they are paid more to appear than for voice-overs. Up until recently, it was viewed as “selling out” to take a paycheck for doing nothing more than using fame to sell a product. Nowadays, it’s viewed as good marketing.</td>
</tr>
</tbody>
</table>
for the passages, \( t(44) = -0.97, p = .34 \).

Figure 2 displays the recognition accuracy for the funny and nonfunny limericks and passages. Paired-samples \( t \) tests found no significant differences between the proportions of sentences correctly recognized for limericks, \( t(44) = -0.01, p = .99 \), or passages; \( t(44) = -0.35, p = .73 \). Performance appeared to be near ceiling.

Although, overall, the type of limerick did not appear to influence recall or recognition of the passage, it is possible that the influence of the limerick might have been restricted to just the first few propositions of the passage. Figure 3 shows recall accuracy for the first three propositions following each type of limerick. A series of paired-samples \( t \) tests showed no significant differences between any adjacent propositions, suggesting that the limerick (funny or nonfunny) had no influence on memory for the adjacent material, \( t(44) = 0.65, p = .52; t(44) = 0.53, p = .60; t(44) = 0.53, p = .60 \), for first three propositions after the limerick, respectively.

Participants were asked to rate the relationship between the limerick and passage, whether the limerick increased their interest in the passage, and whether the limerick distracted them from the passage on a 1–5 scale. None of the ratings differed significantly between funny and nonfunny conditions, \( t(44) = -1.61, p = .12; t(44) = .18, p = .86; t(44) = .98, p = .33 \), relation, interested, and distraction listed respectively.

In Experiment 1, we successfully replicated the standard humor benefit by showing that funny limericks were remembered via free recall significantly better than nonfunny limericks. Further, the humor of the limericks in this case had no visible effects on memory for the rest of the passages. Participants also did not answer the follow-up questions significantly different across humor conditions. The results, then, do not support the old adage that starting a speech with a joke will render the speech more memorable because the memory enhancement was restricted to the joke.

**Experiment 2**

Experiment 2 was designed to assess whether the effect of the humorous limerick depended on its location within the passage, and whether the primacy location was the only reason for the humor effect seen in Experiment 1. Experiment 2 followed the same basic format as Experiment 1 except that the limerick was presented at the end of the passage. Based on the first experiment’s results, we expected to see the same effects shown
in Experiment 2—that is, a humor benefit for limerick content only and no effects on adjacent material or any of the rest of the passage material.

**Method**

**Participants.** Thirty-seven (27 women, 10 men) undergraduates were recruited from an Introduction to Psychology course at Lake Forest College and were compensated with extra credit in the course. Participants were run in groups of one to four people and completed the task at separate computers.

**Materials and Design.** The same limericks and passages used in Experiment 1 were used for Experiment 2; they appear in Tables 1 and 2. The recordings from Experiment 1 were edited such that the limerick followed the passage rather than leading into it in both the audio recording and the visual display. The same propositions, free recall and recognition instructions, were used in Experiment 2 as well. To ensure that the experiment was completely automated, the experiment was created using E-Prime and conducted on IBM-compatible computers.

**Procedure.** The procedure used was the same as that in Experiment 1.

**Data Analysis.** A rater assessed recall performance by determining which of the propositions in each limerick/passage combination were accurately remembered. The rater used lenient scoring criteria in which participants were awarded credit if they recalled the general idea of the proposition. The rater discussed any uncertain cases with a fellow researcher until consensus was reached and kept a running list of “acceptable phrases” to ensure consistent scoring of propositions throughout the process. For analysis, the mean proportion of propositions correctly recalled was calculated for accurate comparison between limericks and passages and among the different limericks.

**Results and Discussion**

One participant’s data set was removed from the analysis because of a procedural mistake—the participant received only one level of the manipulation. Hence, all subsequent analyses were performed using 36 participants. As seen in Figure 4, the mean proportions of the propositions accurately remembered in the free recall portion of the experiment for the funny and nonfunny limericks were .51 (SD = .12) and .36 (SD = .14), respectively. A paired-samples *t* test showed that the difference between the means was statistically significant, *t*(35) = 2.59, *p* = .014, Cohen’s *d* = 1.15.

The funny limericks were again remembered better than the nonfunny limericks. A paired-samples *t* test showed no significant difference between the mean proportion correct for the passages, *t*(35) = -0.03, *p* = .98.

Figure 5 displays the recognition accuracy for the funny and not-funny limericks and passages. Consistent with Experiment 1, paired-samples *t* tests found no significant differences between the proportions of sentences correctly recognized for limericks, *t*(35) = -1.59, *p* = .12, or passages, *t*(35) = -1.19, *p* = .24.

Again, although the type of limerick did not appear to influence recall or recognition of the passage overall, it is possible that the influence of the limerick might be restricted to just the last few propositions of the passage in this condition. Figure 6 shows recall accuracy for the last three propositions preceding each type of limerick. A series of paired-samples *t* tests showed no significant differences between any adjacent propositions, suggesting that the limerick was the only part of the stimulus affected by its humor or lack thereof for the last 3 propositions, in order: *t*(35) = 0.00, *p* = 1.0; *t*(35) = -1.44, *p* = .15; *t*(35) = -0.77, *p* = .45.

Only the comparison between conditions of the question, “Did the limerick distract you from the passage that preceded it?” showed a significant difference, *t*(35) = 2.19, *p* = .04, Cohen’s *d* = 0.50. The limericks were rated as significantly more distracting in the funny condition (*M* = 4.22, *SD* = .84) than in the nonfunny condition (*M* = 3.69, *SD* = 1.25). None of the other ratings differ significantly between funny and nonfunny conditions, *t*(35) = -0.82, *p* = .42; *t*(35) = -1.07, *p* = .29.

In Experiment 2, the standard humor benefit was successfully replicated by showing that funny limericks were remembered via free recall significantly better than nonfunny limericks. Consistent with Experiment 1, the presence of humor had no visible effects on memory for the rest of the passages. Participants did not answer significantly different across humor conditions with answers to the follow-up questions, although they did report feeling that the funny limericks were significantly more distracting. The results, again, do not support the old adage that starting with a joke will enhance the memorability of the speech.

**General Discussion**

Experiment 1 revealed a significant positive effect of humor on memory only when free recall was tested. Any residual effects (positive or negative) of
humor on the rest of the passage were not statistically significant. The humor benefit was not evident with the recognition test, which likely occurred due to a ceiling on performance levels. Additionally, the presence of humor at the beginning of a speech did not influence participants’ interest in the passage and the presence of humor did not seem to distract participants. These results were replicated in Experiment 2, even when the location of the limerick was changed to the end of the speech.

These experiments demonstrate that the basic humor benefit is replicable and robust, even with a novel procedure and stimuli. Though there was no influence of humor on the passages themselves, they were not manipulated for humorous content. They are still, however, an important part of the stimulus, as the lack of difference between memory for the passages in the humorous and nonhumorous conditions adds interesting information to the presence of the effect in this study. Our results are consistent with those reported by Kaplan and Pascoe (1977), who showed a humor effect for humorous examples given within a lecture but only for those specific examples and not for the actual lecture material.

Some other strengths of the current study are that it used a less typical methodology in that participants learned intentionally (the memory test was not a surprise) and another more typical method in that the humor condition was manipulated within-subjects and produced the effect. The current study also used the propositional method of scoring data, which is a more objective and operationalized way to score recall data than by number of words within a sentence (Kintsch, 1974). Moreover, one could also argue that the nonfunny limericks were quite bizarre, and the lack of memory enhancement for them provides evidence of the strength of the humor effect over the bizarreness effect (and any taboo effects on memory due to the counterbalancing of taboo words across limerick type).

The results of this study provide a challenge to the encoding-based explanations of the humor benefit which seemed to predict that the extra attention (or rehearsal) paid to the humorous material would be at the expense of the immediately surrounding material (e.g., Schmidt, 2002; Takahashi & Inoue, 2009). The results, however, may be consistent with a retrieval-based explanation of the humor effect, although these theories do not directly address memory for surrounding information.

One possible weakness of this study is the
lack of a manipulation check on the humorous material for each participant. In other words, the study did not include a check to see whether each limerick was funny or nonfunny as intended for each participant. Such a check was avoided in order to keep the manipulation unknown to participants until after the experiment was over. However, the pilot study pretested the humor of the limericks, so the manipulation was nonetheless expected to work and did work, as a humor effect was shown.

Another potential limitation of the current study is the lack of a strong content relationship between the limerick and its passage. While leading with information couched in a joke might only help memory for basic concepts, such enhancement would be very helpful for the most critical information in a presentation or lecture situation, and so a more closely related limerick might improve general concept memory in the desired manner. Similarly, the use of a shorter joke might also make the nature of the experiment less obvious and add to the strength of the effect. The short length of the stimuli may also limit the generalizability of the study to presentations or lectures made in a classroom or at a convention. In such real-life situations, a speech would last much longer than one paragraph, so the manipulation necessary to show a humor effect might need to be stronger if it is even, in fact, possible for a joke to affect memory for content of a speech or paragraph in a different study or situation.

Further, the type of joke employed here could also have limited generalization, because limericks are a very specific and obvious type of joke. We chose limericks because of the ease of matching them to each other and the restrictions put on this study by time and resources. However, the length and type of the joke did not appear to be a major issue because the effect was produced twice. Finally, the taboo words, though repeated in all limericks, may not have maintained taboo meanings in different contexts (specifically in 2F/2NF) and were not retested for taboo status after the addition of new words. Despite these limitations, the content remembered in free recall, that the humor of the limerick, was the important manipulation in all conditions, not its content of taboo words.

Future Directions
The present experiments could potentially open up some new possibilities for the field of research on the humor benefit. First, it might be interesting to control for individual difference in humor perception by creating stimuli that are tailored to each individual’s sense of humor. Although difficult, such a study could add much information to our knowledge of the effect as researchers could be more sure that each participant would find the stimuli humorous. It would also address a potential limitation of the current study, which is that even though the stimuli were rated as funny, they were not rated as highly funny. Despite the possibility of a more robust effect (or an effect on the passage) with funnier stimuli, the current study did replicate the effect quite successfully.

Next, it would be nice to examine the effect of humor in more real world situations (e.g., Kaplan and Pascoe, 1977). If researchers are to answer the question of how to truly remember things better, they must first test the theories in situations like lectures and speeches. For instance, one might show participants videos of lectures or presentations. Indeed, initially this research began with a pilot test of a funny and a nonfunny video of a short speech, but the experiment never got past the piloting stage because the funny stimulus was not rated as sufficiently funny. However, further research using the recall methods here and the propositional method of scoring might lead to further validation of the effect and suggestions for ways to use it in the real world.

References