Memory & Cognition
2001, 29 (1), 117-126

The relation of tip-of-the-tongue states and retrieval time

BENNETT L. SCHWARTZ
Florida International University, Miami, Florida

The tip-of-the-tongue state (TOT) is the phenomenological experience that a word is on the verge of being recalled. Participants rated TOTs as either emotional or nonemotional. In Experiment 1, given general-information questions, participants spent more time attempting retrieval during emotional TOTs than during nonemotional TOTs or n-TOTs (retrieval failures not accompanied by TOTs). Experiment 2 replicated the effect that TOTs show longer retrieval times than n-TOTs. In Experiment 3, with word definitions as stimuli, retrieval times were longer for emotional TOTs. Experiment 4 showed the same relation between retrieval times and TOTs even when participants made retrospective decisions about whether they had experienced a TOT before they retrieved the correct target. Valence of emotion was correlated with correct resolution of the TOT. These results are discussed in the context of a metacognitive model, in which TOTs serve to monitor and control cognition.

Most people report experiencing a tip-of-the-tongue state (TOT) on a common basis (A. S. Brown, 1991; Reason & Lucas, 1984). To the rememberer, a TOT is a *strong feeling* that a particular target word can be retrieved, and the feeling that the retrieval of that target is imminent. All studies to date have shown that TOTs are correlated with objective indices of memory such as retrieval of the first letter, syllable, grammatical gender, synonyms, homonyms, or semantically related information (Brennan, Baguley, Bright, & Bruce, 1990; A. S. Brown, 1991; R. Brown & McNeill, 1966; Koriat & Lieblich, 1974; Miozzo & Caramazza, 1997; Perfect & Hanley, 1992; Vigliocco, Antonini, & Garrett, 1997). However, researchers have attempted in only a few studies to relate TOTs to behaviors that people engage in to control retrieval (see Gardiner, Craik, & Bleasdale, 1973; Ryan, Petty, & Wenzlaff, 1982, for exceptions). The present study was designed to address whether TOTs serve to control aspects of retrieval.

Metacognitive control refers to when rememberers use the output of their monitoring systems to regulate or change cognitive processes or behavior (e.g., Mazzoni, Cornoldi, & Marchitelli, 1993; Nelson, 1996; Nelson & Narens, 1990, 1994). Thus, metacognitive control is a form of self-regulation, because the person uses feedback from one system to affect change in another (Nelson, 1996). Examples of control phenomena abound in the metacognitive literature. For example, Mazzoni et al. found a negative correlation between judgments of learning and study time, suggesting that participants devote more study time to items that they consider to be difficult. Thiede and Dunlosky (1999) showed that judgments of learning also predicted the selection of items for restudy. Koriat and Goldsmith (1996) found that with high incentives for accuracy, participants could successfully screen out answers as incorrect that they had earlier made under forced-report conditions.

The relation of retrieval times to feeling of knowing has been examined in several studies. Nelson, Gerler, and Narens (1984), using feeling-of-knowing judgments, not TOTs, found that retrieval latencies were longer for unrecalled targets given higher feeling-of-knowing judgments. Costermans, Lories, and Ansay (1992) replicated this finding, finding longer retrieval latencies following higher feeling-of-knowing judgments. They also found that high confidence in correct answers was related to quicker retrieval latencies. Barnes, Nelson, Dunlosky, Mazzoni, and Narens (1999) argued that one of the most important control functions of metacognition is the regulation of search time during retrieval. They too found that longer retrieval latencies were associated with higher feeling of knowing. These studies suggest that rememberers are capable of using their metacognitive judgments to alter behavior, which in turn can strengthen memory representations. However, because of differences between TOTs and feelings of knowing, it is not clear that these results will necessarily generalize to TOTs (see Schwartz, 1999).

Nelson and Narens (1994) argued that one of the major functions of any metacognitive system is to allow effective control of ongoing cognition. Nelson and Narens (1994) employed an analogy comparing metacognition to the working of a thermostat. A thermometer measures the ambient temperature of the room, which serves the monitoring function. When the temperature rises above or falls

The author thanks Alan S. Brown, Leslie Frazier, Peter Graf, Thomas O. Nelson, and one anonymous reviewer for commentary on preliminary drafts of this paper. The author also thanks Yenys Castillo Andrade, Jean Richmond, Michelle Vigoa, and Daniel Yero for assistance with data collection. Correspondence should be addressed to B. L. Schwartz, Department of Psychology, Florida International University, Miami, FL 33199 (e-mail: schwartb@fiu.edu).

below a threshold temperature, the thermostat mechanism turns on the heat or cooling system, which serves as the control function. By analogy, metacognitive judgments, such as judgments of learning or feelings of knowing, can serve the purpose of monitoring the cognitive system. However, it does not logically follow that monitoring can influence the control system. For example, consider a freestanding thermometer in the same room—all it can do is monitor. Thus, it is important to empirically demonstrate that metacognition serves both control and monitoring functions. Although research has shown that feelings of knowing are related to retrieval decisions (e.g., Costermans et al., 1992; Nelson et al., 1984), only limited research has suggested that TOTs control retrieval decisions. The purpose of this paper is to link TOTs' monitoring function to a potential control function.

Two studies on TOTs do shed some light on this topic. Ryan et al. (1982) induced people into TOT states by asking them to retrieve targets of recently learned paired associates. Then, while participants were trying to retrieve the target words for initially unretrieved items, they were asked to perform a secondary task. Ryan et al. found that people were slower at doing the secondary task when they were experiencing a TOT for a target word than when they were not. They argued that participants were still devoting retrieval time to the TOT target even though they were supposed to be concentrating on the secondary task. The study suggests that TOT experiences may be correlated with longer retrieval efforts. In an earlier study, Gardiner et al. (1973) found that participants were better at retrieving target words for which they had earlier experienced a TOT. These studies, along with the Nelson et al. (1984) and Costermans et al. (1992) findings of a positive relation between feeling of knowing and retrieval latency, suggest that TOTs invite further processing or retrieval attempts.

The present study addresses the question of what the functional role of a TOT experience might be. The TOT serves a monitoring function by informing us that the word is known and may be retrieved sooner or later. Thus, it is likely that the TOT also serves the control function of directing our retrieval efforts. However frustrating a TOT may be, rememberers ought to be more likely to attempt additional retrieval of TOT words than of words that do not elicit TOTs. In the experiments described here, it is predicted that participants will spend more time attempting retrieval of unrecalled words when in a TOT than when they are not in a TOT (n-TOT). A caveat is that the present studies were designed only to test a correlation between TOTs and retrieval time because no variable affecting retrieval times was introduced.

TOT substates

In a recent study, Schwartz, Travis, Castro, and Smith (2000) found that rememberers divided TOTs into substates based on phenomenological characteristics, and that these characteristics were predictive of performance. Schwartz et al. found that rememberers could divide TOTs into strong and weak TOTs, emotional and nonemotional TOTs, and imminent and nonimminent TOTs. Each of these phenomenological variables predicted different aspects of retrieval. Imminence was a particularly good predictor of resolution (subsequent recall), whereas emotionality was particularly good at predicting recognition in an eight-alternative forced choice test. Schwartz et al. concluded that the phenomenological variables, while overlapping, tapped into different phenomenological substates of the rememberer.

Because these phenomenological substates appear to have predictive monitoring functions, it is likely that TOT substates may control retrieval behavior in a differential manner. In the present study, I asked each participant to decide whether each TOT was emotional or nonemotional. I chose the emotionality variable because Schwartz et al. (2000) found that it was a good predictor of target recognition (hence, target knowledge), but was not a good predictor of resolution—that is, the recall of initially unretrieved items. Therefore, it seemed to represent a more conservative test of the hypothesis that TOTs would positively correlate with retrieval time. If emotional TOTs show longer retrieval times than n-TOTs, it suggests that an important function of a TOT is to increase the time a person spends engaged in retrieving the target. This should occur regardless of whether the target is actually retrievable or not. Moreover, the present study will serve as a replication of the findings concerning emotionality and TOTs.

Four experiments were conducted in this study to address the relation between retrieval time and TOTs. In Experiments 1, 2, and 4, TOTs were induced with general-information questions, whereas in Experiment 3, TOTs were induced with word definitions. The set of general-information questions used in the present study has proven to be a successful generator of TOTs (Schwartz, 1998; Schwartz et al., 2000). The set of word definitions has also been used in prior TOT research (Harley & Bown, 1998). In past research, general-information questions have been used to address issues of interest to metacognition research (e.g., Schwartz, 1998), whereas word definitions are more likely to be used in TOT research interested in psycholinguistic issues of lexical retrieval (e.g., Harley & Bown, 1998; Vigliocco et al., 1997). Using two different kinds of stimuli to demonstrate the effect should broaden the potential application of the findings. In fact, as the experiments demonstrate, the studies show some interesting similarities and some subtle differences.

EXPERIMENT 1

The chief hypothesis of Experiment 1 was that participants will take more time before terminating a retrieval attempt when they are experiencing a TOT for an unrecalled item than when they are not experiencing a TOT for an unrecalled item. Therefore, in Experiment 1, retrieval times were measured and then compared with the likelihood of experiencing a TOT. Furthermore, this study also

looked at whether this relation is modified by the kind of TOT the participant is experiencing. Thus, each participant was asked whether the TOT was accompanied by emotion or not.

Method

Participants. The participants were 80 Florida International University students who received partial course credit for their participation. Each participant was tested individually on a Macintosh computer during a session that lasted approximately 1 h.

Materials. The stimuli for the experiment were 80 general-information questions taken from the Nelson–Narens norms (Nelson & Narens, 1980). For example, one question was "For which country is the rupee the monetary unit? (India)" Previous testing indicated that the 80 questions generated a correct percent rate of around 36% in the student population tested (Schwartz, 1998; Schwartz et al., 2000). The list of items was randomized for each participant.

Procedure. Participants were first given detailed instructions about the procedure. They were told that they would be answering a series of general-information questions, some of which would be easy and some of which would be more difficult. They were given an explanation of what the term *tip of the tongue* meant. All participants reported being familiar with the experience and with the term. The instructions were as follows,

> If you do not answer the question correctly or leave the answer blank, you will be asked whether or not you are in a tip-of-the-tongue state for the target answer. A tip-of-the-tongue state (abbreviated TOT) means that you feel as if it is possible that you could recall the target answer, and that you feel as if its recall is imminent. Sometimes, you may feel frustrated or emotional that you cannot recall a word that you are sure you know. If you are in a TOT state that is accompanied by this emotional content, indicate a TOT with emotion by pressing the E key. If you are in a TOT state, but do not feel an emotional state, press the P key. If you cannot recall the answer and are not in a TOT state, press the N key.

This was explained first by the experimenter and was then repeated on the computer just before the experiment began. The participants were instructed only to indicate an emotional TOT if they were experiencing emotion as part of the TOT. The experimenter then started the computer program that ran the experiment. The participants were instructed to go through the questions at a pace that was comfortable for them. No instructions were given to go either fast or slow.

Each question appeared on the middle of the screen, and a prompt appeared beneath the question. The question remained on the screen until the participants typed in their responses. As soon as the question appeared on the screen, the program started the timer. The retrieval time was measured as the amount of time that transpired between the onset of the questions and the initiation of the participant's response (i.e., typing of the first letter). The participants typed in their responses, or they indicated that they did not know by typing in a question mark. If the participants typed in the correct response, they simply moved on to the next question. If they indicated that they did not know (omission error) or answered incorrectly (commission error) (i. e., Krinsky & Nelson, 1985), they were asked whether or not they were in a TOT. The participants were told in advance that if they were asked about a TOT after they had typed in a response, the response was either incorrect or misspelled. The instructions were to report a TOT only if one was experiencing one. Participants typed in "E" when they were in an emotional TOT, "P" when they were in a nonemotional TOT, and "N" when they were not experiencing a TOT.

After the participants had attempted retrieval for all 80 questions and made TOT judgments for those that they did not recall correctly, they were given more instructions. The participants were given a final recognition test for the answerable questions. They were again shown the question followed by eight alternatives, one of which was the correct answer (Wilkinson & Nelson, 1984). A number accompanied each alternative. The participants typed in the number associated with the answer that they thought was correct. They were then presented with the next question. This continued for all the originally missed questions. At the end of the session, the participants were fully debriefed and were given credit in their introductory psychology course.

Results

Procedures for analysis. Statistical reliability was measured at $p < .05$ in all experiments, as adjusted by Greenhouse–Geisser epsilon, in all the analyses of variance (ANOVAs) discussed here (Stevens, 1986). ANOVAs were used throughout, even though the main variables were not true independent variables. Use of ANOVA allowed the examination interactions; however, causality cannot be inferred from any of these analyses. The degrees of freedom vary across the analyses because some participants did not make commission errors whereas others did not make omission errors. Generally, TOT judgments are made only after errors of omission—those items for which participants indicate that they do not know. In the present study, TOT judgments were also elicited after errors of commission, but for the most part, they were analyzed separately from the TOT judgments made after errors of omission. To further examine data when a significant omnibus F was found, Newman–Keuls post hoc tests (Winer, 1971, p. 528) were used.

In all of the analyses on retrieval times, analyses were done on both the means and the medians. The first question was always treated as practice and was not included in any of the analyses. Outliers for each participant were excluded from the analyses. Outliers were considered to be any retrieval time that was greater than three standard deviations from the mean of that participant. For example, one participant with a median retrieval time of 12.5 sec had a single retrieval time of over 360 sec! Although the retrieval time distributions were positively skewed, the analyses for means and medians showed identical patterns. For the sake of brevity, only the analyses of the means are reported, and only the means are included in the tables.

Recall and recognition. Correct recall at the first test was 38%, which is consistent with earlier findings with these stimuli (Schwartz, 1998; Schwartz et al., 2000). Correct recall was computed with a weak criterion for spelling. The computer program was set up to accept many misspellings. However, it could not catch all misspelling (e.g., "sixteen" for "Sistine"). These odd misspellings were removed postexperimentally by hand (they represented 9% of all retrieval attempts in this experiment). Misspelled items were included in the total percent correct. Of items not answered correctly, 60% were commission errors and 40% were omission errors. The ratio of commission errors to omission errors varied considerably from experiment to experiment, possibly be-

Table 1
Mean Proportion of Unrecalled Items (Standard Errors in Parentheses) in TOT or n-TOTs for Experiments 1–3

	Experiment 1		
	Emotional	Nonemotional	n-TOT
Commission	.32 (.03)	.20 (.02)	.48 (.03)
Omission	.24 (.02)	.14 (.02)	.62 (.03)
	Experiment 2		
	TOT		n-TOT
Commission	.26 (.02)		.74 (.02)
Omission	.48 (.03)		.52 (.03)
	Experiment 3		
	Emotional	Nonemotional	n-TOT
Commission	.42 (.05)	.19 (.03)	.39 (.05)
Omission	.33 (.05)	.28 (.05)	.39 (.05)

	Experiment 4				
	Overall	Both	Frustrating	Exciting	No Emotion
Commission	.41 (.04)	.13 (.04)	.34 (.05)	.21 (.04)	.32 (.05)
Omission	.37 (.02)	.10 (.02)	.59 (.04)	.07 (.02)	.22 (.03)
Retrospective	.22 (.03)	.16 (.04)	.17 (.03)	.34 (.05)	.32 (.05)

Note—For Experiment 4, TOTs were made on recalled items and unrecalled items.

cause of slight variants in instructions. Overall, recognition of initially unrecalled items was 39%, significantly higher than chance (12.5%) [$t(79) = 23.5$].

Emotional and nonemotional TOTs. Participants indicated that they experienced more emotional than nonemotional TOTs [$F(1,67) = 12.96, MS_e = 0.08$]. TOTs were more common after commission errors than after omission errors [$F(1,67) = 18.72, MS_e = 0.02$]. The interaction between the two was not significant ($F = 1.35$). The TOT rates are shown in Table 1. These rates are close to those observed in Experiment 2 of Schwartz et al. (2000), who also looked at emotional TOTs.

TOTs and retrieval time. The computer program measured the amount of time that participants spent before initiating the typing of their answers in the recall phase of the experiment. For an omission error, retrieval time was measured as the amount of time that it took participants to indicate a "don't know" response by typing in a "?" on the computer keyboard. Response times were a function of retrieval status—that is, of whether the target was correctly recalled, incorrectly recalled, given a "don't know" response, in an emotional TOT state or in a nonemotional TOT state. There was an overall main effect of response type [$F(2,134) = 3.62$, $MS_e = 18.93$]. Post hoc tests indicated that this effect was due to long response times for commission errors (9.5 sec) and omission errors (i.e., "don't know" responses) (8.8 sec) relative to correct responses (7.7 sec) (see Krinsky & Nelson, 1985, for a similar analysis).

Table 2 shows the mean retrieval times across conditions. Response times for omission errors of emotional TOTs were longer than response times for omission errors of nonemotional TOTs, which in turn showed longer response times than did n-TOTs following omission errors [$F(2,84) = 36.46$, $MS_e = 16.53$]. For commission errors, emotional TOTs also showed a longer response time than did nonemotional TOTs or n-TOTs [$F(2,136) = 7.70$, $MS_e = 40.23$]. Overall, correct responses were faster than incorrect responses [$F(1,79) = 17.60, MS_e = 2.53$]. Thus, the evidence suggests that TOTs and initial retrieval time are related, with longer response times occurring when rememberers are experiencing emotional TOTs.

Accuracy of TOTs for answerable questions. Table 3 shows the correct recognition percents as a function of TOTs and n-TOTs. The final recognition test was used as a measure of accuracy. If TOTs are accurate predictors of memory storage, recognition should be higher following a TOT than an n-TOT. Using the first TOT as the criterion TOT, recognition was higher after an emotional TOT than a nonemotional TOT, which was higher than an n-TOT [$F(2,138) = 96.7, MS_e = 0.03$]. A parallel method of examining accuracy of the TOTs is to look at the gamma correlation between the TOTs and the likelihood of subsequent recognition (see Nelson, 1984, for a justification of the gamma correlation). To compute the gamma correlations, emotional TOTs and nonemotional TOTs were both dummy coded as a 2 and n-TOTs as a 1. Emotional and nonemotional were coded equivalently because neither is a stronger subjective prediction of knowing. Correct recognition was dummy coded as a 2 and incorrect recognition as a 1. The TOT judgments showed a .61 gamma correlation with recognition, significantly higher than chance.

Summary and Discussion

As predicted, retrieval times were longer following TOTs than following n-TOTs. This effect was particularly large for the emotional TOTs. Thus, it is likely that

Table 2
Retrieval times in Seconds (Standard Errors in Parentheses) as a Function of Correct Retrieval, TOT Substate, and Type of Error

	Experiment 1			
	Correct	Emotional	Nonemotional	n-TOT
Omission	13.3 (1.07)	8.1 (.56)	7.3 (1.00)	
Commission	12.3 (1.17)	8.2 (.56)	8.8 (.33)	
Correct	7.7 (.28)			
	Experiment 2			
	Correct	TOT		n-TOT
Omission		13.8 (1.13)		7.6 (.42)
Commission		12.3 (.89)		11.0 (.72)
Correct	8.9 (.48)			
	Experiment 3			
	Correct	Emotional	Nonemotional	n-TOT
Omission		25.2 (2.60)	16.1 (1.77)	12.7 (1.11)
Commission		16.9 (1.26)	15.0 (1.70)	17.3 (1.53)
Correct	12.6 (.87)			
	Experiment 4			
	Correct	TOT		n-TOT
Omission		14.9 (.92)		7.7 (.34)
Commission		16.3 (1.03)		11.3 (.53)
Retrospective		16.0 (.84)		10.0 (.54)
Correct	11.1 (.46)			

Table 3
Proportion Recognized (Standard Errors in Parentheses) as a Function of TOT or n-TOT

	Experiment 1		
	Emotional TOTs	Nonemotional TOTs	n-TOTs
Recognition	.51 (.02)	.42 (.03)	.14 (.01)
	Experiment 2		
	TOTs		n-TOTs
Recognition	.55 (.03)		.31 (.02)
	Experiment 3		
	Emotional TOTs	Nonemotional TOTs	n-TOTs
Recognition	.69 (.04)	.61 (.04)	.30 (.03)
	Experiment 4		
	TOTs		n-TOTs
Recognition	.49 (.02)		.16 (.01)

being in a TOT influences retrieval behavior and that participants spend more time trying to retrieve the missing word. This was true, in this experiment, even for commission errors in which people make an answer, albeit an incorrect one. Moreover, TOTs accurately predicted performance in the final recognition test. Thus, the present study suggests that TOTs serve as accurate markers of monitoring and serve a control function as well, although only a correlation between TOTs and retrieval times has been demonstrated.

EXPERIMENT 2

In most experiments on TOTs, participants are simply asked to make dichotomous judgments as to whether they are experiencing a TOT or not. Only one other study (Schwartz et al., 2000) used the methodology of Experiment 1, and just a handful have used continuous judgments (e.g., Gardiner et al., 1973; Yaniv & Meyer, 1987). Thus, a potential concern about Experiment 1 is that the process by which rememberers decided whether a TOT was emotional or not, or perhaps even the process of finding the right keys to press, may have interfered with the normal process of retrieval and the normal induction of TOTs. Therefore, in Experiment 2, a more traditional TOT design was employed in which the participants merely had to decide whether or not they were experiencing a TOT.

Method

Participants. The participants were 50 Florida International University students who received partial course credit for their participation. Each participant was tested individually on a Macintosh computer during a session that lasted approximately 1 h.

Materials. The same materials were used in Experiment 2 as were used in Experiment 1.

Procedure. Participants were first given detailed instructions about the procedure. There were two differences in procedure between Experiment 1 and Experiment 2. First, there were no instructions about the differences between emotional and nonemotional TOTs. Second, participants were instructed to type a "Y" when they were experiencing a TOT and an "N" when they were not experiencing a TOT.

Results

Recall and recognition. Correct recall at the first test was 36%. Correct recall was computed with a weak criterion for spelling. Odd misspellings, representing 7% of all responses made, were removed postexperimentally by hand and were counted as correct responses. Of items not answered correctly, 48% were commission errors and 52% were omission errors. Overall, recognition of unrecalled targets was 39%, significantly higher than chance (12.5%) [$t(49) = 11.4$].

TOTs and retrieval time. Response time in this initial recall phase was tracked as a function of whether the response was correct, a commission error, or an omission error. There was an overall main effect of response type [$F(2,98) = 16.72$, $MS_e = 5.86$]. Post hoc tests indicated that this effect was due to long response times for commission errors (11.6 sec) relative to correct responses (8.9 sec) and omission errors responses (9.4 sec). Overall TOT rates are reported in Table 1.

Table 2 shows the mean retrieval times across conditions. Response times of omission errors followed by TOTs were longer than the response times of omission errors followed by n-TOTs [$F(1,47) = 48.8$, $MS_e = 18.93$]. For commission errors, the mean of TOTs did not differ from the means of commission errors not followed by TOTs ($F = 3.8$). Overall, correct responses were faster than incorrect responses [$F(1,49) = 8.0$, $MS_e = 3.8$]. The evidence suggests that TOTs and initial retrieval time are strongly correlated, with longer response times occurring when rememberers are experiencing TOTs, at least for those following omission errors.

Accuracy of TOTs. Table 3 shows the mean recognition performance across conditions. Recognition was higher after a TOT than after an n-TOT [$F(1,49) = 80.8$, $MS_e = 0.02$]. The TOT judgments showed a .44 gamma correlation with recognition, significantly higher than chance.

Summary and Discussion

As predicted, retrieval times were longer following TOTs as opposed to n-TOTs, at least for omission errors, even when participants were not asked to distinguish between emotional and nonemotional TOTs. Thus, it is unlikely that the additional judgment category (the distinction between emotional and nonemotional TOTs) interfered with the participants' ability to efficiently retrieve target answers or assess their metacognitive state.

It is worth noting that the TOTs in Experiment 2 resembled the emotional TOTs of Experiment 1 rather than the combination of both emotional and nonemotional TOTs. First, the rate of TOTs for omission errors in Experiment 2 was almost exactly the same as the rate of emotional TOTs in Experiment 1. Second, the retrieval times for the TOTs in Experiment 2 were more similar to the

retrieval times for the emotional TOTs than they were to the averaged rate of emotional and nonemotional TOTs in Experiment 1. Thus, the demand characteristics of Experiment 1 might have created a category (nonemotional TOTs) that participants might otherwise have judged to be n-TOTs. This issue was addressed further in Experiment 4.

EXPERIMENT 3

Experiments 1 and 2 support the hypothesis that TOTs serve a control function. In both experiments, TOTs were correlated with longer retrieval latencies than were n-TOTs. However, both Experiments 1 and 2 used the same kind of stimuli—general-information questions. Experiment 3 was an attempt to broaden the generalizability of the relation between TOTs and retrieval times. Therefore, in Experiment 3, word definitions were used as the stimuli.

Method

Participants. The participants were 45 Florida International University students who received partial course credit for their participation. Each participant was tested individually on a Macintosh computer during a session that lasted approximately 1 h.

Materials. The stimuli for the experiment were 40 word definitions adapted from Harley and Bown (1998). For example, one question was "The dense, soft, often curly hair of some animals, used to make yarn, cloth, and clothing" (wool). The list of items was randomized for each participant.

Procedure. Participants were first given detailed instructions about the procedure. They were told that they would be answering word definitions, some of which would be easy and some of which would be more difficult. Otherwise the instructions and procedures were identical to those in Experiment 1. The distractors in the final recognition test were semantic and were created by the author. There were only three distractors for each target word.

Results

Recall and recognition. Correct recall at the first test was 28%. Correct recall was computed with a weak criterion for spelling. Odd misspellings, representing 7% of all responses made, were removed postexperimentally by hand and were counted as correct responses. Of items not answered correctly, 76% were commission errors and 24% were omission errors. Overall, recognition of unrecalled targets was 67%, significantly higher than chance (25%) [$t(44) = 41.2$].

Emotional and nonemotional TOTs. Table 1 shows the distribution of TOTs across conditions. Participants experienced more emotional than nonemotional TOTs [$F(1,30) = 5.18$, $MS_e = 0.13$]. TOTs were as likely after omission errors as they were after commission errors ($F < 1$). However, there was a significant interaction [$F(1,30) = 5.04$, $MS_e = 0.05$]. The interaction is caused by the relatively high number of emotional TOTs for commission errors and the relatively high number of nonemotional TOTs for omission errors.

TOTs and retrieval time. Response time in this initial recall phase was tracked as a function of whether the response was correct, a commission error, or an omission error. There was an overall main effect of response type [$F(2,58) = 16.88$, $MS_e = 22.04$]. Post hoc tests indicated that this effect was due to long response times for omission errors (18.2 sec) relative to correct responses (11.3 sec) and commission errors (15.7 sec). Unlike in Experiment 1, here the omission errors took more time than did the commission errors.

Table 2 shows the mean retrieval times across conditions. Response times of omission errors followed by emotional TOTs were longer than the response times of omission errors followed by nonemotional TOTs, which did not significantly differ from omission errors followed by n-TOTs [$F(2,42) = 14.40$, $MS_e = 62.61$]. For commission errors, however, the mean of emotional TOTs did not differ from either the mean of nonemotional TOTs or the mean of commission errors not followed by TOTs ($F < 1$). Overall, correct responses were faster than incorrect responses [$F(1,43) = 38.19$, $MS_e = 11.14$]. The evidence suggests that TOTs and initial retrieval time are related, with longer response times occurring when rememberers are experiencing emotional TOTs, at least for those following omission errors.

Accuracy of TOTs. Table 3 shows mean recognition performance across conditions. Recognition was higher after an emotional TOT than after a nonemotional TOT, which was higher than that after an n-TOT [$F(2,74) = 27.68$, $MS_e = 0.06$]. The TOT judgments showed a .54 gamma correlation with recognition, significantly higher than chance.

Summary and Discussion

In Experiment 3, participants had longer retrieval latencies for target words when they were experiencing an emotional TOT than when they were not experiencing an emotional TOT or when they were not in a TOT. However, unlike in Experiment 1, but as in Experiment 2, this effect was observed only following omission errors. There were no differences in response times for commission errors for emotional TOTs, nonemotional TOTs, and n-TOTs. Another difference between Experiment 1 and Experiment 3 is that in Experiment 3, the retrieval times were generally longer. In fact, during an emotional TOT for omission errors, participants averaged 25.2 sec in retrieval time, almost twice as long as was observed in Experiment 1. It is also interesting to note that whereas Experiment 1 produced higher recall scores in the initial phase, Experiment 3 showed higher recognition scores in the final phase. It is possible that the word definitions, because of their low recallability but high recognizability, produced a generally more frustrating set of items, leading to the overall increase in response times. This speculation is supported by the observation that there were more emotional TOTs in Experiment 3 than in Experiment 1.

The main methodological difference between Experiment 3 and the other experiments was that word definitions were used as stimuli in Experiment 3 whereas in all of the other experiments reported here, general-information

questions were used. In addition to the generally longer retrieval latencies, there was one other difference that seemed to stem from the difference in stimuli. In Experiments 1 and 2 (and 4), retrieval times for commissions were longer than retrieval times for omission errors. In Experiment 3, the effect reversed, and retrieval times for omissions were significantly longer than with commission errors. This may have been, in part, due to the difficulty in recalling the stimuli in this experiment and also the high rate of commission errors. It may also have been partially due to an inference that there may be fewer competing targets for a word definition than for a general-information question. In any event, the pattern seems to have reflected the stimuli used.

EXPERIMENT 4

The data from the first three experiments demonstrate a relation between TOTs and retrieval time. In all three experiments, there was a strong difference in retrieval times for TOTs and n-TOTs, particularly for the emotional TOTs, at least for those following omission errors. However, the TOTs following commission errors were problematic because many participants might have thought that their incorrect answers were correct. Thus, it was not quite clear what the TOTs were for when they followed commission errors. Therefore, in Experiment 4, participants were asked to make retrospective TOTs following correct answers. These TOTs were called *retrospective* TOTs, because the judgments did not occur until after the correct answer had been retrieved.[1] After every correct retrieval, participants were asked whether they had temporarily experienced a TOT before they retrieved the target answer. Then, the presence of a TOT could be correlated with retrieval time of correct answers. Many people have reported that TOTs can be quite short lived and quickly followed by resolution (e.g., Burke, MacKay, Worthley, & Wade, 1991), but no study has examined retrospective TOTs.

Experiment 2 raised the possibility that only emotional TOTs were "true" TOTs, and that the nonemotional TOTs might be more likely to result from demand characteristics of the experiment and would not ordinarily be considered TOTs by the participants (see Widner, Smith, & Graziano, 1996, for a discussion of demand characteristics and TOTs). Thus, in Experiment 4, participants were first asked to decide whether they were experiencing TOTs or not, and only after they had made this decision were they asked about emotion. Analysis, therefore, could be conducted on dichotomous TOTs ("yes" for TOT, "no" for n-TOT) or emotionality in TOTs. Furthermore, emotionality was divided as a function of valence.[2] The participants could indicate either that their TOT was emotionally frustrating (negative) or that their TOT was emotionally exciting (positive). It was hypothesized that emotionally exciting TOTs might be more associated with the retrospective TOTs because the TOTs were on the verge of being resolved.

Thus, this experiment differed from the previous three in three substantial ways. First, after a correct response, the participants were asked to retrospectively judge whether they had experienced a TOT before they correctly retrieved a target. Second, the participants judged emotionality after they had indicated that they were in a TOT. Third, the participants were required to make a judgment about emotional valence for all TOTs.

Method

Participants. The participants were 49 Florida International University students who received partial course credit for their participation (most of these participants were from an Honors section of Introductory Psychology). One participant was not included in the analysis because she typed her answers in small letters when the instructions indicated that all answers were to be typed in capital letters. Each participant was tested individually on a Macintosh computer during a session that lasted approximately 1 h.

Materials. The same materials were used as in Experiment 1 and Experiment 2.

Procedure. Participants were first given detailed instructions about the procedure, including the distinction between emotional valence and TOTs and concerning making TOT judgments retrospectively following correct answers. The instructions concerning TOTs were as follows:

> If you do not answer the question correctly or leave the answer blank, you will be asked whether or not you are in a tip-of-the-tongue state for the target answer. A tip-of-the-tongue state (abbreviated TOT) means that you feel as if it is possible that you could recall the target answer, and that you feel as if its recall is imminent. If you are in a TOT state, indicate a TOT, by pressing the Y key. If you cannot recall the answer and are not in a TOT state, press the N key.
>
> If you did answer the question already, you may have momentarily experienced a TOT just in advance of retrieving the word. The TOT may have only lasted an instant, but may have been felt prior to the retrieval of the answer. If you experienced such a momentary TOT, press the Y key. If you did not experience a momentary TOT, press the N key.
>
> Press the return key for more instructions.

The instructions for emotional valence were as follows;

> If you do experience a TOT, it may be accompanied by an emotional response. Some TOTs may be accompanied by emotional frustration, because you know the answer, but the failure to retrieve it causes you to be frustrated. Some TOTs may be accompanied by emotional excitement because you sense that you are about to retrieve the answer at any second. Some TOTs may not be accompanied by any experience of emotion at all.
>
> Please indicate your emotional state following a TOT.
>
> F = frustration
>
> E = excitement
>
> B = both frustration and excitement
>
> T = no emotion

Results

Recall and recognition. Correct recall at the first test was 41%. Correct recall was computed with a weak criterion for spelling. Odd misspellings were removed post-experimentally by hand (5% of all responses) and counted as correct responses. Of items not answered correctly, 27% were commission errors and 73% were omission errors. Overall, recognition of unrecalled targets was 41%, significantly higher than chance (12.5%) [$t(47) = 25.6$].

Emotional and nonemotional TOTs. First, the TOTs for unretrieved items will be considered. For omission errors, TOTs were experienced for 37% of unrecalled targets. Of those TOTs, the majority were frustrating TOTs [$F(3,141) = 46.5, MS_e = 0.06$; see Table 1]. Post hoc tests showed that no-emotion TOTs outnumbered those that were exciting and those that were both exciting and frustrating. For commission errors, TOTs were experienced for 41% following incorrect targets. Of these TOTs, frustrating TOTs and no-emotion TOTs outnumbered those that were exciting and those that were both frustrating and exciting [$F(3,126) = 3.2$, $MS_e = 0.13$; see Table 1]. For the retrospective judgments, TOTs were experienced 22% of the time following correct responses. Of these TOTs, exciting TOTs and no-emotion TOTs outnumbered TOTs that were frustrating and those that were both frustrating and exciting [$F(3,132) = 3.9, MS_e = 0.11$].

TOTs and retrieval time. Response time in this initial recall phase was tracked as a function of whether the response was correct, a commission error, or an omission error. There was an overall main effect of response type [$F(2,94) = 45.9, MS_e = 5.1$]. Post hoc tests indicated that this effect was due to long response times for commission errors (14.6 sec) relative to correct responses (11.1 sec) and omission errors responses (10.1 sec).

Table 2 shows the mean retrieval times across conditions. Retrieval times of omission errors for TOTs were longer than the retrieval times of n-TOTs [$F(1,46) = 98.7$, $MS_e = 12.5$]. For commission errors, the mean retrieval time was longer for TOTs than for n-TOTs [$F(1,42) = 27.6, MS_e = 20.0$]. For retrospective TOTs, retrieval times were also longer following TOTs as opposed to n-TOTs [$F(1,44) = 60.2$, $MS_e = 13.6$].

Analyses were also conducted separately for emotionality of TOTs and its relation to initial retrieval time. For the TOTs following retrieval failure, omissions and commissions were combined in order to give sufficient power to the analyses. The analyses were also conducted separately on the retrospective TOTs. For both TOTs following retrieval failure and assessed retrospectively, there were no differences in response times among the varying emotional states (Fs = 1.0 and 1.3, respectively). However, for both types of TOTs, the retrieval times were numerically fastest in the no-emotion condition. However, because of missing cells, this could not be explored statistically.

Accuracy of TOTs. Table 3 shows mean recognition performance across conditions. Recognition was higher after a TOT than after an n-TOT [$F(1,47) = 258.0, MS_e = 0.01$]. Recognition was highest after a "both" judgment (.62), but this did not reach a level of statistical significance ($F = 1.89$). The TOT judgments showed a .62 gamma correlation with recognition, significantly higher than chance.

Summary and Discussion

As predicted, retrieval times were longer following TOTs than following n-TOTs, as in each of the previous three experiments. In this experiment, retrieval times were longer for TOTs than for n-TOTs, even for TOTs judged retrospectively after the target word had been correctly recalled. Thus, these results support the idea that TOTs are linked to the metacognitive control of retrieval.

Experiment 4 introduced two procedures new to the TOT literature. First, as just mentioned, after every correct response, the participants were asked if they had experienced a TOT before they retrieved the target. Indeed, retrospective TOTs were indicated on 22% of all correct responses. By contrast, TOTs following omission errors were greater (37%). Second, after every TOT, participants were asked about the emotional valence of the TOTs. These retrospective TOTs were more likely to be judged as "exciting" than the TOTs that followed retrieval failure, which were more likely to be judged as "frustrating." The rate of no-emotion TOTs was constant across the retrospective TOTs and those following retrieval failure. Future studies are planned for the purpose of examining why this pattern occurs.

GENERAL DISCUSSION

I begin the discussion with a review of the main results. In all of the experiments, TOTs were predictive of recognition, demonstrating accurate metacognitive monitoring. TOTs were also predictive of initial response time, demonstrating good metacognitive control, at least for omission errors. For these omission errors, participants consistently spent more time trying to retrieve items for which they gave TOTs than for items that they did not experience a TOT. In Experiments 1 and 3, participants distinguished between emotional and nonemotional TOTs (see Schwartz et al., 2000). In both of these experiments, emotional TOTs were associated with longer retrieval times than were the nonemotional TOTs. This trend was supported in Experiment 4, but it did not reach statistical reliability. In Experiment 4, retrospective TOTs were also correlated with slower retrievals than were n-TOTs.

An interesting question that the present study cannot address is whether rememberers engage in different kinds of retrieval strategies when they are experiencing a TOT than when they are not (A. S. Brown, 1991). It is possible that the increase in retrieval times is a function of using specific or more elaborate retrieval strategies when a TOT is experienced. If this is the case, the link between retrieval time and TOTs is increased by a step. TOTs may induce the selection of different retrieval strategies, which in turn cause a lengthening of the response times. However, retrieval strategies were not addressed in the present study, so the existence of strategy selection as an intermediate step in the control process is only speculation.

An important caveat to these findings must be acknowledged, and that is that the study demonstrates only a correlation between retrieval times and the likelihood of TOTs. Causality cannot be inferred from this study. It is likely that being in a TOT causes the person to spend more time attempting retrieval, but other explanations are

possible. First, because TOT judgments were made after the retrieval attempt, it is conceivable that participants might have given TOT judgments to the items for which it took them a long time to decide that they did not know. In fact, recent research suggest that fluency at the time of retrieval is a marker for metacognitive judgments (e.g., Barnes et al., 1999; Benjamin, Bjork, & Schwartz, 1998). Fluency could also account for the retrospective TOTs, which were also associated with slower retrievals than were n-TOTs. It is, of course, also possible that a third unknown factor, such as familiarity, drives both the long retrieval latencies and the likelihood of a TOT (Metcalfe, Schwartz, & Joaquim, 1993; Nelson et al., 1984).

This point is brought home by the recent success in using perceptual fluency, retrieval fluency, and familiarity as mechanisms for accounting a host of memory phenomena (see Bjork, 1999; Jacoby, 1991). In these models, fluency is used as a guide to judge memory performance; a fluently perceived target is one that is likely to have been seen before. In the present context, these nonanalytic judgments might also serve as the basis between TOTs and retrieval time. Nonfluent retrieval might serve to inform the rememberer that the retrieval might be difficult but possible, and a TOT would be experienced before retrieval is complete.

Nonetheless, the present data are best interpreted with Nelson and Narens's (1990, 1994) model of metacognition. In the Nelson–Narens model, a "meta" level exists, which monitors and controls the "object" level. In the present experiments, the metalevel processes were those by which TOTs were generated (see Schwartz, 1999, for a review of theories of TOT etiology), whereas the object-level processes were the processes that affected retrieval. In the Nelson–Narens model, the output of the monitoring system allows for effective control of the object-level system. Thus, the presence of a particular TOT substate triggers the rememberer to spend more or less time in attempting the retrieval of an unrecalled word. Indeed, Nelson et al. (1984) and Costermans et al. (1992) found a similar relation between retrieval time and a related metacognitive judgment, feeling of knowing. Thus, it is appropriate to consider the TOT as a metacognitive state with both monitoring and control functions.

In the Nelson–Narens model, metacognitive monitoring is an "imperfect" measurement of the object-level processes. Nelson and Narens (1994) argued that metacognitive judgments are generally predictive of memory performance, but seldom perfectly so. Indeed, in the present experiments, TOTs were highly correlated, roughly around .6, with memory performance, but the correlation was not 1.0. Imperfect monitoring, of course, means imperfect control. Nonetheless, because the monitoring is generally effective, it is adaptive to spend more retrieval time on items for which one is experiencing a TOT. Indeed, numerous studies have shown that relative to control conditions, resolution is higher for TOT items than for n-TOTs (e.g., Burke et al., 1991; Schwartz et al., 2000; Smith, 1994).

An additional finding from this experiment is the seemingly high rate (22%) of retrospective TOTs in Experiment 4. The methodology employed in this study has never been used before to examine TOTs, and therefore the data must be viewed with caution. However, a momentary experience of a TOT followed by resolution is frequently reported informally by rememberers and has also been documented in diary studies (e.g., Burke et al., 1991). Indeed, for researchers interested in lexical retrieval failures (Miozzo & Caramazza, 1997), these retrospective TOTs may offer a more clear window on examining momentary failures to retrieve words.

Furthermore, the present study highlights the importance of considering phenomenological aspects of human cognition (e.g., Nelson, 1996; Schwartz, 1999; Tulving, 1989). Nelson (1996) argued that current methodology in metacognition research provides a way of addressing age-old questions concerning human phenomenology. In metacognition research, human experience is queried but always compared with objective performance. This methodology allows us to examine the relation between phenomenological experience and cognitive processes. Thus, in the present experiments, phenomenological judgments were TOT substates, which were correlated with objective retrieval behavior. Schwartz et al. (2000) showed differences in monitoring effectiveness between different phenomenological aspects of the TOT, and the present study showed that TOT substates might also serve control functions for the rememberer. Consistent with the metacognitive perspective, rememberers' experience is an important aspect of human cognition because people use their feelings and experiences to control their behavior. Thus, the investigation of subjective aspects of human cognition may afford an examination of control processes.

REFERENCES

Barnes, A. E., Nelson, T. O., Dunlosky, J., Mazzoni, G., & Narens, L. (1999). An integrative system of metamemory components involved in retrieval. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 287-314). Cambridge, MA: MIT press.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, **127**, 55-68.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 435-459). Cambridge, MA: MIT Press.

Brennan, T., Baguley, T., Bright, J., & Bruce, V. (1990). Resolving semantically induced tip-of-the-tongue states for proper nouns. *Memory & Cognition*, **18**, 339-347.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, **109**, 204-223.

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning & Behavior*, **5**, 325-337.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory & Language*, **30**, 542-579.

Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and the feeling of knowing in question answering: The weight of infer-

ential processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 142-150.

Gardiner, J. M., Craik, F. I. M., & Bleasdale, F. A. (1973). Retrieval difficulty and subsequent recall. *Memory & Cognition*, **1**, 213-216.

Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology*, **89**, 151-174.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, **103**, 490-517.

Koriat, A., & Lieblich, I. (1974). What does a person in a "TOT" state know that a person in a "don't know" state doesn't know. *Memory & Cognition*, **2**, 647-655.

Krinsky, R., & Nelson, T. O. (1985). The feeling of knowing for different types of retrieval failure. *Acta Psychologica*, **58**, 141-158.

Mazzoni, G., Cornoldi, C., & Marchitelli, G. (1993). Do memorability ratings affect study-time allocation? *Memory & Cognition*, **18**, 196-204.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 851-861.

Miozzo, M., & Caramazza, A. (1997). Retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 1410-1423.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, **95**, 109-133.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, **51**, 102-116.

Nelson, T. O., Gerler, D., & Narens, L. (1984). Accuracy of feeling-of-knowing judgments for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General*, **113**, 282-300.

Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and feeling-of-knowing ratings. *Journal of Verbal Learning & Verbal Behavior*, **19**, 338-368.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In A. C. Graesser & G. H. Bower (Ed.), *The psychology of learning and motivation: Inferences and text comprehension* (Vol. 25, pp. 125-173). San Diego: Academic Press.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1-25). Cambridge, MA: MIT Press.

Perfect, T. J., & Hanley, J. R. (1992). The tip-of-the-tongue phenomenon: Do experimenter-presented interlopers have any effect? *Cognition*, **45**, 55-75.

Reason, J. T., & Lucas, D. (1984). Using cognitive diaries to investigate naturally occurring memory blocks. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory, actions and absent-mindedness* (pp. 53-70). London: Academic Press.

Ryan, M. P., Petty, C. R., & Wenzlaff, R. M. (1982). Motivated remembering efforts during tip-of-the-tongue states. *Acta Psychologica*, **51**, 137-147.

Schwartz, B. L. (1998). Illusory tip-of-the-tongue states. *Memory*, **6**, 623-642.

Schwartz, B. L. (1999). Sparkling at the end of the tongue: The etiology of tip-of-the-tongue phenomenology. *Psychonomic Bulletin & Review*, **6**, 379-393.

Schwartz, B. L., Travis, D. M., Castro, A. M., Smith, S. M. (2000). The phenomenology of real and illusory tip-of-the-tongue states. *Memory & Cognition*, **28**, 18-27.

Smith, S. M. (1994). Frustrated feelings of imminent recall: On the tip-of-the tongue. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 27-46). Cambridge, MA: MIT Press.

Stevens, J. (1986). *Applied multivariate statistics for the social sciences*. Hillsdale, NJ: Erlbaum.

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1024-1037.

Tulving, E. (1989). Memory: Performance, knowledge, and experience. *European Journal of Cognitive Psychology*, **1**, 3-26.

Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, **8**, 314-317.

Widner, R. L., Smith, S. M., & Graziano, W. G. (1996). The effects of demand characteristics on the reporting of tip-of-the-tongue states and feeling-of-knowing states. *American Journal of Psychology*, **109**, 525-538.

Wilkinson, T. S., & Nelson, T. O. (1984). FACTRETRIEVAL2: A PASCAL program for assessing someone's recall of general-information facts, confidence about recall correctness, feeling-of-knowing judgments for nonrecalled facts, and recognition of nonrecalled facts. *Behavior Research Methods, Instruments, & Computers*, **16**, 486-488.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.

Yaniv, I., & Meyer, D. E. (1987). Activation and metacognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 187-205.

NOTES

1. The author thanks Alan Brown and an anonymous reviewer for suggesting the investigation of prospective TOTs.

2. The author thanks Alan Brown for suggesting the investigation of emotional valence and TOTs.

(Manuscript received February 23, 1999;
revision accepted for publication March 10, 2000.)