

# The Effect of Context on Humor: A Constraint-Based Model of Comprehending Verbal Jokes

Heather H. Mitchell

*Department of Behavioral and Social Sciences  
Webster University*

Arthur C. Graesser and Max M. Louwerse

*Department of Psychology  
University of Memphis*

Two experiments were conducted to assess the effects of various constraints on the processing of jokes. Participants read humorous jokes and nonhumorous alternatives of the jokes, which were presented in 3 conditions that manipulated discourse context (comedy, political, and control). In Experiment 1, participants rated the funniness of texts and provided some recall data. In Experiment 2, participant's eye movements were collected to examine the effects of the different contexts on the online processing of the texts. Results confirmed that context constraints reliably impact how jokes are appreciated and processed, but the patterns of results were sometimes counterintuitive—namely, constraints of the discourse context appear to influence the processing of verbal jokes early on, whereas constraints associated with the text type strongly affect later processing and the results of processing, including recall. A constraint-based model is offered as a theoretical account for these findings.

A verbal joke is often described by researchers as having some form of opposition or incongruity followed by a resolution (Attardo, 1994; Schultz, 1976; Suls, 1972, 1977). Consider, for example, the following joke, which has some novel and unexpected perspectives on social customs:

---

Correspondence concerning this article should be addressed to Heather H. Mitchell, Department of Behavioral and Social Sciences, Webster University, St. Louis, MO 63122. E-mail: hmitchell33@webster.edu

A friend of mine had just returned from two weeks of vacation, and asked his boss for two more weeks off to get married. "What!," shouted his boss, "I can't give you more time now. Why didn't you just get married while you were off?" "Are you nuts?," replied my friend. "That would have ruined my entire vacation!"

Now imagine the difference between this joke being told by a prominent comedian versus a well-known politician. Although verbal jokes are often considered to be a comparatively decontextualized form of humor, compared with highly contextualized forms of wit and irony that respond to immediate local opportunities, we nonetheless suspect that comprehenders would interpret and remember the joke rather differently. The joke might come off as amusing when told by a comedian, like Jerry Seinfeld or Rita Rudner, but there is a risk that it would come off as tacky or socially awkward when told by George Bush or Hillary Clinton.

This study examines how and when context affects the interpretation and time course of joke comprehension. Limited research has been conducted investigating these phenomena. For example, Vaid, Hull, Heredia, Gerkens, and Martinez (2003) investigated when the humorous interpretation of a joke becomes accessible to a reader independent of contextual information. The humorous interpretation appears to first become activated when a reader processes an incongruity, or the part of the text that does not match a reader's expectations, such as the word *ruin* in the opening joke.

Additional research, which is more directly related to context, that has investigated the effect of authorship on language comprehension suggests that participants do interpret the same utterances differently when delivered by different speakers. For example, Asch (1952) reported that the identity of a speaker can affect the interpretation of language in persuasive messages and propaganda. Participants were asked to interpret excerpts of texts written by prominent public figures such as, "I hold it that a little rebellion, now and then, is a good thing, and as necessary in the political world as storms are in the physical." When students were told that this excerpt was written by Thomas Jefferson (the real author), American students interpreted the quotation as being about politics, whereas when the same excerpt was attributed to Lenin, it was interpreted as being about bloodshed.

Similarly, Katz and Pexman (1997) found that a speaker's reported occupation influenced participants' interpretations of an utterance. Participants read ambiguous statements that could be perceived as either ironic or metaphorical, such as "Her mind is an active volcano." Readers interpreted the same utterances by speakers from occupations rated as stereotypically high-irony (e.g., comedians, police, and factory workers) as more sarcastic and mocking than the same utterances made by speakers from occupations stereotypically rated as high-metaphor (e.g., clergymen, doctors, and teachers). If readers interpret jokes

in the same way they interpret political speeches (Asch, 1952) or the ironic or metaphoric statements used in Katz and Pexman, then jokes may well be interpreted differently when attributed to different speakers.

One might also predict that the process of reading jokes will be different when the jokes are attributed to different speakers. In a follow-up study, Pexman, Ferretti, and Katz (2000) investigated the online processing effects of the speaker's occupation on nonliteral utterances. When a target utterance was preceded by a mentioned occupation, participants spent longer reading the target utterance than when no occupation was mentioned. Further, the occupation conditions had significantly longer reading times (RTs) for the last word in the target statement compared to the no occupation condition. These data support the claim that participants are sensitive to the speaker's occupation and that this information influences the reading of the entire target statement.

A curious interaction was also found between speaker type and context type in the Pexman et al. (2000) study. When a target was uttered by a high-metaphor speaker, RTs were remarkably similar for both irony-inviting and metaphor-inviting contexts across all RT locations. There seems to be little ambiguity to resolve for a nonliteral target statement when it is spoken by a person from a high-metaphor occupation. In contrast, when a target was uttered by a high-irony speaker, different results occurred. When a high-irony speaker presented a statement in an irony-inviting context, additional processing time appeared to be needed immediately. The RT for the space following the last word in the target sentence was longer for the irony-inviting context than the metaphor-inviting context. However, when a target was uttered by a high-irony speaker in a metaphor-inviting context, additional processing time appeared later; the RT of the first word of the sentence following the utterance was longer. These findings suggest that in an irony-inviting context, the high-irony speaker information is processed early and aids in interpretation quickly after reading the target utterance. Alternatively, in the metaphor-inviting context, the ironic nature of the speaker is initially ignored but seems to be processed by the beginning of the following sentence. If the online encoding of jokes is similar to the encoding of other nonliteral texts, such as ironic statements (Pexman et al., 2000), then readers would be expected to encode jokes differently when there are different speakers.

Findings from Katz and Pexman (1997) also suggest that the amount of information recalled from a joke will differ because of the different speakers. Participants recalled utterances made by members of high-irony occupations better than the same utterances were made by members of high-metaphor occupations. In fact, some data suggest that recall improved when the speaker's occupation and context type were incongruent. Statements made by a member of a high-irony occupation were recalled best when given in a metaphor-inducing context, whereas statements made by a member of a high-metaphor occupation

were recalled best when given in the irony-inducing context. Katz and Pexman suggest that when the speaker's occupation and the discourse context led to different interpretations, then memory for the utterance was facilitated because of the effort required to disambiguate the utterance and interpret the speaker's intention. Similarly, Kemper and Thissen (1981) found that participants were more likely to remember the wording of a request that was incongruous with the social status of the receiver. For example, the request of a low-status person made in an impolite way to a high status person was recalled well. Findings like these lead to the prediction that the incongruity of a politician's telling a joke should facilitate recall of the joke compared to the congruous expectation of a comedian telling a joke.

More generally, broader theorizing about context and contextualization leads to the prediction that context should affect the comprehension of verbal jokes. It is, of course, difficult, if not elusive, to offer a precise technical definition of context (Duranti & Goodwin, 1992; Grimshaw, 2003), but it is frequently argued that talk, as a contextualized event, cannot be appropriately described, understood, or interpreted unless one examines the context (e.g., shared assumptions, cultural setting, or speech situation) in which the event is embedded. The notion of context involves the fundamental juxtaposition between a focal event and its surroundings (Goffman, 1974).

Nevertheless, other areas of research suggest that different speakers will not have a differential effect on how a joke is interpreted, read, or recalled. Specifically, according to the influential incongruity-resolution theory of humor comprehension and appreciation, no specific comprehension differences would be expected for the joke because of different speakers (Suls, 1972, 1977). The incongruity-resolution theory proposed a two-stage model of humor comprehension. The first stage is a semantic processing stage in which a cognitive representation of the material continues to be built until an incongruity is encountered. In jokes, this incongruity detection phase typically occurs at the punchline. Consider the joke we started with. Initially, a ruined vacation seems to have nothing to do with marriage. However, in the second stage, resolution, the comprehender attempts to make sense of the incongruity created at the point of the punchline: Why would a marriage ceremony ruin a vacation when both vacations and marriages are supposed to be positive experiences? Problem-solving mechanisms are needed to reinterpret various elements in the joke to explain the incongruity.

According to Sulz (1972), situational factors are assumed to influence humor appreciation in the same ways that they affect other kinds of information processing. These situational factors are not specifically defined, however, so no distinctive predictions are made about the impact of speaker identity on joke processing. Sulz (1972) did claim that the situational factors would influence the second resolution stage of processing, if at all; but, again, there are not enough

commitments in the model to generate any predictions. Consequently, if readers process jokes in two distinct processing stages, without special consideration of situational or external factors, then any differences in the interpretation, reading, or recalling of the joke would be because of differences in the joke itself and not because of different speakers. The general theory of verbal humor (GTVH), proposed by Attardo and Raskin (1991), provides a similar theoretical approach for analyzing verbal jokes. Although the GTVH is highly detailed regarding humor-eliciting parameters within a joke itself (i.e., the text stimulus), external factors are not specifically addressed. Therefore, differences in the reading, interpreting, or recalling of a joke because of its speaker would not be predicted by this position.

### THIS STUDY

Two primary manipulations were used in this study to empirically investigate the effect of context on the processing of verbal jokes. The first independent variable manipulated the *text* by presenting either a humorous joke with a punchline or a nonhumorous alternative to the joke that substituted the punchline with a culminating adage (i.e., a nonhumorous proverb). The second independent variable included three alternative *framing narratives*: (a) one in which a text was told by a comedian in a standup routine, (b) one in which a text was told by a politician in a political arena (e.g., speech, interview, or book), or (c) one in which a text was presented with no specified teller or situation. To maximize the sensitivity of this manipulation, the assigned framing narrative was reinforced by a convergence of three information channels: (a) the name of a famous comedian or politician, (b) a corresponding photograph of the famous person, and (c) a corresponding social setting.

Framing narratives of comedians and politicians were chosen because of the assumed constraints associated with each group of individuals. Whereas politicians are expected to say something informative, comedians are expected to be humorous, as well as informative. In other words, politicians have more options about what they are expected to present: A politician may either be funny in an effort to impress possible voters or serious to make a point. Note that other possible speaker occupations would be less likely to be associated with joke telling (e.g., American evangelists).

Based on previously discussed research, we outline two general theoretical positions: a *de-contextualized* incongruity–resolution position (based on Sul's, 1972, seminal incongruity–resolution theory of humor) and a *constraint satisfaction* position. Table 1 shows these two theoretical positions, along with their dependent measures and their corresponding predictions for the treatment conditions.

TABLE 1  
Predictions Based on Theoretical Positions

<i>Theoretical Position</i>	<i>Dependent Measure</i>	<i>Predictions for Conditions</i>	
De-contextualized incongruity-resolution	Funniness ratings, interest ratings, and recall	JC = JP = JN > AC = AP = AN	
	Total fixation time on context	JC = JP = AC = AP	
	Total fixation time on final line	JC = JP = JN > AC = AP = AN	
Constraint satisfaction position	Violated-constraints hypothesis	Funniness ratings	AP = AN > AC
		Regression fixation times on context	AC > JC = JP = AP
	Regressive fixation times on cotext	AC > JC = JP = AP = JN = AN	
	Validated-constraints hypothesis	Funniness ratings	JC > JP = JN > AC = AP = AN
		Regression fixation times on context	JC > = JP = AP > AC
		Regression fixation times on cotext	JC > = JP = AP = JN = AN > AC

*Note.* C = comedy context; P = political context; N = no context; J = joke text; A = nonhumorous proverb text with adage.

The constraint satisfaction position is similar to other constraint-based approaches proposed as models for language and discourse processing (MacDonald & Seidenberg, 2006; McRae, Spivey-Knowlton, & Tanenhaus, 1998). Constraint satisfaction models have been used to explain structural ambiguity resolution within nonhumorous texts (MacDonald & Seidenberg, 2006; McRae et al., 1998) and to specify when context affects the interpretation of nonliteral statements (Katz & Ferretti, 2003). Such models assume that the understanding of linguistic messages involves constructing an interpretation that fits the constraints of available sources of information better than alternative interpretations. Linguistic, as well as nonlinguistic, sources of constraints are used to evaluate multiple interpretations of a text. As in all types of language and discourse, a constraint based model could be applied to the processing of humorous text (Katz, 2005; Katz & Ferretti, 2003).

As previously discussed, the speaker of the text (Fitneva & Spivey, 2004) is one important constraint, as well as the discourse context preceding the text (Pexman et al., 2000). These two constraints correspond to the framing narratives of this study. Moreover, not all categories of constraints affect text processing equally (MacDonald & Seidenberg, 2006). Highly constraining sources of information have a greater impact than minimally constraining sources. For example, the comedy narrative is considered a highly constraining factor because the presented information is expected to be humorous. In contrast, a control narrative is less constrained and is not expected to greatly affect the reading, interpretation, or recall of the texts.

Regarding offline measures of processing, the de-contextualized position predicts differences only between the jokes and nonhumorous proverbs, as shown in Table 1. Jokes in all conditions were predicted to be seen as equally funny and significantly funnier than the adages, but not necessarily to be remembered better; increased memory for humorous material has not always been a consistent finding in the literature (Schmidt, 1994). For example, early studies that investigated the effects of humor on memory in education (e.g., Kaplan & Pascoe, 1977; Kintsch & Bates, 1977) and advertising (e.g., Duncan, Nelson, & Frontzak, 1984; Gelb & Zinkhan, 1986) obtained mixed results, perhaps because they did not control for potentially confounding factors. However, Schmidt (1994) conducted several well-designed studies that confirmed facilitative effects of humor on memory for humorous versus nonhumorous sentences. An example humorous sentence was, "Suburbia is where the developer bulldozed out the trees, then named the streets after them," and a corresponding nonhumorous sentence was, "In suburbia, the developer bulldozes out the trees, and the names of the streets come from trees."

The de-contextualized position also makes predictions about comprehension processes when inspecting eye movements. First, no differences are predicted between the total fixation times spent on the comedy and political narratives. Comparisons among fixation times on all three framing narratives are not possible because the control condition contain significantly fewer words than the other two. However, participants are predicted to spend significantly more time processing the punchlines compared to the adages. The incongruity-resolution theory (Suls, 1972, 1977) proposes jokes are processed in two phases, which should yield an increase in time to interpret the punchline compared with the adage. Increased processing times are predicted for the punchlines because of the additional processing needed to detect and resolve the incongruity associated with the jokes.

As shown in Table 1, there are two competing sets of predictions associated with the constraint satisfaction position with respect to encoding, appreciation, and recall: a *violated-constraints* hypothesis versus a *validated-constraints* hypothesis. The difference between these hypotheses stems from previous studies that have examined Baker's (1984) model of "standards" of comprehension. These standards include (a) the lexical standard, (b) the internal consistency standard, and (c) the external consistency standard. The simplest level of comprehension is the lexical standard that is met if a reader is able to read and understand all words in a text. A deeper level of understanding, the internal consistency standard of evaluation, is met when a reader is able to construct a meaningful, unified, situation model for the text. The deepest level of comprehension, the external consistency standard of evaluation, occurs when readers understand the relational sense of the text with respect to context and general world knowledge. Violated constraints should occur in this study when the external consistency

standard associated with a framing narrative is not met, whereas validated constraints should occur when the external consistency standard of a framing narrative is satisfied. For example, if a joke is not presented in a comedy narrative, there is a violated constraint, whereas a validated constraint occurs when a joke is presented in a comedy narrative.

Several predictions for appreciation and encoding are made based on the violated-constraints hypothesis. The obvious prediction would be that the substituted adage should significantly decrease funniness ratings compared to the joke with a punchline. Participants are also predicted to have increased regressive fixations to the framing narrative of a comedian and the context, or setup, of the nonhumorous adage condition. These regressive fixations may reflect a search for humor or for the solution of a violated constraint. For example, in a study investigating lexical-decision times between context-consistent, context-inconsistent, and control probes, participants had significantly slower lexical decision times for context-inconsistent (e.g., speaker was a "salesclerk"; probe was "jury") compared to context-consistent conditions (e.g., speaker was a "judge"; probe was "jury"; Fitneva & Spivey, 2004). If there is any similarity between how readers process the incongruity between speaker and context-inconsistent probe and the incongruity between the comedy narrative and the nonhumorous adage, then we could expect increased regressive fixations to the comedian narratives.

The validated-constraints hypothesis offers a different set of predictions for the interactive effects that framing narratives and type of text have on appreciation and encoding. According to the validated-constraints hypothesis, when a joke is presented in a comedian narrative, the expectation for humor is satisfied, and participants should rate the joke as reliably funnier than jokes in all other conditions. The humor ratings are not predicted to reliably differ between the political and control narratives. Neither of these conditions is constrained by the expectation of humor, so the intrinsic funniness of the text should primarily affect how funny these jokes are perceived to be.

An additional prediction of the validated-constraints hypothesis is that participants will have increased regressive fixation times to the comedy framing narrative and the context, or setup, of the joke when presented by a comedian. This eye-movement prediction is based on findings from an online study investigating the processing and recall of original (i.e., funny) versus weird or literal (non-funny) cartoons (Schmidt, 2002). In the study, participants' heart rates were recorded as they processed original, weird, and literal cartoons. Original cartoons, compared to weird and literal cartoons, led to significantly greater secondary heart-rate deceleration (i.e., difference between an initial acceleration component and a secondary deceleration component). According to Schmidt (2002), this increased secondary heart-rate deceleration provided evidence that the original cartoons stood out as a unique class of stimuli because this condition



is the only one in which the expectation of humor is fulfilled within the cartoon genre.

The secondary deceleration heart-rate measure also signaled the increased elaboration of humorous material late in the processing of the cartoons because the elaboration was not reflected in the initial deceleration rate. If there is any similarity between how readers processed the expected humor in original cartoons and how they process the expected humor in jokes presented by a comedian, then we predict increased processing times evidenced by increased regressive fixations to the framing narrative and joke cotext. As participants read a punchline to initially recognize the humor in a text, their additional processing time for the unique class of humorous jokes told by comedians would be manifest in regressive fixations back to the cotext or setup of the joke.

Conversely, the validated-constraints hypothesis predicts that nonhumorous adages presented by comedians will have significantly decreased regressive fixation times. Just as the secondary heart-rate deceleration measure did not suggest any elaborative processing for the literal and weird cartoon versions, the nonhumorous adage should not receive any additional processing when presented by a comedian. Participants are not expected to reread either the cotext of a nonhumorous adage or the framing narrative of an unfunny comedian. Last, no differences were predicted between the fixation times given to the cotexts presented in the political and control conditions because politicians have options about what they could present, and no constraints existed for the control condition.

## EXPERIMENT 1: SCALING TEXTS ON HUMOR

Experiment 1 was conducted primarily to scale materials on funniness so that only the funniest would be used in the eye-tracking experiment. Also, if texts presented in one framing narrative (e.g., a comedy narrative) were rated significantly funnier than the same texts presented in another (e.g., a control narrative), then there is evidence supporting the validated-constraints hypothesis. Experiment 1 was conducted secondarily to assess recall for jokes versus nonhumorous proverbs to test if recall for humorous material is better than nonhumorous material as suggested by Schmidt (1994).

### Method

#### *Participants*

The participants were 54 undergraduate students enrolled in psychology courses at The University of Memphis. They participated in the experiment as

an option to fulfill a course requirement or to receive extra credit. Participants were randomly assigned to one of six groups of equal size ( $n = 9$ ).

### Materials

*Texts.* All jokes were taken from *Reader's Digest: Laughter, The Best Medicine* (DeFord, Goodman, & LaFlaur, 1999). These jokes allowed for the same joke to be told in a variety of situations, including a control condition with no manipulated external factors. Thirty of these jokes were used with their original wording, including the punchline. Thirty nonhumorous proverb versions of the original jokes were created by removing the last line of the joke (i.e., the punchline) and inserting an appropriate adage. This manipulation provided nonhumorous texts that were optimally comparable to the original joke in content, style, and length, and that had a comparable culminating point and substantive meaning. These substitutions created instructive texts that loosely resembled fables, as in this example joke and corresponding nonhumorous proverb:

*Joke:* I was once hired by a supermarket and reported for my first day of work. The manager greeted me with a warm handshake and a smile, and then gave me a broom and said, "Your first job will be to sweep out the store."

"But I'm a college graduate," I replied indignantly.

"Oh, I'm sorry. I didn't realize that," said the manager. "Here, give me the broom—I'll show you how."

*Nonhumorous proverb:* I was once hired by a supermarket and reported for my first day of work. The manager greeted me with a warm handshake and a smile, and then gave me a broom and said, "Your first job will be to sweep out the store."

"But I'm a college graduate," I replied indignantly.

"Oh, I'm sorry. I didn't realize that," said the manager. "But you must start at the bottom and climb your way to the top."

*Framing narratives (speaker and setting information).* Framing narratives contextual information associated with each text. The following are examples of the comedian, politician, and control introductions:

*Comedian framing narrative:* "Following the finale of Jerry Seinfeld's top-rated primetime comedy TV series 'Seinfeld', Jerry embarked on an international tour of the classic stand-up material that originally made his name famous. The following joke is part of his stand-up material used in this tour:"

*Politician framing narrative:* "Colin L. Powell was a professional soldier for 35 years and rose to the rank of 4-star General. Secretary Powell has even written

his best-selling autobiography, *My American Journey*, in which he writes of the short story given below.”

*Control condition:* “Please read the following text carefully and naturally.”

*Recall measure.* Two different methods of recall were used: speaker-cued and cotext-cued. The speaker-cued recall was used to assess any differences in recall due to differences in speaker characteristics. In the speaker-cued recall, participants were asked to write about the text a particular comedian or politician told. The photographs and names of the alleged speakers were provided. Cotext-cued recall included all words of the cotext with the last line left blank. Participants were asked to fill in the correct punchline or adage for the text. Speaker-cued recall was collected prior to the cotext-cued recall.

### *Procedure*

In a paper-and-pencil task, each participant read 15 jokes and 15 nonhumorous proverbs. For any given participant, there were five texts presented in each of the six treatment conditions. On each of the 30 trials, participants first read the introduction information, then the actual text, and finally made a funniness rating about the text on a 6-point scale ranging from 1 (*not funny at all*) to 6 (*really funny*). After making a funniness rating, participants were not allowed to return to that text. After all ratings were made, participants completed the speaker-cued recall and then the cotext-cued recall. The ordering of the 30 texts was random and different for each participant.

### *Design*

This study followed a 2 (Text)  $\times$  3 (Framing Narrative) factorial design. The assignment of texts to the six treatments was counterbalanced across participants using a 6  $\times$  6 Latin square design, with 9 participants per counterbalanced assignment. The dependent variables included a funniness rating and two recall measures: (a) speaker-cued recall and (b) cotext-cued recall.

## Results and Discussion

All data from the 54 participants were included in analyses. The data were analyzed using a 2  $\times$  3 repeated-measures analysis of variance (ANOVA), as in prior studies that also used rating scales to investigate humorous material and its effects on memory (e.g., Berg & Lippman, 2001; Gunter, Baluch, Duffy, & Furnham, 2002; Schmidt & Williams, 2001). In all analyses,  $F_1$  refers to tests of subject variance, and  $F_2$  refers to tests of item variance. Follow-up tests of

TABLE 2  
Means and Standard Deviations for Humor Ratings, Speaker Cued Recall, and  
Cotext Cued Recall for Jokes and Nonhumorous Proverbs in Experiment 1

Context	Humor Ratings				Total	Speaker Cued Recall				Cotext Cued Recall			
	Jokes		Proverbs			Jokes		Proverbs		Punchlines		Adages	
	M	SD	M	SD		M	SD	M	SD	M	SD	M	SD
Comedian	4.29	1.36	3.06	1.45	3.68	1.28	0.57	1.42	0.76	3.22	0.90	2.96	1.20
Politician	4.21	1.33	2.94	1.44	3.56	1.11	0.35	1.51	0.75	3.40	0.88	2.68	1.30
Control	3.90	1.56	3.04	1.56	3.47					3.11	1.15	2.98	1.24
Total	4.03	1.42	3.01	1.48		1.20	0.46	1.47	0.76	3.24	0.98	2.87	1.25

interactions were calculated using  $F_1$  tests comparing contrasts. An alpha level of .05 was used for all statistical tests.

As shown in Table 2, participants rated the jokes as reliably funnier than the nonhumorous proverbs,  $F_1(1, 53) = 113.75, p < .01$  ( $MSE = 0.921, d = 0.69$ ) and  $F_2(1, 29) = 115.61, p < .01$  ( $MSE = 0.478, d = 1.73$ ), supporting the decontextualized position. However, participants also rated texts given in some framing narratives differently than when they were given in other framing narratives,  $F_1(2, 106) = 3.64, p < .05$  ( $MSE = 0.38, d = 0.32$ ) and  $F_2(2, 58) = 3.08, p = .05$  ( $MSE = 0.25, d = 0.48$ ). Specifically, participants rated jokes presented in a comedy narrative reliably funnier than jokes presented in the control condition,  $F_1(1, 106) = 23.11, p < .01$  ( $MSE = 0.387, d = 0.47$ ). Participants also rated jokes presented in a political narrative as reliably funnier than jokes presented in the control condition,  $F_1(1, 106) = 15.47, p < .01$  ( $MSE = 0.387, d = 0.39$ ). There was not a reliable difference between jokes presented in comedy and political narratives, however,  $F < 1$ . Participants also did not rate the nonhumorous proverbs differently based on the different framing narratives in which they were presented, as shown in Table 2.

In all, participants' humor ratings did not fully support the de-contextualized incongruity-resolution position, which predicted jokes in all conditions would be equally funny. Instead, participants rated jokes presented in the comedy and political narratives as reliably more humorous than those presented in the control condition. The humor ratings also did not support the violated-constraints hypothesis, which predicted participants would rate nonhumorous proverbs presented by comedians as reliably less funny than those presented in other conditions. The participants' humor ratings of the nonhumorous proverbs did not reliably differ among conditions.

The humor ratings best supported predictions associated with the validated-constraints hypothesis, although all predictions of this hypothesis were not

confirmed (see Table 1). Participants rated jokes presented in the comedy narrative, in which the comedian's constraint for being funny was met, as the funniest. However, participants also rated jokes presented in political narratives as reliably funnier than those given in a control narrative. (Although the humor ratings for comedy jokes had higher means than those for political jokes, this difference was not statistically reliable.) The validated-constraints hypothesis did not predict any differences between jokes given in political and control narratives. Participants apparently assigned increased humor ratings for jokes associated with any substantial framing narrative that accommodates jokes, as a political context allows.

### *Recall Measure*

As there is no well-defined analytical system for rating recall of humorous texts, a scale was devised for this study. (Detailed systems for analyzing recall of nonhumorous texts exist [e.g., Gernsbacher & Robertson, 2002], but the categories in these systems do not apply to humorous texts.) The speaker-cued recall of this experiment was evaluated using a scale with four values: (a) gist (general idea) of text definitely not captured, (b) gist of text not approximately captured, (c) gist of text approximately captured, and (d) gist of text definitely captured and humor or lesson of the text preserved. When interjudge reliability was obtained by calculating Cohen's kappa for the speaker-cued recall scores, the reliability between two raters was .87. The cotext-cued recall scale also had four values: (a) meaning of final line definitely not captured, (b) meaning of final line not approximately captured, (c) meaning of final line approximately captured, and (d) meaning of final line definitely captured. Reliability between two raters on cotext-cued recalls was .73. Means and standard deviations for both recall types are shown in Table 2.

For speaker-cued recall, participants remembered the nonhumorous proverbs significantly better than jokes,  $F_1(1, 53) = 6.15, p < .05$  ( $MSE = 0.392, d = 0.43$ ), as shown in Table 2. (Items analyses could not be conducted because so few items were included in the two types of recall.) Participants did not recall the gist reliably differently in the comedy and political conditions, and the interaction between framing narrative type and texts was not significant. For the cotext-cued recalls, participants recalled the punchlines significantly better than adages,  $F_1(1, 53) = 9.58, p < .01$  ( $MSE = 1.02, d = 0.46$ ), as shown in Table 2. Cotext-cued recalls did not reliably differ among the framing narratives, and the interaction between framing narratives and text was not significant.

These data also did not fully support the prediction of the de-contextualized position that memory for jokes would be better than memory for nonhumorous proverbs. Participants recalled nonhumorous proverbs reliably better in the speaker-cued recall, whereas punchlines were recalled reliably better in the

cotext cued recall. It should be noted that recall was not the primary interest in this experiment, however. The lack of reliable recall effects may be explained by how few texts were included in the recall phase of the study.

## EXPERIMENT 2: EYE-MOVEMENT STUDY

As eye movements can provide an important window for understanding processes and representations in cognitive tasks, we collected eye-tracking data while the participants read the jokes and the surrounding framing narratives. Eye-tracking technology allows investigation of comprehension at deeper levels of processing, as evidenced in eye-tracking studies on the comprehension of sentences (e.g., Rayner, 1998; Tanenhaus & Trueswell, 1995), texts (e.g., Just & Carpenter, 1980; O'Brien, Raney, Albrecht, & Rayner, 1997), and illustrated texts on everyday devices (Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005).

In this study, we also included a photograph of the speaker along with the framing narratives to assure participants were fully aware of who the speaker was. Although eye-movement studies on text and picture stimuli have typically segregated the two modalities, eye movements can also help assess how readers integrate pictures and text. For example, readers' eye fixations while processing information in illustrated texts on mechanical devices are initially driven by the text, but the illustrations are intermittently explored to help create deeper understandings (Graesser et al., 2005; Hegarty & Just, 1993). Similarly, when processing print advertisements, viewers' processing is mostly guided by the print information (Rayner, Rotello, Stewart, Keir, & Duffy, 2001). When participants read cartoons with captions, they process the picture and caption as relatively isolated events and only occasionally integrate the two (Carroll, Young, & Guertin, 1992). Available research supports the conclusion that participants typically read the text during early phases of comprehension and strategically inspect pictorial information.

Experiment 2 collected eye-movement data for two reasons. First, the framing narrative, photograph, and the text were presented simultaneously because we wanted to have some record of the allocation of processing resources among these components. Second, segregating initial, regressive, and total fixations allowed us to better test some of the predictions summarized in Table 1. We were able to identify when the framing narratives or photograph may influence the comprehension of texts by investigating all three fixation types. To answer the primary questions in this study, four regions of interest in the stimuli were examined: (a) the framing narrative, including speaker and setting information, if applicable; (b) the photograph of the speaker, if applicable; (c) the setup of the joke or nonhumorous proverb; and (d) the final line of joke or nonhumorous proverb. The framing narrative and photograph appeared in the top fourth of the

computer screen, whereas the text appeared in the rest of the screen. These four regions allowed sufficient discriminability for our purposes, although one could, of course, dissect the regions of interest into smaller units of analysis.

## Method

### *Participants*

The participants were 43 undergraduate students enrolled in psychology courses at The University of Memphis. The students participated as an option to fulfill a course requirement or to receive extra credit. Participants were randomly assigned to one of six groups. Data from 30 participants were included in analyses after removing the cases with calibration problems in the eye-tracker or motivation problems in the participants. As this study had a repeated-measures design, attrition rate should not seriously compromise the generality of the findings.

### *Materials*

*Equipment and calibration.* Eye movements were recorded by a Model 501 Applied Science Laboratory eye tracker (Bedford, MA). The equipment has a magnetic head tracker with a head mounted apparatus, which is advantageous because it allows participants to move their head during data collection and respond in a relatively natural way to the humor. Calibration could be maintained if a participant smiled, laughed, or nodded his or her head. An eye tracker with a chin rest or bite bar would not have allowed this natural type of reaction.

Computer software recorded the eye movements at a fine-grained level: The temporal resolution of the Model 501 eye tracker was 60 Hz, and the visual range was 50° horizontally and 40° vertically. Participants were calibrated both before the experimental session began and throughout the session to insure reliable data. The font size of the stimulus material ranged from Courier 12 to 24 points to allow for all of the information to fit on a screen without scrolling. The jokes and nonhumorous proverbs appeared in the 24-point font.

*Recall measure and interest ratings.* Two types of cotext-driven recall were collected. The first was a cloze recall task that deleted a sample of content words and the entire final line. The second recall task included the cotext and deleted the final line. For the cloze recall task, texts were presented on paper with keywords left blank. Keywords were diagnostic of the final lines—that is, the final lines were based on the semantic concepts that the keywords represented. The final line of the text was also left blank for the participant to fill in the correct punchline or adage. The same set of keywords was left blank for the

joke and nonhumorous proverb. Directly following the cloze recall task, the second recall for the same text was presented.

Interest ratings on the texts were also collected, rather than the humor ratings from Experiment 1, to index participants' motivation. Participants evaluated each text immediately after reading it. One prediction based on the de-contextualized position was that jokes in all conditions would be significantly more interesting than all nonhumorous proverbs (see Table 1).

### *Procedure*

Participants initially worked through practice slides before being calibrated to become familiar with the equipment. After calibration, participants worked through the remaining 24 texts at their own pace. The funniest 24 jokes and their corresponding nonhumorous proverbs from Experiment 1 were used in this experiment. For all texts, participants first viewed a framing narrative and photograph, if applicable, and then clicked the mouse to read a text that appeared on the computer screen directly beneath the framing narrative and photograph. Immediately after reading a text, participants rated the interest level of the text by clicking on the appropriate rating, which ranged from 1 (*not at all interesting*) to 6 (*really interesting*). Immediately following the eye-movement session of the experiment, participants completed a demographic questionnaire as a filler task and then completed the surprise recall tasks.

### *Design*

This study used a 2 (Text)  $\times$  3 (Framing Narrative) factorial design. The assignment of texts to treatments was counterbalanced across subjects using a 6  $\times$  6 Latin square design. We adopted the suggestion of Liversedge, Paterson, and Pickering (1998) to measure total fixation times, gaze durations (first pass fixation times), and regressive fixation times. Thus, the dependent variables included these three fixation time measures, interest ratings, and recall measures.

## Results and Discussion

### *Eye-Movement Data*

The eye-movement and recall data were analyzed using a series of 2 (Text)  $\times$  3 (Framing Narrative) repeated-measures ANOVAs. Eye movements on the framing narratives were analyzed using 2 (Framing Narrative)  $\times$  2 (Text) repeated-measures ANOVAs. Because the control framing narrative contained significantly fewer words, comparisons among all framing narratives were not possible.



Item analyses were conducted using the same corresponding designs. Follow-up tests of interactions and planned comparisons were calculated using  $F$  tests comparing pairwise contrasts. Table 3 displays means and standard deviations for eye fixations on the picture and framing narratives, whereas Table 4 displays means and standard deviations for eye fixations on the cotexts and final lines. These tables include gaze durations, total fixation times, and regressive fixation times for each of the four regions of interest (i.e., picture, framing narrative, cotext, and final line). Table 5 displays the means and standard deviations for the interest ratings and three types of recall.

The eye-movement data better supported predictions associated with the constraint satisfaction position than the de-contextualized incongruity-resolution position. The de-contextualized position predicted the framing narratives would not affect any eye movements associated with the jokes and nonhumorous proverbs (see Table 1). Contrary to this prediction, distinctive patterns of eye movements were found for the texts presented in different framing narratives, as shown in Table 4.

According to the validated-constraints hypothesis, participants should spend significantly more time on the comedy framing narratives when the constraint of being humorous was satisfied than on the other conditions. The data supported this hypothesis when comparing regression fixation times on the framing narratives of comedians and politicians. The interaction between framing narratives and text type was significant for the regressive fixation times on the framing

TABLE 3  
Mean Fixation Times (in Seconds) and Standard Deviations on Picture and Introduction Areas of Interest for Experiment 2

Areas of Interest	Dependent Measures	Comedian		Politician	
		Joke	Proverb	Joke	Proverb
Picture	Total fixations	0.51 (0.35)	0.44 (0.34)	0.50 (0.48)	0.52 (0.45)
Picture	Gaze durations	0.33 (0.24)	0.24 (0.15)	0.25 (0.17)	0.29 (0.19)
Picture	Regressive fixations	0.17 (0.23)	0.19 (0.28)	0.25 (0.38)	0.23 (0.37)
Introduction	Total fixations	10.14 <sup>a,c</sup>	7.84 <sup>a,c</sup>	8.76 <sup>a,c</sup>	8.56 <sup>a,c</sup>
	Total fixation	(4.19)	(3.33)	(3.36)	(3.21)
Introduction	Gaze durations	2.32	2.79	2.65	2.43
	Gaze duration	(1.20)	(3.13)	(1.78)	(2.19)
Introduction	Regressive fixations	7.83 <sup>a,c</sup>	5.05 <sup>a,c</sup>	6.11 <sup>a,c</sup>	6.12 <sup>a,c</sup>
	Regressive fixation	(3.65)	(3.65)	(2.66)	(3.09)

*Note.* Standard deviations are in parentheses.

<sup>a</sup>Denotes significant difference between texts (i.e., jokes vs. proverbs) at  $p < .01$ . <sup>b</sup>Denotes significant difference among contexts (i.e., comedian vs. politician) at  $p < .05$ . <sup>c</sup>Denotes significant interaction between texts and contexts at  $p < .05$ .

TABLE 4  
Mean Fixation Times (in Seconds) and Standard Deviations on Cotext and Final Line Areas of Interest for Experiment 2

Areas of Interest	Dependent Measures	Comedian		Politician		Control	
		Joke	Proverb	Joke	Proverb	Joke	Proverb
Cotext	Total fixations	13.30	13.51	13.31	13.50	12.72	13.09
		(2.85)	(3.08)	(3.08)	(2.95)	(2.93)	(3.48)
Cotext	Gaze durations	3.29*	3.77*	3.07*	3.99*	4.29*	4.63*
		(2.30)	(3.48)	(2.42)	(3.59)	(2.47)	(3.67)
Cotext	Regressive fixations	10.00*	9.74*	10.24*	9.52*	8.43*	8.46*
		(3.98)	(3.63)	(3.81)	(3.44)	(3.72)	(4.33)
Final line	Total fixations	2.50**	2.71**	2.45**	2.89**	2.33**	2.62**
		(0.70)	(0.71)	(0.71)	(0.93)	(0.69)	(0.97)
Final line	Gaze durations	1.36	1.41	1.25	1.37	1.30	1.23
		(0.78)	(0.61)	(0.61)	(0.63)	(0.73)	(0.62)
Final line	Regressive fixations	1.14**	1.60**	1.20**	1.51**	1.03**	1.38**
		(0.75)	(0.73)	(0.82)	(1.00)	(0.71)	(1.03)

Note. Standard deviations are in parentheses.

\*Denotes significant difference between texts (i.e., jokes vs. proverbs) at  $p < .05$ . \*\*Denotes significant difference between texts (i.e., jokes vs. proverbs) at  $p < .01$ .

narratives, as shown in Table 3:  $F_1(1, 29) = 41.28, p < .01$  ( $MSE = 5.78, d = 1.27$ ) and  $F_2(1, 23) = 7.88, p = .01$  ( $MSE = 5.82, d = 1.14$ ). As predicted by the validated-constraints hypothesis, participants spent extra time looking back to the framing narrative of a comedian if the comedian fulfilled the expectation of being funny by presenting a joke as compared to the framing narrative of a comedian who did not fulfill the humorous expectation and instead presented a nonhumorous proverb:  $F_1(1, 29) = 17.13, p < .01$  ( $MSE = 3.41, d = 1.44$ ). Participants' regressive fixation times back to the political framing narratives did not reliably differ when a joke versus nonhumorous proverb was given.

TABLE 5  
Mean Interest Ratings and Recall Scores for Experiment 2

Dependent Variable	Comedian		Politician		Control	
	Joke	Proverb	Joke	Proverb	Joke	Proverb
Interest ratings	4.29* (1.06)	3.74* (1.13)	4.42* (1.26)	3.42* (1.46)	4.61* (0.91)	3.44* (1.35)
Cloze recall	0.73 (0.29)	0.70 (0.28)	0.70 (0.32)	0.68 (0.31)	0.66 (0.28)	0.39 (0.31)
Final line recall one	2.77* (1.56)	1.01* (1.48)	2.70* (1.67)	1.31* (1.75)	2.99* (1.54)	1.03* (1.39)
Final line recall two	3.00* (1.44)	1.13* (1.47)	2.91* (1.65)	1.31* (1.75)	3.22* (1.40)	1.06* (1.40)

Note. Standard deviations are in parentheses.

\*Denotes significant difference between texts (i.e., jokes vs. proverbs) at  $p < .01$ .

Results from the total fixation times on the framing narratives further support this finding. Again, the interaction between framing narratives and text type was significant for the total fixation times on the framing narratives, as shown in Table 3:  $F_1(1, 29) = 14.45, p < .01$  ( $MSE = 2.29, d = 0.77$ ) and  $F_2(1, 23) = 5.52, p < .01$  ( $MSE = 3.90, d = 0.64$ ). Participants fixated on a comedian's framing narrative significantly more when s/he presented a joke rather than a nonhumorous proverb:  $F_1(1, 29) = 69.3, p < .01$  ( $MSE = 2.29, d = 0.61$ ). In contrast, participants did not spend reliably different amounts of total fixation time on political framing narratives when a joke or a nonhumorous proverb was presented.

The validated-constraints hypothesis also predicted that participants would have significantly increased regression fixation times to the cotexts of jokes in a comedy narrative framing. This prediction was (partly) supported by the eye-movement data. Participants had reliably different regressive fixation times to the cotext based on the different framing narratives,  $F_1(2, 58) = 3.39, p < .05$  ( $MSE = 10.53, d = 0.48$ ) and  $F_2(2, 46) = 1.57, p > .05$  ( $MSE = 13.71, d = 0.37$ ), as shown in Table 4 with means of 9.87 s, 9.88 s, and 8.45 s in the comedian, politician, and control conditions, respectively. Although participants had reliably longer regressive fixation times to the cotexts of jokes given by comedians than those given in a control narrative,  $F_1(1, 58) = 19.41, p < .01$  ( $MSE = 7.62, d = 0.41$ ), participants also had reliably longer regressive fixation times to the cotexts of jokes presented in political narratives than those presented in the control condition,  $F_1(1, 58) = 12.91, p < .01$  ( $MSE = 7.62, d = 0.48$ ). Participants did not reliably differ in the amount of time they looked back to the cotexts of jokes presented in comedy and political narratives. In all, the constraint that comedians should produce jokes did not affect regressive fixation times to the cotext as much as it affected the regressive fixation times to the framing narratives.

Participants' gaze durations on the cotext region of interest significantly differed (see row 3 of Table 4) based on the different framing narratives,  $F_1(2, 58) = 6.27, p < .05$  ( $MSE = 4.15, d = 0.66$ ) and  $F_2(2, 46) = 3.55, p < .05$  ( $MSE = 3.77, d = 0.56$ ), with means of 3.53 s, 3.53 s, and 4.46 s in the comedy, political, and control conditions, respectively. Pairwise comparisons revealed that participants had reliably longer fixation times on the cotexts presented in the control condition than in the comedy,  $F_1(1, 58) = 5.33, p < .01$  ( $MSE = 4.87, d = 0.31$ ); and the political conditions,  $F_1(1, 58) = 5.33, p < .01$  ( $MSE = 4.87, d = 0.31$ ). An explanation based on external constraints of the framing narratives is proposed to account for the longer gaze durations on the control condition cotexts. Gaze durations represent the first reading of the cotext before moving on to the final line. The control condition did not have any specific constraints about what type of information would be presented, whereas in the comedy and political narratives, participants were oriented to consider a particular type

of domain. In the control condition, the reader had to construct any pragmatic interpretation entirely from the text alone because none was provided. The size and placement of the texts in the control condition were the same as the other two conditions. Therefore, we suggest the extra cognitive effort required for constructing the pragmatic model for these control texts caused the inflated gaze durations.

Predictions associated with the violated-constraints hypothesis were not supported. According to this hypothesis, when a comedian tells something that is not funny, standards of comprehension might be violated (Baker, 1984), causing participants to regress more to the framing narrative of the comedian. However, participants did not look back to the framing narratives of a comedian when something nonhumorous was presented. In fact, participants virtually ignored the violation.

An interesting finding in these measures of online processing was that jokes, and humorous content in general, attracted the attention of participants significantly more than nonhumorous content. The additional fixations on “humorous elements” provide evidence for how the constraints of humor affect processing. For example, participants fixated significantly longer on the framing narratives of speakers in jokes than in nonhumorous proverbs, with means of 9.45 and 8.20, respectively:  $F_1(1, 29) = 7.65, p = .01$  ( $MSE = 79.75, d = 0.35$ ) and  $F_2(1, 23) = 5.131, p < .05$  ( $MSE = 13.43, d = 0.23$ ; see the fourth row of Table 3). Participants also seemed to be searching for humor in the final lines of texts that were not funny. They had significantly longer total fixation times on the adages than the punchlines with means of 2.74 and 2.43, respectively:  $F_1(1, 29) = 9.14, p = .01$  ( $MSE = 0.47, d = 0.39$ ) and  $F_2(1, 23) = 25.04, p < .01$  ( $MSE = 0.91, d = 0.54$ ; see the seventh row of Table 4). However, the first pass fixation times between the punchlines and adages did not reliably differ. Participants must have reread the adages because they were searching for any humor they may have missed during their initial reading. The adages were comparable to the punchlines in style and total number of words. Therefore, inflated fixation times for the adages must have been because of their lack of humor.

### *Interest Rating Data*

As predicted by the de-contextualized position, participants rated the jokes as significantly more interesting than their corresponding nonhumorous proverbs across all conditions, as shown in Table 5:  $F_1(1, 29) = 22.17, p < .01$  ( $MSE = 1.75, d = 0.71$ ) and  $F_2(1, 23) = 8.67, p < .05$  ( $MSE = 2.63, d = 0.72$ ). The impact of framing narratives was not significant, and the interaction between framing narratives and text was not significant. Apparently, a joke is interesting no matter who tells it.

### Recall Data

To analyze the recall data, each word left blank in the cloze procedure was scored on a dichotomous scale. The participant either remembered the correct keyword (or synonym) or did not, and the proportion of keywords filled in correctly was the measure of cloze recall. In contrast, the 4-point scale used in the cotext-cued recall of Experiment 1 was also used to grade both final line recalls.

Three measures of recall (i.e., cloze recall, final line recall one, and final line recall two) were analyzed (see Table 5). For the cloze recall, participants' recall did not reliably differ between texts or framing narratives. There also was not a reliable interaction between the two. However, participants did remember punchlines significantly better than the adages for final line recall one,  $F_1(1, 29) = 84.80, p < .01$  ( $MSE = 1.76, d = 1.09$ ) and  $F_2(1, 23) = 85.24, p < .01$  ( $MSE = 1.34, d = 0.97$ ); and final line recall two,  $F_1(1, 29) = 113.16, p < .01$  ( $MSE = 2.06, d = 1.23$ ) and  $F_2(1, 23) = 74.23, p < .01$  ( $MSE = 1.77, d = 0.89$ ). The framing narratives did not affect participants' recall for either final line. In essence, differences in participants' recall seemed based on only constraints of the texts themselves, as predicted by the de-contextualized position. This finding is congruent with previous research (Schmidt, 1994, 2002) that has shown humor facilitates memory. A curious finding in these data was the significant interaction between the final line recalls and text type:  $F_1(2, 28) = 6.59, p < .05$  ( $MSE = 0.15, d = 0.15$ ) and  $F_2(2, 22) = 1.84, p = .09$  ( $MSE = 0.17, d = 0.45$ ). Participants recalled the punchlines reliably better when all keywords of the joke cotexts were provided (i.e., recall two) than when they were missing (i.e., recall one) with means of 3.04 and 2.82, respectively. However, participants' recall for adages did not change between the two recalls.

Constraints associated with the literary and stylistic mechanisms in the setup of a joke, or a contextual finesse constraint, is offered as a possible explanation of why punchlines are better remembered than adages. Cotextual finesse describes the intricate combination of elements that are cleverly crafted, beginning with the first word of a joke and continuing until the punchline. The precise wording of the cotext of jokes appears crucial for comprehension and later recall of punchlines. The cotextual finesse constraint only applies to jokes, however, as recall of the adages was not significantly affected by the presence or absence of the keywords and ideas in the cotext.

Constraints associated with the GTVH (Attardo & Raskin, 1991) are similar to the idea of a cotextual finesse constraint. According to the GTVH, there are several key parameters of jokes (e.g., narrative strategy, target of the joke, or situation described in the joke) that all work toward the crucial punchline. Consequently, a mistake in any of the parameters can cause the punchline to fail and not be considered humorous. Further evidence for the intricate construction of jokes and other humorous texts is provided elsewhere (see Attardo, 2001).

Memory for punchlines significantly improved when the complete setup of the joke was given. Consider the following joke in which keywords of the cotext are underlined:

I attended a wedding reception, where a priest and a rabbi met at the buffet table. "Go ahead," said the priest, "try one of these delicious ham sandwiches. Overlook your divine rule just this once; it won't do you any harm." "That I will do, dear sir," the rabbi replied, "on the day of your wedding!"

Memory for the punchlines appears to have been facilitated because the keywords (in addition to other deeper constraints) were preparing the reader for the punchline rather than the adage. The final line of this nonhumorous proverb was, "That I can't do, dear sir," the rabbi replied, "God helps those who help themselves." The keywords and ideas in the cotext do not prepare the reader for this culminating adage.

This explanation assumes that cotext plays an important role in both setting up and recalling a joke. The punchline is not the only memorable part of a joke. Instead, the interaction between the constraints of the cotext and the punchline seem responsible for a joke's enhanced memory. This interaction is similar to Hunt and McDaniel's (1993) relational specific memory: Memory is influenced by the organization of groups of sentences within a text, not just the properties of individual sentences.

## GENERAL DISCUSSION

In two experiments, participants provided evidence for a constraint-based model of processing verbal jokes. The experiments show how a number of text constraints and the interaction between constraints of framing narratives and texts had systematic influences on the processing as well as the results of processing verbal jokes. Verbal jokes appear to be encoded, recalled, and interpreted according to a model of specific constraints. The interaction between constraints of the framing narratives and the texts appeared to have a significant impact on the online *processing* of jokes. In contrast, constraints of the texts themselves had the greatest effect on the *results* of processing the texts, as in the case of recall, humor ratings, and interest ratings.

In Experiments 1 and 2, the majority of readers' offline measures of processing best supported predictions based on the decontextualized incongruity resolution position. Jokes were more humorous, interesting, and were recalled better than the nonhumorous proverbs. In Experiment 2, however, participants' eye-movement data best supported predictions based on the validated-constraints hypothesis. Validated constraints were expected when the external consistency

standard (Baker, 1984) of a framing narrative is satisfied, such as when a comedian presents a joke. As predicted, participants' eye movements significantly differed in the condition when humor was expected rather than unexpected. For example, participants looked back to the framing narratives significantly more when comedians presented jokes than all other conditions. In contrast, participants relatively ignored the comedy framing narrative if a joke was not presented.

The results of these experiments have theoretical implications for how conceptually driven and data-driven processing are both used in the processing of verbal jokes. Initially, verbal jokes appear to be read differently based on different constraints associated with different framing narratives. This finding is compatible with the findings that (a) proverbial statements are initially read differently depending on the type of context in which they are presented (Katz & Ferretti, 2001, 2003) and (b) participants process figurative language differently based on different speaker intentions (Pexman et al., 2000). As long as the constraints associated with the framing narratives were met, then the constraints related to the texts were most important. Consequently, data-driven processing guided by the features and elements in the text were evident in later processing of the texts.

The results also suggest a humor facilitation effect. Jokes were not only funnier, more interesting, and recalled better: There was also evidence that the context, or setup of a joke, is constrained such that the precise wording of the joke is necessary to best facilitate participants' recall. In addition, participants spent significantly more time fixating on the adages of the proverbs than the punchlines of jokes. During this additional processing time, participants were likely searching for any humor they may have missed in the nonhumorous proverbs. Future research should empirically investigate potential applications of this humor facilitation effect. For example, could humor be used to facilitate the effect of textbooks on readers' comprehension (Mitchell, 2007)?

Future research should also investigate the effect of context on jokes in a more ecologically valid way. For the exploratory purposes of this research, our framing narratives used to contextualize the texts were extremely limited and were only a crude approximation of authentic context. Nonetheless, this operational definition provided a feasible first step in investigating this phenomenon. Normative constraints associated with laughter in particular environments have been noted elsewhere (Glenn, 2003). Thus, investigating the effect of context constraints in more natural environments, such as when jokes are told between friends versus professional acquaintances, would be of interest.

The possible combinations of context that surround the occurrence of humor or the telling of jokes are infinite. Imagine the difference between the sights and sounds of a comedian performing in a nightclub versus the sensory experiences associated with watching a comedian's performance via DVD in the comfort of

a home. To investigate a full-fledged context of humor, one may wish to hire comedians and have them deliver jokes and compare that to local politicians who deliver jokes. Future work could very well address such authentic context situations. A better understanding of how contextual factors influence all types of humor is certainly needed to obtain a better understanding of humor, in general. The constraint-based model of verbal jokes appears to only scratch the surface when considering the impact that context has on the processing of humorous texts. We have no intention in claiming that this model will explain everything about processing verbal jokes. Rather, as we systematically investigate and understand the components involved in processing verbal jokes, we will be better able to understand the constraints of these jokes, the constraints of contexts in which the jokes are told, and the interactions between the two.

### ACKNOWLEDGMENTS

This research was funded by the Office of Naval Research (ONR; Grant No. N00014-98-1-0331). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of ONR. We thank Richard Van Eck for his guidance on this project and Bethany McDaniel for her help in collecting and analyzing data.

### REFERENCES

- Asch, S. E. (1952). *Social psychology*. New York: Prentice Hall.
- Attardo, S. (1994). *Linguistic theories of humor*. Berlin: de Gruyter.
- Attardo, S. (2001). *Humorous texts: A semantic and pragmatic analysis*. Berlin: de Gruyter.
- Attardo, S., & Raskin, V. (1991). Script theory revisited: Joke similarity and joke representation model. *Humor, 4*, 293–347.
- Baker, L. (1984). Spontaneous versus instructed use of multiple standards for evaluating comprehension: Effects of age, reading proficiency, and type of standard. *Journal of Experimental Child Psychology, 38*, 289–311.
- Berg, E. M., & Lippman, L. G. (2001). Does humor in radio advertising affect recognition of novel product brand names? *The Journal of General Psychology, 128*, 194–205.
- Carroll, P. J., Young, J. R., & Guertin, M. S. (1992). Visual analysis of cartoons: A view from the far side. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 444–461). New York: Springer-Verlag.
- DeFord, D., Goodman, P., & LaFlaur, M. (1999). *Reader's Digest: Laughter, the best medicine*. Pleasantville, NY: The Reader's Digest Association, Inc.
- Duncan, C. P., Nelson, J. E., & Frontzak, N. L. (1984). The effect of humor on advertising comprehension In T. C. Kinnear (Ed.), *Advances in consumer research* (pp. 432–437). Chicago: Association for Consumer Research.
- Duranti, A., & Goodwin, C. (Eds.). (1992). *Rethinking context*. Cambridge, England: Cambridge University Press.



- Fitneva, S. A., & Spivey, M. J. (2004). Context and language processing: The effect of authorship. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 317–328). Cambridge, MA: MIT Press.
- Gelb, B. D., & Zinkhan, G. M. (1986). Humor and advertising effectiveness after repeated exposures to a radio commercial. *Journal of Cognitive Psychotherapy*, *10*, 235–254.
- Gernsbacher, M. A., & Robertson, R. (2002). The definite article *the* as a cue to map thematic information. In M. Louwerse & W. van Peer (Eds.), *Thematics: Interdisciplinary studies* (pp. 119–136). Amsterdam: Benjamins.
- Glenn, P. (2003). *Laughter in interaction*. Cambridge, MA: Cambridge University Press.
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. New York: Harper & Row.
- Graesser, A. C., Lu, S., Olde, B. A., Cooper-Pye, E., & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory & Cognition*, *33*, 1235–1247.
- Grimshaw, A. D. (2003). Genres, registers, and contexts of discourse. In A. C. Graesser, M. A. Gernsbacher, & S. R. Golman (Eds.), *Handbook of discourse processes* (pp. 25–81). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Gunter, B., Baluch, B., Duffy, L. J., & Furnham, A. (2002). Children's memory for television advertising: Effects of program-advertisement congruency. *Applied Cognitive Psychology*, *16*, 171–190.
- Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*, 717–742.
- Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and a distinctiveness. *Journal of Memory and Language*, *32*, 421–445.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354.
- Kaplan, R. M., & Pascoe, G. C. (1977). Humorous lectures and humorous examples: Some effects upon comprehension and retention. *Journal of Educational Psychology*, *69*, 61–65.
- Katz, A. N. (2005). Discourse and social-cultural factors in understanding nonliteral language. In H. Colston & A. Katz (Eds.), *Figurative language comprehension: Social and cultural influences* (pp. 183–207). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Katz, A. N., & Ferretti, T. R. (2001). Moment-by-moment comprehension of proverbs in discourse. *Metaphor and Symbol*, *16*, 193–221.
- Katz, A. N., & Ferretti, T. R. (2003). Reading proverbs in context: The role of explicit markers. *Discourse Processes*, *36*, 19–46.
- Katz, A. N., & Pexman, P. M. (1997). Interpreting figurative statements: Speaker occupation can change metaphor to irony. *Metaphor and Symbol*, *12*, 19–41.
- Kemper, S., & Thissen, D. (1981). How polite? A reply to Clark and Shunk. *Cognition*, *9*, 305–309.
- Kintsch, W., & Bates, E. (1977). Recognition memory for statements from a classroom lecture. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 150–159.
- Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 55–75). Oxford, England: Elsevier.
- MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (2nd ed., pp. 581–611). London: Elsevier.
- McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312.

- Mitchell, H. H. (2007). No laughing matter: The use of humor in texts. *Information Design Journal*, 15, 17–20.
- O'Brien, E. J., Raney, G. E., Albrecht, J. E., & Rayner, K. (1997). Processes involved in the resolution of anaphors. *Discourse Processes*, 23, 1–24.
- Pexman, P. M., Ferretti, T. R., & Katz, A. N. (2000). Discourse factors that influence online reading of metaphor and irony. *Discourse Processes*, 29, 201–222.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Rayner, K., Rotello, C., Stewart, A. J., Keir, S., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied*, 7, 219–226.
- Schmidt, S. R. (1994). Effects of humor on sentence memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 953–967.
- Schmidt, S. R. (2002). The humour effect: Differential processing and privileged retrieval. *Memory*, 10, 127–138.
- Schmidt, S. R., & Williams, A. R. (2001). Memory for humorous cartoons. *Memory & Cognition*, 29, 305–311.
- Schultz, T. R. (1976). A cognitive–developmental analysis of humor. In A. J. Chapman & H. C. Foot (Eds.), *Humor and laughter: Theory, research and applications* (pp. 11–36). New York: Wiley.
- Suls, J. M. (1972). A two-stage model for the appreciation of jokes and cartoons: An information-processing analysis. In J. H. Goldstein & P. E. McGhee (Eds.), *The psychology of humor* (pp. 81–100). New York: Academic.
- Suls, J. M. (1977). Cognitive and disparagement theories of humor: A theoretical and empirical synthesis. In A. J. Chapman & H. C. Foot (Eds.), *It's a funny thing, humor* (pp. 41–45). Oxford, England: Pergamon.
- Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), *Handbook of perception and cognition*, Vol. 11: *Speech, language and communication*, (pp. 217–262). New York: Academic.
- Vaid, J., Hull, R., Heredia, R., Gerkens, D., & Martinez, F. (2003). Getting a joke: The time course of meaning activation in verbal humor. *Journal of Pragmatics*, 35, 1431–1449.