APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **20**: 633–639 (2006)
Published online in Wiley InterScience
(www.interscience.wiley.com). DOI: 10.1002/acp.1215

Remembering Pragmatic Inferences

JASON C. K. CHAN* and KATHLEEN B. McDERMOTT

Washington University, St. Louis, MO, USA

SUMMARY

After hearing that 'the flimsy shelf weakened under the weight of the books,' people often remember having heard that the shelf broke–an inference that is not logically necessitated by the statement. Indeed, when asked to choose which word was heard—*weakened* or *broke*—people often mistakenly choose *broke*, a pragmatic inference, as having been heard (McDermott & Chan, 2006). On what basis are these decisions made? The present study explored the subjective experience accompanying memory for pragmatic inferences with remember/know/guess classifications (Gardiner & Java, 1991; Tulving, 1985). The phenomenological experiences accompanying such false memories were indistinguishable from those of true memories. The dual process signal detection account proposed by Wixted and Stretch (2004) provides a framework for understanding our results. Copyright © 2006 John Wiley & Sons, Ltd.

In everyday conversation, people often do not take the time to choose their words so precisely as to capture the exact meaning intended. As a result, the role of the listener is both to take in what is said and to infer the intended meaning. Hence, if one was told that 'the infant stayed awake all night; the karate champion hit the cinderblock; and the flimsy shelf weakened under the weight of the books,' it might be pragmatically reasonable—although logically incorrect—to infer that the baby cried, the champion broke the block, and the shelf collapsed. Most pertinent to the present report is that these inferences can become incorporated into one's memory for what was overtly stated (e.g., Brewer, 1977; Johnson, Bransford, & Solomon, 1973). On a two-alternative forced choice recognition test, in which the studied information (e.g., stayed awake) is paired with the pragmatic inference (e.g., cried), people often mistake the pragmatic inference as being the option that was previously encountered (McDermott & Chan, 2006).

A critical property of a pragmatic inference is that it changes the meaning of the original information, leading one to infer 'something neither explicitly stated nor necessarily implied' (Brewer, 1977, p. 673). Under circumstances when absolute accuracy is required, such as in legal settings, generating pragmatic inferences and later remembering them can lead to potentially damaging consequences.

Investigation of the mechanisms behind memory for pragmatic inferences received major attention during the 1970s and early 1980s (Harris & Monaco, 1978) but has since given way to other popular false memory research paradigms (e.g., Loftus, Miller, & Burns, 1978; Roediger & McDermott, 1995). Given the obvious relevance of pragmatic

*Correspondence to: Jason C. K. Chan, Department of Psychology, Washington University, Campus Box 1125, St. Louis, MO 63130-4899, USA. E-mail: ckchan@wustl.edu

inference to the understanding of false memories, surprisingly little is known about the phenomenological experience underlying them.

The remember/know procedure (Tulving, 1985) has been used to examine the phenomenological experience accompanying other types of false memories. False memories arising from exposure to semantic associates (e.g., remembering *sleep* after studying *bed, rest, awake*, etc.) are usually characterized by remember responses; that is, people claim to be able to vividly recollect the instance of presentation of the nonpresented word (Roediger & McDermott, 1995). In contrast, false memories arising from misleading postevent information are more typically characterized by know responses (Frost, 2000). In the current experiment, we investigated the phenomenological experience of erroneous recognition of pragmatic inferences using the remember/know/guess procedure.

METHOD

Participants

Twenty-four undergraduate students at Washington University participated for partial fulfillment of a course research requirement.

Materials and procedure

Variants of the 48 sentences reported in McDermott and Chan (2006) were presented in one of two conditions during the encoding phase. Sentences presented in the *implication* condition tend to induce pragmatic inferences (e.g., 'the new baby *stayed awake* all night'), whereas sentences presented in the *target* condition explicitly represent the meanings that were pragmatically implied by sentences in the implication condition (e.g., 'the new baby *cried* all night'). Note that these latter sentences did not contain pragmatic implications and were not expected to generate high levels of false recognition. Of primary interest was the phenomenological experience accompanying false recognition for the target sentences after encoding of implication sentences (e.g., studied *stayed awake*, recognized *cried*). That is, will people report vividly remembering the instance of presentation of a nonpresented word such as *cried*?

Participants studied 48 sentences for 3 seconds each. Twenty-four sentences were presented in the implication condition and 24 in the target condition. During the retrieval phase, three types of sentences were presented. In addition to the implication and target sentences, a set of *baserate* sentences was created to assess baserate false alarm probabilities. These sentences were syntactically similar to but meaningfully inconsistent with the implication sentences (e.g., the new baby *slept* all night).[1]

Participants saw 48 sentence frames presented one at a time during the retrieval phase. These sentence frames were shown with a blank, and the probe information appeared below the sentence frame. The following is an example of a test trial as seen by participants.

The karate champion _____ the cinder block.

- hit

[1]The terms implication, target, and baserate were chosen to simplify explication; we realize that the terms do not apply in any absolute sense and that all sentences contain some implications.

Participants were told to mentally fit the probe information into the blank of the sentence frame and to indicate whether the resulting sentence was identical to the one presented during encoding. They were further told that all sentences shown during the test phase were similar to sentences they had studied; therefore, they should only label a sentence 'old' when the sentence was, word-for-word, identical to the one they studied.

All subjects saw the same 48 sentence frames at test; what differed was the probe word(s) below the sentence frames: 16 frames were presented with the probe from the implication version, 16 from the target version, and 16 from the baserate version. Hence, there were six different study-test trial type combinations, and materials were counterbalanced across subjects.

Phenomenological experience accompanying recognition decisions were assessed using the remember/know/guess (RKG) task. A remember response indicated that subjects could recollect something specific about the moment at which they studied the sentence (e.g., if they heard a noise, the way they felt at the moment, etc.). They were instructed to say remember only if they could recollect the moment at which they encountered the *probe information* (i.e., the information that fit into the blank). A know response indicated that participants knew they had studied the probe but could not remember anything specific about the study episode. A guess response indicated that participants did not remember or know that they had studied the probe but guessed the probe had been studied.

RESULTS

The raw recognition probabilities and RKG results are presented in Table 1. We first discuss the overall recognition probabilities before we turn to the RKG data. Alpha level was set at $p = 0.05$. Partial eta squared (*pes*) indicates effect size.

Overall recognition probabilities

As can be seen in Table 1, the hit rates following study of implication sentences (0.73) and target sentences (0.84) were both quite high. These probabilities should be considered

Table 1. Proportion claimed 'old' and raw probabilities of remember/know/guess responses

	Response sentence type		
	Implication (e.g., stayed awake)	Target (e.g., cried)	Baserate (e.g., slept)
Study type			
Implication (e.g., stayed awake)	0.73 (0.18)	0.65 (0.15)	0.06 (0.10)
R	0.30 (0.27)	0.25 (0.18)	
K	0.31 (0.20)	0.22 (0.13)	
G	0.12 (0.14)	0.18 (0.12)	
Target (e.g., cried)	0.32 (0.06)	0.84 (0.10)	0.09 (0.11)
R	0.08 (0.14)	0.40 (0.25)	
K	0.09 (0.11)	0.24 (0.16)	
G	0.15 (0.17)	0.20 (0.13)	

Note: Standard deviations are presented in parentheses. Remember/know/guess data are not presented for baserate false alarms due to their extremely low probabilities.

against the baserate false alarm rates of 0.06 and 0.09, respectively. Hence, there was good discrimination between studied sentences and other plausible sentences that were not studied. Of primary importance was the false alarm rate in the condition in which participants studied implication sentences (e.g., the baby *stayed awake*) and then had to make a decision about a target sentence (the baby *cried*). The false alarm rate in this case was, as predicted, quite high (0.65). Further, it was much higher than in the complementary case (studying the target sentence and being tested with the implication sentence, 0.32). Hence, the data were asymmetrical such that there was more movement towards a false alarm from an implication sentence to a target sentence (than the reverse). These impressions are considered in greater depth as follows.

Prior to conducting inferential statistical analyses, the probability of claiming a baserate sentence 'old' was subtracted from the recognition probabilities for each subject (see Table 1 for raw probabilities). The resulting probabilities (presented here) were analyzed using a 2 (study: implication, target) $\times$ 2 (response: implication, target) repeated measures analysis of variance (ANOVA), which revealed a main effect for study condition, $F(1, 23) = 10.36$, $pes = 0.31$, a main effect for response, $F(1, 23) = 42.23$, $pes = 0.65$, and an interaction between the two factors, $F(1, 23) = 114.4$, $pes = 0.83$. The two main effects indicate that the implication study condition led to more *old* responses at test than did the target study condition and that the target test condition resulted in more *old* responses than did the implication test condition. In other words, when given a target sentence during retrieval, people tended to recognize that sentence as having been encoded, regardless of whether they studied it or a similar sentence with a pragmatic implication. The interaction can be understood as follows. If subjects were given a target sentence at retrieval (*cried*), the hit rate and false alarm rate were quite high (0.76 and 0.58, respectively, with the baserate probabilities subtracted); further, this difference between the hit rate and false alarm rate was smaller than the difference between hit and false alarm rates for the implication sentence presented at retrieval (e.g., stayed awake, hit and false alarm rates of 0.67 and 0.23, respectively).

Remember/know/guess

When overall recognition probabilities differ across item types, comparing their raw RKG probabilities can lead to scaling issues (see Rajaram, 1993, for discussion). To compensate for this problem, conditional probabilities are considered. Specifically, conditional RKG probabilities were calculated by dividing raw RKG probabilities by the recognition probabilities of the corresponding response type (for each subject). For example, if a subject's hit rate for the implication sentences was 0.80 and his/her corresponding 'remembering' probability was 0.40, then the conditional probability would be $0.40/0.80 = 0.50$. For cases in which no false alarm (FA) occurred (four participants did not have false alarms for the implication sentences at retrieval), zeros were applied to the conditional probabilities of RKG. Due to the large number of possible comparisons available for this data set, we discuss only the comparisons that directly address our research questions. In addition, we did not statistically analyze the probabilities of guessing because they were mathematically restricted by the combined probabilities of remembering and knowing.

The most important point to be gleaned from Figure 1 is in the two leftmost bars. After studying implication sentences (e.g., stayed awake), subjects were as likely to 'remember' the false alarms to the target sentences (e.g., cried, a pragmatic inference, $M = 0.34$) as they were to 'remember' the correct response ($M = 0.38$), $t(23) < 1$. Similarly, the probabilities

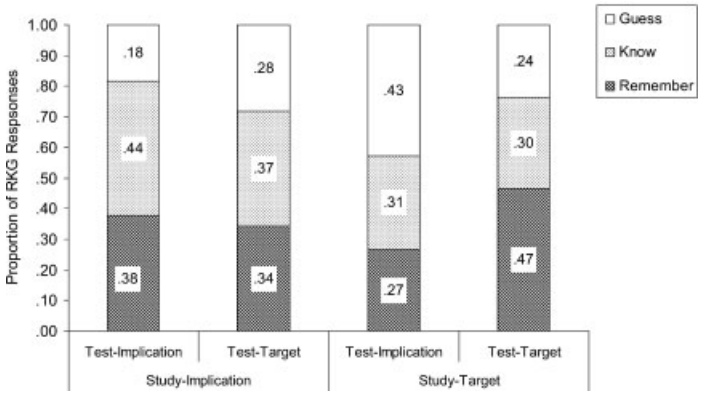

Figure 1. Distribution of remember/know/guess responses (based on conditional RKG probabilities) as a function of study sentence types and response sentence types

of 'knowing' did not differ between these sentence types, $t(23) = 1.25$, $p > 0.22$. A different pattern emerged, though, after studying the target sentences (e.g., cried). In that case, hits were accompanied by more remember response ($M = 0.47$) than were false alarms ($M = 0.27$), $t(23) = 2.92$, $SEM = 0.08$; know responses again did not differ between these sentence types, $t(23) < 1$.[2]

DISCUSSION

Two major findings emerge from these data. First, participants often erroneously recognized pragmatic inferences on free choice recognition even though they were told to claim a sentence old only if it was identical to the one they had studied. Second, after studying the implication sentences, participants were as likely to later remember similar sentences containing pragmatic inferences as they were to remember the actually studied sentence.

In the following discussion, we attempt to answer the question of why subjects would claim to remember the specific episode of encountering a pragmatic inference. We place our results into the broad perspective of how false remembering occurs. Traditionally, false recognition has been treated as noise or guesses in the dual-process framework of recognition (Gardiner & Java, 1991). However, experiments that show high levels of remembering nonpresented information (such as the current one) pose a problem for traditional dual-process theories. These problems stem mainly from two assumptions that certain variants of the dual-process theory hold. For example, one such assumption is that remember responses reflect the underlying recollective process (Gardiner & Java, 1991; Jacoby, Yonelinas, & Jennings, 1997); another assumption is that recollection is an errorless process (Jacoby, 1991). When one combines the above two assumptions, it is difficult to understand how false recognition can be accompanied by remember responses. If a subject mistakenly 'remembers' a pragmatic inference, and if remember responses reflect recollection, one must conclude that such responses can be erroneous. Alternatively, if we assume that recollection is an errorless process, one must conclude that erroneous remember responses can reflect familiarity, not just recollection (see Brainerd, Wright,

[2]Conclusions for remember responses were identical to those reported here if one considers raw (not conditional) probabilities.

Reyna, & Mojardin, 2001; Higham & Vokey, 2004, for other discussions on this topic). Hence, it is difficult to reconcile the results reported here within this framework.[3]

Recently, Wixted and Stretch (2004) proposed a *dual-process* signal-detection account that may help reconcile this paradox. They suggested that separate processes correspond to recollection and familiarity, but people combine the information provided by these processes into a continuous, unidimensional memory strength. When confronted with a recognition test containing a remember-know component, subjects set separate response criteria on this memory strength that correspond to remembering, knowing, and guessing. Note that we invoke this framework to understand our results on the basis of its theoretical appeal (rather than the quantitative aspect of signal-detection theory). That is, our interest in using this framework is not based on modeling responses using signal-detection theory.

There are two key points in this framework: first, information provided by recollection and familiarity can be combined, and second, remember and know responses are not process-pure. Because people combine the information provided by recollection and familiarity into a single dimension, any stimulus that is strong on this memory strength continuum will be classified as 'remembered,' regardless of whether the memory strength is based on a strong sense of recollection, a relatively weak sense of recollection combined with very strong familiarity, or a very strong sense of familiarity alone. Applying this logic to the current findings, participants could provide erroneous remember responses to a pragmatic inference either because the pragmatic inference was recollected or was very familiar, or both. Further, if people indeed combine the information provided by recollection and familiarity, it is not surprising that they frequently misattribute details associated with studied sentences to pragmatic inferences.

In line with this point of view, recent studies have suggested that remember responses cannot be considered process-pure. For example, the same items can be classified as remembered or known depending on how well other items on a recognition test are remembered (Bodner & Lindsay, 2003). Free recall, which is often considered the quintessential measure of recollection, has both 'remember' and 'know' components (Hamilton & Rajaram, 2003; McDermott, 2006). Variables that are traditionally considered to affect familiarity (e.g., perceptual priming) have also been shown to affect remember judgments (Higham & Vokey, 2004). Further, participants are unable to provide meaningful remember and know judgments when these judgments are made in isolation (e.g., making remembered/not-remembered judgments, Dolan, Donaldson, & Gallo, 2004). It is possible that remember/know judgments do accurately capture the subjective experience of recognition decisions, but to equate remember/know responses with the underlying retrieval processes such as recollection and familiarity can be problematic (i.e., the process-purity assumption, Jacoby, 1991).

To put the current argument into a more applied perspective, remember/know judgments may not be an ideal way to distinguish veridical from false memories, especially for false memories elicited by conceptual processes (such as the ones in the current paradigm). Similar to the pragmatic inference paradigm, false memories elicited by constructive, conceptually driven processes have repeatedly demonstrated levels of remember responses that are virtually indistinguishable from those for true memories (e.g., Roediger & McDermott, 1995). As Loftus (1979, p. 101) pointed out, 'one should not take

[3]We are not questioning the general usefulness of using remember/know classifications as an index of subjective experience. The point is that it is difficult to determine what processes underlie these subjective experiences (Jacoby, 1991).

high confidence [or in this case, remember responses] as any absolute guarantee of anything.''

REFERENCES

Bodner, G. E., & Lindsay, D. S. (2003). Remembering and knowing in context. *Journal of Memory & Language, 48*, 563–580.

Brainerd, C. J., Wright, R., Reyna, V. F., & Mojardin, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 307–327.

Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition, 5*, 673–678.

Dolan, P. O., Donaldson, D. I., & Gallo, D. A. (2004). Irrelevant-context effects on subjective states of remembering. Paper presented at the The 45th Annual Meeting of the Psychonomic Society, Minneapolis, MN.

Frost, P. (2000). The quality of false memory over time: Is memory for misinformation 'remembered' or 'known'? *Psychonomic Bulletin & Review, 7*, 531–536.

Gardiner, J. M., & Java, R. I. (1991). Forgetting in recognition with and without recollective experience. *Memory & Cognition, 19*, 617–623.

Hamilton, M., & Rajaram, S. (2003). States of awareness across multiple memory tasks: Obtaining a 'pure' measure of conscious recollection. *Acta Psychologia, 112*, 43–69.

Harris, R. J., & Monaco, G. E. (1978). Psychology of pragmatic implication: Information processing between the lines. *Journal of Experimental Psychology: General, 107*, 1–22.

Higham, P. A., & Vokey, J. R. (2004). Illusory recollection and dual-process models of recognition memory. *Quarterly Journal of Experimental Psychology, 57*, 714–744.

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

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 13–47). Mahwah, NJ: Erlbaum.

Johnson, M. K., Bransford, J. D., & Solomon, D. K. (1973). Memory for tacit implications of sentences. *Journal of Experimental Psychology, 98*, 203–205.

Loftus, E. F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19–31.

McDermott, K. B. (2006). Paradoxical effects of testing: Repeated retrieval attempts enhance the likelihood of later accurate and false memory. *Memory & Cognition*, in press.

McDermott, K. B., & Chan, J. C. K. (2006). Effects of repetition on memory for pragmatic inferences. *Memory & Cognition*, in press.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21*, 89–102.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803–814.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26*, 1–12.

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review, 11*, 616–641.

