Journal of Experimental Psychology:
Learning, Memory, and Cognition
2001, Vol. 27, No. 1, 14-33

0278-7393/01/$5.00 DOI: 10.1037//0278-7393.27.1.14

The Discrepancy-Attribution Hypothesis: II. Expectation, Uncertainty, Surprise, and Feelings of Familiarity

Bruce W. A. Whittlesea and Lisa D. Williams
Simon Fraser University

In the accompanying article (B. W. A. Whittlesea & L. D. Williams, 2001), surprising violation of an expectation was observed to cause an illusion of familiarity. The authors interpreted that evidence as support for the discrepancy-attribution hypothesis. This article extended the scope of that hypothesis, investigating the consequences of surprising validation of expectations. Subjects were shown recognition probes as completions of sentence stems. Their expectations were manipulated by presenting predictive, nonpredictive, and inconsistent stems. Predictive stems caused an illusion of familiarity, but only when the subjects also experienced uncertainty about the outcome. That is, as predicted by the discrepancy-attribution hypothesis, feelings of familiarity occurred only when processing of a recognition target caused surprise. The article provides a discussion of the ways in which a perception of discrepancy can come about, as well as the origin and nature of unconscious expectations.

Memory has two primary functions: the production of responses to stimuli and the evaluation of the quality of those responses (Whittlesea, 1997; Whittlesea & Leboe, 2000). The production function provides people with information about the world and appropriate responses; the evaluation function is the source of subjective reactions to that information. The experiments of this article investigate the evaluation process, tracing a chain of inference and attribution that leads from the initial perception of a word in a meaningful context to a feeling of familiarity for that word.

People are chronically involved in constructing percepts and cognitions about their environment. They encounter stimuli in various contexts and for various purposes; those contexts and purposes cause people to impose different organizations and interpretations on the stimuli. For example, on encountering the stimulus *ROSE,* people identify the letters, unitize them, and impose the meaning "red flower." In a different context, *ROSE AROSE,* they impose a different meaning on that unit ("girl's name"); asked to count the letters in *ROSE,* generate a rhyme, or judge the beauty of its referent, different perceptual organizations and cognitions arise. In extended acts of processing, such as when encountering *A ROSE IS A ROSE,* people integrate the informational contents of various parts of the experience into larger organizations, producing more elaborate cognitions. In this act of constructing an understanding of their environment, people also chronically evaluate the quality or coherence of their processing. This evaluation or monitoring function enables them to detect errors or inconsistencies in their processing, such as when encountering *A ROSE IS AROSE.*

Whittlesea and Williams (2000) proposed that, in evaluating their processing, people arrive at one of three major conclusions: the perception that everything went well (coherence), that something went wrong (incongruity), or that something strange happened (discrepancy). These perceptions are not properties of the stimulus per se: They are interpretations of the quality and content of the processing performed on a stimulus within a specific context. These three states lead to quite different subjective experiences and behavioral outcomes, as demonstrated later.

According to the discrepancy-attribution hypothesis (Whittlesea & Williams, 1998, 2000), the feeling of familiarity is selectively produced by the perception of discrepancy. In the accompanying article (Whittlesea & Williams, 2001), we described the origins of that hypothesis in the literature on heuristic decision making (e.g., Kahneman & Tversky, 1973; Nisbett & Ross, 1980; Schachter & Singer, 1962), heuristic remembering (e.g., Johnson, Hashtroudi & Lindsay, 1993; Lindsay & Read; 1994; Roediger & McDermott, 1995), and fluency-based familiarity (e.g., Jacoby, Kelley, & Dywan, 1989; Lindsay & Kelley, 1996; Whittlesea, Jacoby, & Girard, 1990). We demonstrated how the perception of discrepancy can produce an illusion of familiarity for orthographically regular nonwords such as *HENSION* and for words encountered after rhymes; we also showed the effect of that perception on judgments about music and about words presented in the "revelation" paradigm. We suggested that the illusion of familiarity for the *HENSION* items occurs because people experience the violation of an indefinite expectation, that the fluently pronounced regular nonwords would turn out to be some known word. In the experiments of this article, we examined more closely the importance of expectation and uncertainty in producing the feeling of familiarity.

The *HENSION* paradigm produces powerful illusions of familiarity and has been very useful in generating ideas about the discrepancy-attribution hypothesis. However, it has several drawbacks, including the fact that the perception of discrepancy is caused by two aspects of the same item (the fluency and mean-

Bruce W. A. Whittlesea and Lisa D. Williams, Department of Psychology, Simon Fraser University, Burnaby, British Columbia, Canada.

This research was supported by Grant 99826 from the Natural Sciences and Engineering Research Council of Canada and by Unilever Research Port Sunlight, Port Sunlight, England. We thank Jessica Yeung, Trevor Walls, and Mara Billings for their help in the collection of data.

Correspondence concerning this article should be addressed to Bruce W. A. Whittlesea, Department of Psychology, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada. Electronic mail may be sent to bruce_whittlesea@sfu.ca.

inglessness of the *HENSION* items), so that they cannot be manipulated independent of the items. In the experiments of this article, we instead used the "sentence-stem completion" paradigm. It consists of presenting a sentence stem, followed by a terminal or completing word, which is judged for recognition. The stem can be of low or high semantic constraint (predictiveness of the terminal word), the terminal item can complete the stem sensibly or not, the temporal relationship between presentation of stems and completions can be varied, and so on. This paradigm has the advantage of permitting great control over the person's contextually based expectations and the relationship between those expectations and the outcome. We suggest that variation in these factors can parallel the variety of interconnections between well- and ill-known stimuli occurring in ordinary or unusual settings.

To illustrate that paradigm, Whittlesea (1993) presented recognition probes after either a predictive sentence stem (e.g., The stormy seas tossed the *BOAT*) or a stem that was consistent with the outcome but not predictive of it (e.g., She saved her money and bought a *BOAT*). Probes presented in predictive contexts were named more quickly than probes presented after consistent stems; they were also more often judged old (i.e., as having been presented in the study list), resulting in elevated hits and false alarms. Whittlesea concluded that the subjects misattributed the extra fluency of reading probe items in predictive contexts to a prior experience of the items in training.

However, the discrepancy-attribution hypothesis suggests that the sequence of psychological events leading to the feeling of familiarity was not quite so simple as the perception and misattribution of fluency. Instead, it suggests that when consistent probes followed predictive contexts, the subjects experienced the event as a whole as being discrepant in some way, causing a surprise and consequent attribution of familiarity. In contrast, it suggests that when consistent probes followed nonpredictive contexts, the event was experienced as being merely coherent.

In the *HENSION* paradigm, surprise is induced by violation of an expectation. Fluent processing of items like *HENSION* causes an initial anticipation that they will turn out to be known units; that expectation is violated by realizing that *HENSION* is a nonword. However, in the sentence-stem paradigm, we believe the surprise comes about in a different and more interesting way. Presentation of a consistent probe (*BOAT*) after a predictive stem (Stormy seas tossed the . . .) does not violate the initial expectation; it actually confirms that expectation. We believe that the illusion of familiarity caused by such presentations results from a feeling of uncertainty that accompanies the expectation.

By this account, four elements are required to cause an illusion of familiarity in the sentence-stem paradigm. First, the person must generate a general expectation about what is to come; second, they must experience uncertainty about that expectation; third, the original expectation must be confirmed. The combination of expectation, uncertainty, and an unanticipated but consistent termination produces surprise. The final critical element is that the person is unable to ascribe his or her surprise to internal characteristics of the event. In that case, the person experiences a perception of discrepancy; unconscious attribution of that perception to a source in the past causes the conscious experience of a feeling of familiarity. In the experiments to follow, we investigated each of the components of this processing sequence. In Experiment 1, we manipulated the subjects' opportunity to experience uncertainty, by presenting probes with or without a pause after the stem. In Experiments 2 and 3, we varied the subjects' initial expectations by presenting predictive and consistent stems; in Experiments 3 and 4, we manipulated the subjects' ability to identify the source of surprise by presenting coherent versus incongruous stem-probe pairings. Finally, in Experiments 5 to 7, we manipulated the subjects' expectation and uncertainty by asking them to guess the outcome of sentence stems in foresight and in hindsight. Taken together, these experiments demonstrate a critical and varied role for expectation, uncertainty, and surprise in causing the feeling of familiarity.

Experiment 1: Expectation and Uncertainty

In Whittlesea's (1993) studies, sentence stems were presented for 2 s, during which the subject was to read the stem and then the probe was presented. Because the stems were short and could be read quickly, there was often a pause between the end of reading and the presentation of the probe. The length of that pause was not manipulated; it depended on the subject's reading speed. However, the existence of that pause meant that the probe was often episodically separated from the stem. We suspected that that pause might be important in creating the illusion of familiarity created by predictive contexts.

To test this idea, we conducted a more complex experiment. The subjects were initially exposed to a list of natural words (e.g., *SMILE, BENCH, BROOM*). In a subsequent recognition test, subjects were shown sentence stems followed by a new or old probe. The stems were of two types, consistent or predictive, differing in their constraint on what the terminal word could be and the semantic relatedness of the stem and probe. Consistent stems (e.g., "She couldn't find a place to put the . . . ") could be sensibly completed by any of a large number of words (e.g., *DOG, CLOCK, CAR, DESK*), including the actual test word (e.g., *BROOM*). In contrast, predictive stems (e.g., "She cleaned the kitchen floor with a . . . ") could only be sensibly completed by one of a small set of words (e.g., *BROOM, MOP, RAG,* or *CLOTH,* perhaps *WHISK* or *BRUSH*).

Half of the test items were presented after consistent stems (e.g., Later that day they went to the *BEACH;* there was a lot of noise in the *CROWD*). The other half were presented after predictive stems (e.g., They swam and played at the *BEACH;* the little boy got lost in the *CROWD*). Crossed with that factor, half of the test items were presented with a pause between presentation of the stem and presentation of the probe; the remainder were presented without a pause. By this means, we could discover whether the predictiveness of the sentence contexts was sufficient in itself to create feelings of familiarity or whether, instead, the predictiveness only had an effect when the subject had time to form an impression of his or her ability to anticipate what was yet to come.

Method

Subjects. Thirty-one Simon Fraser University students participated for course credit.

Procedure. We created a set of stimuli for this experiment consisting of 120 monosyllabic natural words, five to six letters long (see Appendix). These words were taken from the MRC Psycholinguistic Database (1998), using words of Thorndike-Lorge frequency between 40 and 100 and concreteness rating between 500 and 700. Examples include *crown, drink,*

glass, mouse, priest, and *smile.* We generated 120 pairs of sentence stems for these items; one stem of each pair predicted the probe word and the other was merely consistent with that word. Examples include "After the accident he was covered in *BLOOD*"; "On the corner table there was a spot of *BLOOD,*" and "The hikers got lost when they left the *TRAIL*"; "He helped the company build a new *TRAIL.*"

Sixty of the words, selected at random, were presented to be read in a training phase. These items were presented in a random order, one word at a time. Subjects were asked to read each word aloud in anticipation of an unspecified memory test. This training task was self-paced; the subjects struck a key to expose each successive word.

At test, all training items were shown again, in a fresh random sequence, interspersed with an equal number of unstudied items. Half of the old and half of the new items were presented with a predictive stem, and half were presented with a consistent stem. Crossed with those factors, the probe was presented along with the stem, in lowercase letters, on half the trials (e.g., He secured the gate with a lock and chain), centered on the screen. On the other half, the stem was presented with three dots (e.g., He secured the gate with a lock and . . .). On these trials, the subjects hit a key when they finished reading the stem; the screen then cleared and the probe (*CHAIN*) was presented after a further pause of 250 ms. As indicated by those examples, isolated terminations were presented in uppercase, and probes shown with their stems were presented in lowercase. Subjects were instructed to read each sentence aloud. They were told to watch for the three dots (" . . . "), judging the terminal word of the sentence as new or old if there were no dots but waiting for the probe to appear when dots were present. We assume a .05 level of significance throughout this article.

Results and Discussion

Overall, the subjects discriminated well between old and new items, $F(1, 30) = 331.90$, $MSE = 0.04$ (Table 1). There was a reliable main effect of predictiveness, $F(1, 30) = 4.57$, $MSE = 0.04$, but no main effect of pausing, $F(1, 30) = 1.42$, $MSE = 0.01$. More important, there was an interaction between predictiveness and pausing, $F(1, 30) = 3.77$, $MSE = 0.02$, but none of the other interactions approached reliability ($F < 1$ in all cases).

Unpacking the prediction-pause interaction, we found that when there was a pause the predictiveness of the stem had a reliable effect, $F(1, 30) = 7.59$, $MSE = 0.01$. When there was no pause, the predictiveness of the stem had no reliable effect, $F(1, 30) = 0.14$. The presentation factor (old/new) was reliable in both cases, $F(1, 30) = 170.80$, $MSE = 0.04$, and $F(1, 30) = 380.12$, $MSE = 0.02$, but the interaction between presentation and pausing was not, $F(1, 30) = 1.31$ and $F(1, 30) = 0.21$. We concluded that the predictiveness of the sentence contexts did indeed have an effect on recognition claims, as reported by Whittlesea (1993), but that it did so only when a pause occurred between the stem and probe. To test that conclusion further, we compared claims for the predictive-stem-and-pause cell against the means of the other three conditions (predictive-no pause, consistent-pause, and consistent-no pause). Items presented in predictive contexts after a pause were claimed old more often than items in the other three conditions, $F(1, 30) = 7.21$, $MSE = 0.01$. This was true both for items that had been presented in training, $F(1, 30) = 4.18$, $MSE = 0.01$, and items that were new at test, $F(1, 30) = 4.38$, $MSE = 0.01$. We concluded that claims of recognition were selectively influenced by the combination of predictiveness of the stem and a pause before the probe.[1]

We believe that the pause exerted its effect on familiarity by allowing the subject time to experience uncertainty about the outcome of predictive contexts. We perform a preliminary interpretation of the data in those terms; we test the idea further in the remaining experiments.

We suggest that reading a predictive context readies the person to incorporate the terminal word into the general meaning of the stem. In that sense, the subjects have an active expectation in advance of seeing the probe. However, the predictive contexts, although constraining, do not predict a specific word. If pressed, the person could use that general expectation to generate a specific word, but have no motive to do so in the circumstances of this test. The expectation is framed in terms of the meaning of the stem, not in anticipation of the specific meaning of the probe. For example, given "He secured the gate with a lock and . . . ", they are prepared for a meaning having to do with a "closing-locking" theme but not the specific meaning of the termination *CHAIN.* (These assumptions are tested later, in Experiments 5 to 7).

Thus, the subjects have a general expectation, a readiness to encounter any of a class of words that fit that meaning. When a predictive stem is presented followed by a pause, the subjects can experience this readiness, a sense of knowing roughly what the meaning of the word will be. However, because no specific word comes to mind to complete the stem, they also experience uncertainty. This uncertainty is a product of having an active general expectation. Without some expectation of what is to come, as in the case of the unconstrained consistent contexts ("She never knew that he had a . . . "), people do not experience uncertainty but instead unreadiness. Just as one cannot experience bafflement without realizing there is a problem to be solved, so one cannot experience uncertainty with no expectation of what is to come.

The terminal word, with its specific meaning, completes the theme very well when it occurs. It also shifts the developing meaning of the sentence in one direction or another (e.g., the

Table 1
Experiment 1: Effects of Predictiveness of Stem and Pause Before Probe, p(Claim "Old")

	Pause				No pause			
Predictiveness	Old items	New items	d_L	C_L	Old items	New items	d_L	C_L
Consistent	.61	.17	2.03	.57	.66	.15	2.39	.54
Predictive	.68	.22	2.01	.26	.63	.16	2.19	.56

Note. d_L and C_L are indexes of discrimination and bias, like d' and β (Snodgrass & Corwin, 1988).

[1] The reader might imagine that the increased claims of recognition in the predictive case resulted from semantic priming produced by the stem, augmenting the fluency of perceiving that item relative to probes in other conditions. Contrary to that idea, the effect of predictive context on recognition does not occur if no pause occurs between the stem and probe. Presumably, the predictive stems would have similar facilitatory effects on perception of the probe in the no-pause case, suggesting that such facilitation is not the source of the effect on recognition when the pause is present. Further, in Experiment 2, some probes were presented after completely predictive stems. Semantic priming should be at least as great after those stems as the predictive stems, but they caused less claims of remembering.

difference between "The young mother kissed her son on the *CHEEK/NOSE/SORE/RUN/SPOT/HOUR*"). The specific shift of meaning that occurs is not anticipated. The person thus passes rapidly from a state of general expectation and uncertainty to a state of specific understanding and resolution. This rapid change is accompanied by the feeling of surprise. In the case of predictive contexts followed by a pause, the surprise is normatively wrong: Actually, the word simply completes the stem sensibly. The surprise is due to the anticipation and uncertainty developed during the pause and the resolution of those feelings by the probe, not the semantic fit per se. We think of this effect as being like "waiting for the other shoe to drop": although strongly anticipated, the dropping of the other shoe is startling, because one does not know when it will occur. Similarly, people are surprised by the completion of a theme, when they have time to experience their anticipation and fail to complete it for themselves.

We thus suggest that when the subjects were shown consistent probes without a pause, they simply integrated it into the developing meaning of the sentence. Consequently, they experienced the whole event as being merely coherent. When shown a probe after a predictive stem without a pause, they again simply integrated it with its context. Although their fluency in processing the probe might have been greater in that case (cf. Whittlesea, 1993), that did not surprise the subjects: Not having time to experience a sense of uncertainty before encountering the probe, they again perceived the whole event as being merely coherent. When a pause was inserted, the subjects had time to experience the expectation raised by the stem. However, when the stem was merely consistent, the expectation was very vague: Hundreds of different words, of very different meanings, could sensibly complete those stems. Having only a vague expectation, they did not experience uncertainty; consequently, they were unsurprised by the probe, again experiencing it as being merely coherent with the stem. In contrast to those three cases, when the stem was predictive and the pause inserted, the subjects had time to experience both a general expectation about the probe and also uncertainty about what it would turn out to be. In this case, they perceived the probe as being surprisingly coherent with the stem: coherent because it fit well with the general expectation, but surprising because they had not anticipated that exact word and shift of meaning.

On experiencing surprise, people immediately attempt to discover its source. In doing so, they are guided by intuitive theories of cause and effect (cf. Ross, 1977) and by those aspects of the current event that are salient, given the task and context. In general, surprise experienced in processing a stimulus can be attributed to some overt or covert aspect of the stimulus, to one's current state, or to one's history. When the surprise can be attributed to some obvious characteristic of the processing event itself, it causes a perception of incongruity, not discrepancy, and does not cause a feeling of familiarity (see Experiments 3 and 4). The surprise occurring in our experiment was caused by the pause, which permitted an opportunity to experience uncertainty. However, that is not a salient aspect of the processing experience: The subjects instead focused on the semantic relationship of the probe with its stem, the meaning of the event as a whole. Unable to attribute the surprise to a source within the event, our subjects experienced a perception of discrepancy. In the context of a recognition experiment, knowing that many of the stimuli had been shown earlier, they often attributed the experience of surprise to a prior experience of the probe in the training phase. In doing so, they consciously experienced a feeling of familiarity.

In the remaining experiments of this article, we explored the boundary conditions of the state of uncertainty and its relationship to the feeling of familiarity. Having established that a pause between the stem and probe is critical in inducing the illusion of familiarity, we presented it in all succeeding experiments. However, we did not always control the pause by requiring subjects to strike a key when finished reading a stem. Instead, in some studies we presented the stem for a specific amount of time (2,000 or 2,500 ms) and then presented the probe automatically. The subjects usually completed reading of the stems within that period. There was thus a pause of varying length, depending on how quickly individuals read the stem. That procedure also caused the subjects to experience an illusion of familiarity for probes shown after predictive stems.

Experiment 2: Low, Moderate, and Complete Predictability

Different contexts lead people to different expectations about what is to come; they also cause people to expect outcomes with differing degrees of certainty. In Experiment 2, the subjects were again shown recognition probes in the consistent and predictive contexts used in Experiment 1. They were also shown some probes presented in completely predictive contexts. These stems were taken from well-known song lyrics or aphorisms, such as "row, row, row your boat" and "a rolling stone gathers no moss."

We suggested earlier that the feeling of familiarity depends on surprise, and that surprise depends on having some expectation about what will occur but also some uncertainty about the outcome. The completely predictive stems permitted the subjects to form an expectation about what word would terminate the stem: However, in that case, they should have no uncertainty. We thus expected that those stems would not produce an illusion of familiarity. In contrast, the predictive stems permitted the subjects to develop a general anticipation of what was to come, without being able to anticipate exactly what word would be presented. In that case, they could have both an expectation and also uncertainty. We excepted those stems to produce an illusion of familiarity.

Method

Subjects. Eleven Simon Fraser University students participated for course credit.

Procedure. We used the 120 words and stems developed for Experiment 1. We also collected 60 familiar phrases, such as "blood, sweat, and tears," "row, row, row your boat," and "raindrops keep falling on my head."

In a training phase, we presented 20 words taken from the known phrases and 40 words taken from the other set. Otherwise, selection of items to be included in training was random. These items were presented as in Experiment 1.

At test, we presented those items plus an equal number of unstudied items taken from each set. On each test trial, we showed the subjects a sentence stem, followed by a recognition probe. Words taken from known phrases were shown with their appropriate stem (e.g., The early bird gets the . . . *WORM*). Words from the other set were presented with either a predictive stem (e.g., The sailor marked their position on the . . . *CHART*) or a consistent stem (e.g., He spilled coffee all over the expensive . . . *CHART*) at random, subject to the condition that half of the new items

and half of the old items were presented in each type of context. The sequence of presenting test items was randomized. All conditions were rerandomized for each successive subject.

Unlike Experiment 1, we did not directly manipulate the pause between stem and probe. However, we ensured that there usually would be a pause. On each test trial, we presented a stem for 2 s, during which the subject was to read the stem aloud. After that interval, the probe stimulus was presented automatically. There was thus a pause of varying length, depending on how quickly the subject read the stem.

After naming the probe, subjects were required to decide whether the complete sentence was well known (such as a proverb or song line) or not and then to judge whether the test stimulus was new or old. We analyzed recognition claims on those trials on which the subjects correctly classified sentences as being well known or not. On average, subjects were about 95% correct on that judgment.

Results and Discussion

As can be seen in Table 2, the subjects discriminated well between new and old items of all three types, $F(1, 10) = 47.53$, $MSE = 0.04$. However, they claimed to recognize more items presented in predictive contexts than in either well-known phrases, $F(1, 10) = 10.49$, $MSE = 0.03$, or low-predictive contexts, $F(1, 10) = 7.60$, $MSE = 0.01$. There was no reliable difference between probes presented in well-known phrases and consistent contexts ($F < 1$). That is, recognition claims were higher when the probe was one of a small number of possible completions than when it was unpredictable (but consistent) or completely predictable. These results demonstrate that, to experience a feeling of familiarity, people must anticipate the outcome to some degree but must also be in a state of uncertainty.

Experiment 3: Violation of Definite Expectations

Earlier, we suggested that surprise can occur either through violation of an expectation or affirmation of a general expectation when one is in uncertainty. Having demonstrated that the latter source of surprise can produce an illusion of familiarity, we now turned to a deeper examination of the former.

In the accompanying article (Whittlesea & Williams, 2001), we suggested that the *HENSION* items induce an illusion of familiarity because their initial fluency creates an expectation that they will turn out to be known units; their lack of meaning violates that expectation, leading to surprise. The expectation in that case was indefinite: The subjects expected fluent items to turn out to be a known word, but not a specific word. We now tested whether violation of a definite expectation also produces a feeling of familiarity. In all of the conditions of this study, we created definite initial expectations by presenting stems taken from the aphorisms and song lyrics used in Experiment 2.

Violations of definite expectations may be more or less surprising. For example, on asking for a glass of water, one might be given a cup of coffee or something really odd, such as a squirrel. We produced violations in two ways: simple violations, in which the expected word was replaced by a word that simply did not fit the context, and clang violations, in which the expected word was replaced by a rhyming word. Simple violations (e.g., row, row, row your *SHEEP*) may not be very surprising, perhaps being experienced as merely wrong. However, clang violations often produce giggles, indicative of surprise (e.g., row, row, row your *GOAT;* let them eat SNAKE; a bridge over troubled *OTTERS*). We included clang violations to ensure that the subjects would experience a strong sense of surprise.

Table 2
Experiment 2: Effects of Indefinite Versus Definite Expectations, p(Claim "Old")

Stem	Old items	New items	d_L	C_L
Consistent	.67	.34	1.37	−.02
Predictive	.75	.46	1.25	−.47
Definite prediction	.70	.28	1.79	.05

Note. Consistent and predictive stems are from made-up sentences; definite-prediction stems are taken from familiar songs or aphorisms. d_L and C_L are indexes of discrimination and bias, like d' and β (Snodgrass & Corwin, 1988).

Method

Subjects. Fourteen Simon Fraser University students participated for course credit.

Procedure. We used the 60 familiar phrases created for Experiment 2. For each of these, we generated two additional types of termination, other than the usual one: clangs ("blood, sweat, and beers"; "row, row, row your goat"; "raindrops keep falling on my bed") and simple violations ("blood, sweat and ales"; "row, row, row your sheep"; "raindrops keep falling on my chair"). The clang outcomes matched the expected outcome on one dimension (they rhymed with the expected word) but mismatched on semantic relatedness to the stem. In contrast, the simple violations did not fit in any way; they were simply and obviously the wrong word for that stem.

Each stem was used only once in the experiment, one third of stems was completed at random by each type of terminal word (e.g., if "a stitch in time saves" was selected to be completed by "mine," the other words from the trio ["nine" and "yours"] would not be used for that subject). Half of the terminal words allocated to each condition were presented in a training phase, presented in the same way as in Experiment 1. At test, those items plus the remaining words of each condition were presented with their respective stems. Each stem was presented for 2 s, during which time the subject was to read it aloud; there was thus a pause of variable duration. After that interval, the probe word was automatically presented. The subject then made a recognition decision.

Results and Discussion

The subjects discriminated effectively between new and old items, $F(1, 13) = 125.05$, $MSE = 0.04$. However, neither the clang nor the simple violations of expectation caused an illusion of familiarity (Table 3). Instead, relative to the expected outcome condition, clangs and simple violations caused about 13% fewer false alarms than standard endings, $F(1, 13) = 8.60$, $MSE = 0.02$ and $F(1, 13) = 9.06$, $MSE = 0.01$, respectively. There was no apparent difference between clangs and simple violations in recognition claims, $F(1, 15) < 1$. We concluded that the subjects perceived their processing of both types of violation as simply incongruous with the stems rather than discrepant, resulting in a feeling of wrongness or silliness rather than familiarity.

The results also suggested that affirmation of definite expectations has little effect on feelings of familiarity. Although in this study the subjects could develop definite expectations about what word *should* complete these stems on all trials, the normatively

Table 3
Experiment 3: Effects of Violating Definite Expectations, p(Claim "Old")

Probe type	Old items	New items	d_L	C_L
Expected outcome	.70	.34	1.51	−.09
Clang	.72	.21	2.26	.19
Simple violation	.72	.21	2.26	.19

Note. d_L and C_L are indexes of discrimination and bias, like d' and β (Snodgrass & Corwin, 1988).

appropriate ending was presented on only one third of trials. The occurrence of that word was thus not a predictable event in the larger experimental context. Nevertheless, the hits and false alarms for such endings were very similar to those in Experiment 2 (see Table 2), in which the occurrence of a well-known stem was a completely valid cue predicting the terminal word. Coupled with Experiment 2, these results suggest that neither violation nor confirmation of a definite expectation affects the feeling of familiarity. Instead, that feeling selectively comes about through the development of a constrained but indefinite expectation.

Experiment 4: Violation of Indefinite Expectations

As discussed previously, we believe that surprise is essential for people to experience familiarity. However, as demonstrated in the last study, surprise does not always cause that feeling. As discussed early in this article, we believe that there are three basic categories of perception that people can arrive at in evaluating their processing: namely, coherence, discrepancy, and incongruity. Both of the latter perceptions occur when the outcome of the event is surprising, given the initial expectation. However, the perception of discrepancy leads to feelings of familiarity, the perception of incongruity does not. We believe that the difference between perception of discrepancy and of incongruity is that the person cannot identify the source of surprise in the first case but can in the second. That is, we believe that the perception of discrepancy consists of the perception of something odd about the quality or content of the event, without being able to identify what it is about that event that is odd.

We think that the *HENSION* stimuli cause an illusion of familiarity because they cause a surprise, and the source of surprise (the inconsistency between an initial expectation of meaningfulness, based on fluency of processing, and the later realization that the item is a nonword) is not clear to the subjects. Similarly, we believe that it is not obvious to people that the surprise they experience when probes follow predictive stems is caused by the pause inserted between them. In contrast, the inconsistency between the stems and violating outcomes of the last study ("To make a long story *SNORT*") is quite obvious. When the source of the surprise is clear, it causes a perception of incongruity; when it is not evident to the subject, it causes a perception of discrepancy and an illusion of familiarity.

However, there was one other major difference between those studies. The conditions so far examined that caused a feeling of incongruity involved a definite initial expectation ("a rolling stone gathers no . . . "). In contrast, in the conditions that produced illusory familiarity, the initial expectation was indefinite. For example, the initial expectation about *HENSION* was that it would turn out to be a word but not any specific word. Similarly, predictive contexts like "the mountain climber clung to the rocky . . . " suggest a constrained set of terminations but not a specific outcome. It is thus possible that the difference between the perceptions of discrepancy and incongruity results from the definiteness of the initial expectation, rather than the subjects' ability to identify the source of their surprise.

To test that idea, in Experiment 4 we examined the outcome of presenting recognition probes after predictive and inconsistent contexts, both of which aroused indefinite expectations. The predictive contexts were the same as those used in Experiment 1 (e.g., "She cleaned the kitchen floor with a . . . "). These contexts were constraining but not predictive of a specific word. For inconsistent contexts, we re-paired stems and items at random, creating such sentences as "She cleaned the kitchen floor with a . . . *TRAIL*" and "The hikers got lost when they left the . . . *BROOM*". Thus, up until the probe was presented, the two conditions were identical, causing identical degrees of expectation and uncertainty. Consequently, any difference in the amount of familiarity caused by those conditions could not be ascribed to a difference in the definiteness of expectation. We also presented some probes after consistent contexts (e.g., "She had never expected to see a . . . ") to demonstrate the effects of all three categories of perception—coherence, discrepancy, and incongruity—within a single study.

Our previous studies of the discrepancy-attribution hypothesis used a traditional train-test procedure, showing subjects a list of items in a study phase followed by a series of test trials presenting new and old items. For the present experiment, however, we followed a different procedure. On each trial of the experiment, the subject was shown a list of seven words presented in a rapid series. The stem and probe for that trial were then presented, and the subject judged whether or not he or she had just seen that probe in the preceding list. This change in procedure was intended to add to the generality of the discrepancy-attribution effect, showing that it occurs not only for remembering remote events but also in decisions about the immediate past.

Method

Subjects. Thirty-one Simon Fraser University students participated for course credit.

Procedure. The stimuli for these experiments were the words and stems created for Experiment 1. In addition, 720 more words having similar characteristics were selected from the database to be used as filler items in the training lists.

The experiment consisted of 120 trials, each trial consisting of a training and test phase. (This procedure was also used by Whittlesea et al., 1990, and by Whittlesea, 1993, in investigations of the fluency-attribution hypothesis of familiarity.) Each trial was initiated by a keypress by the subject. In the training phase of each trial, the subjects were exposed to seven words, presented in the center of the screen in RSVP format, at 284 ms per word. On half the trials, at random, one of the words in this list was also the probe word in the succeeding recognition test. On those occasions, the probe word was presented at random in any location of the training list except the first or last. Subjects were instructed to watch these lists closely in preparation for the test.

In the test phase of each trial, the subjects were initially shown a sentence stem, such as "After the accident he was covered in . . . " Subjects were asked to read the stem aloud, striking a key when finished. The probe item was then presented in capital letters after a further pause of 250 ms.

The subjects were asked to judge whether that word made sense in that context (to ensure they were attending to the meaning of the context) and then asked whether that word was new or old, indicating their choices by keypresses.

One third of the new and one third of the old items were presented in each of three types of context: namely, consistent, predictive, and inconsistent. Inconsistent contexts were created by swapping predictive contexts, producing sentences such as "The mighty wind died down to a . . . *CLOWN*" and "The best act at the circus was the . . . *BREEZE*." The assignment of items to presentation and context conditions and the sequence of trials were freshly randomized for each subject.

Results and Discussion

The subjects discriminated between new and old items fairly well, $F(1, 31) = 258.09$, $MSE = 0.04$ (Table 4). The manipulation of the three contexts also produced a reliable effect, $F(1, 31) = 4.10$, $MSE = 0.01$; the interaction was not reliable ($F < 1$).

Unpacking the main effect of context, the subjects judged items to be old about 5% more often in predictive contexts than consistent contexts, $F(1, 31) = 7.99$, $MSE = 0.01$, and also about 5% more often in predictive contexts than inconsistent contexts, $F(1, 31) = 5.04$, $MSE = 0.02$; the difference between consistent and inconsistent contexts was not reliable ($F < 1$). The effect of predictive contexts relative to the others was about as large for old items as new items; the interaction of context with presentation status was not reliable in either case ($F < 1$ in both comparisons).

These results demonstrate the difference among the perceptions of coherence, discrepancy, and incongruity. Both the perceptions of discrepancy and incongruity result from a surprise: Discrepancy because of the semantic fit of the probe after a pause, incongruity because the probe is semantically inconsistent with the stem. The perception of discrepancy causes a feeling of familiarity; the perception of incongruity does not.

In this study, the perception of both discrepancy and incongruity was induced by indefinite but constraining contexts. This demonstrates that the difference between those perceptions is not a response to the definiteness of the initial expectation. Instead, we concluded that it results from the subject's ability or inability to identify the source of surprise. When people experience inconsistency in their processing but can identify the source of that inconsistency within the event, they attribute it to that source. In that case, they experience the event as merely incongruous. In contrast, when they perceive an inconsistency between initial expectation and outcome but cannot identify the source within the event, they experience a perception of discrepancy. In that case, they attribute the inconsistency to an external cause; in the context of a recognition test, they attribute it to a prior experience of the stimulus and experience a feeling of familiarity.

Table 4
Experiment 4: Effects of Violating Indefinite Expectations, p(Claim "Old")

Context	Presentation status: Old	New	Mean	d_L	C_L
Consistent	.62	.11	.37	2.58	.80
Predictive	.65	.18	.42	2.14	.44
Inconsistent	.60	.12	.36	2.39	.79

Note. d_L and C_L are indexes of discrimination and bias, like d' and β (Snodgrass & Corwin, 1988).

The occurrence of the discrepancy-attribution effect in this study is particularly impressive given that the study sessions occurred only seconds in advance of the recognition decisions and in the context of a clear ability to discriminate new from old items. This demonstrates that the effect is not simply due to confusion about hazy memories of large sets of stimuli. Instead, the effect appears to reflect the operation of a fundamental interpretive and attributive mechanism that is directly responsible for producing feelings of familiarity.

Experiment 5: Eliminating Uncertainty

The experiments thus far demonstrate an illusion of familiarity that depends on the subject experiencing surprise about the goodness of fit of a probe word with its context. The elements that seem to be required to produce that illusion are a predictive context that constrains the set of sensible terminations of the stem, a pause between the sentence stem and the probe, and presentation of a probe that actually is consistent with the context. It does not occur if the context is not semantically constraining or if the probe is entirely predictable from the stem.

Taken together, these findings suggest that, to experience the illusion of familiarity, the subject must be in a state of both anticipation and uncertainty, which is resolved by presentation of a consistent word. However, that state of anticipation and uncertainty could have one of several characters. For example, given a predictive stem such as "The stormy seas tossed the . . . ," consistent terminations could include *BOAT, SHIP, YACHT, RAFT,* or even *SEAGULL.* Given that stem and a further pause, the subject might come to anticipate that the probe will be "something nautical" without going further. In that case, the expectation consists simply of the theme of the stem. Alternatively, in forming that expectation, the subject might think of one or more of the possible alternative terminations, such as "*BOAT,* or something like it." In that case, the subject is still in uncertainty (because of awareness that *BOAT* is not the only possible termination) but has in mind one or more specific possible outcomes. The question is whether the expectation that creates the illusion of familiarity consists of a latent readiness to incorporate one of a variety of consistent words into the theme presented by the context or whether instead it consists of conscious, explicit thinking about possible outcomes.

To test that idea, subjects in Experiment 5 were led to think of a word that was consistent with the stem before being shown the actual recognition probe. If the effect of predictive context on feelings of familiarity observed in previous studies results from explicit anticipation of a particular outcome, then thinking deliberately about possible completions should either augment the effect or not influence it. In contrast, if the effect depends on a latent readiness to integrate the probe with the stem, then deliberate thinking of possible outcomes should reduce the surprise of the probe item, thereby reducing the effect.

Method

Subjects. Twenty Simon Fraser University students participated for course credit.

Procedure. This study used words and sentence stems taken from the same stock as Experiment 4. However, unlike Experiment 4, this study utilized the standard procedure used in most recognition experiments; all training items were shown before any test items.

In the training phase, 60 probe items randomly selected from the stock were presented one at a time to be read aloud. The test was conducted in two phases. The first phase presented probe items after predictive and consistent contexts. In that phase, half of the training items were shown again, interspersed with an equal number of unstudied items, in a fresh random sequence. Half of the old and half of the new items were presented with a predictive stem; the other half of each type was presented with a consistent stem. On all trials, the stem was presented with three dots (e.g., He secured the gate with a lock and . . .). The stem was presented for 2,500 ms; at the end of that interval, the dots were automatically replaced by the probe item for that trial, printed in capitals (e.g., He secured the gate with a lock and *CHAIN*). The subjects were instructed to read the stems aloud as soon as they appeared. On presentation of the probe item, they were asked to judge it as new or old.

In the second phase, the remaining old items were presented along with the remaining unstudied items, again in a freshly randomized order. Half of each were presented after a predictive stem, half with a consistent stem. However, before the probe was shown, the subjects were required to guess a word that would sensibly fit that stem. Their guesses were recorded. In the case of predictive stems (e.g., She kissed her son tenderly on the . . .), the subjects could be expected to guess the actual probe (*CHEEK*) on some, but not all trials; on consistent trials (e.g., She had no desire to look at the . . .), it was expected that subjects would rarely guess the probe (*CLOCK*). After this guess, the probe word was presented as in Phase 1 and judged to be new or old.

The no-guess test phase was conducted before the guess test phase for all subjects. This was due to a concern that, when given experience of guessing outcomes on some trials, the subjects might begin to do so on trials when not asked to. It seemed better, for an initial examination of this issue, to observe how the subjects behaved in the no-guess condition spontaneously, uncontaminated by their experience of the guess condition. The problems associated with uncounterbalanced test order were corrected in Experiments 6 and 7.

Results and Discussion

In the first test phase (no guess), the subjects showed considerable accuracy in discriminating new from old items, $F(1, 19) = 108.41$, $MSE = 0.05$ (Table 5, top panel). More important, they were also about 6% more likely to claim an item old if it was presented in a predictive rather than consistent context, $F(1, 19) = 5.99$, $MSE = 0.01$, replicating the effect of predictive contexts observed in Experiments 1, 2, and 4. (The interaction was not reliable, $F < 1$).

The data for the second test phase (guess) were subdivided into cases in which the subjects completed the stem with the item that would later be presented as a recognition probe and trials on which they guessed a different word. As expected, they produced the actual probe word very rarely (4% of trials) for consistent stems, which could be completed by a huge variety of words. However, on predictive trials, the subjects guessed the word that would later be presented as the probe on 40% of trials: They did so more often for old probes (45%) than new probes (35%), $F(1, 19) = 6.53$, $MSE = 0.02$ (see Table 5, middle panel).[2]

Because of the low rate of guessing the actual probe after consistent contexts, we could not examine the effect of changing context on recognition for correctly guessed probes. However, we could perform that comparison for trials on which the subjects completed the stem with a different word (see Table 5, bottom right panel). In that case, the subjects discriminated fairly accurately between old and unstudied probes, $F(1, 19) = 43.19$, $MSE = 0.08$. However, there was no longer any effect of the type of context; the difference between the means was 0%, $F(1, 19) = 0.01$. Thus, provision of a predictive context rather than a consistent context induced an illusion of familiarity in Phase 1 but not in Phase 2.

Table 5
Experiment 5: Effects of Guessing the Probe

Phase 1: p(judge probe old) without guessing

Context	Presentation status: Old	New	d_L	C_L
Consistent	.68	.16	2.41	.45
Predictive	.73	.23	2.20	.11

Phase 2: p(guess actual probe)

Context	Presentation status: Old	New
Consistent	.05	.02
Predictive	.45	.35

Phase 2: p(judge probe old) after guessing

Context	Correct guesses: Old	New	Incorrect guesses: Old	New	d_L	C_L
Consistent	—	—	.61	.19	1.90	.50
Predictive	.66	.34	.60	.19	1.89	.52

Note. Dashes indicate that no recognition data are presented for correctly guessed trials with consistent contexts, because subjects guessed the probe on only 4% of those trials. d_L and C_L are indexes of discrimination and bias, like d' and β (Snodgrass & Corwin, 1988).

The logic of this study was that, to discover what creates an illusion, find out what increases or eliminates it. Phase 1 showed an illusion of familiarity resulting from provision of predictive contexts; incorrect guessing about the probe in Phase 2 eliminated the illusion. This suggests that in Phase 1 the subjects did not explicitly make guesses about candidates for the termination of the stem. Instead, it appears likely that they simply read the stems and became readied to incorporate consistent terminations into those themes. The expectation that created the illusion of familiarity remained latent rather than actual.

How does guessing a word that is not the same as the probe cancel the effect of predictive context? One possibility is that the subjects believed that they had committed an error in generating a different word. Perhaps that feeling was stronger on predictive

[2] The subjects' accuracy in guessing the probes in predictive and consistent contexts provided a check on our assertion that those two contexts actually differ in predictiveness. The low rate of correct guessing of the probe in the consistent case suggests that that context truly is nonpredictive; the 36% higher rate of guessing the probe in the predictive case suggests that those contexts were considerably more predictive of the probes we used. Similar comparisons can be drawn in Experiment 7.

trials because of the greater coherence of the probe words with their stems. This negative feeling canceled out the positive feeling associated with the predictive contexts, thus removing the illusion. We tested and rejected this possibility in Experiment 6. Instead, we suspect that guessing cancels the effect by eliminating the experience of uncertainty. Finding themselves able to generate a sensible outcome, the subjects experience a resolution of their expectation. In consequence, the probe, when it occurs, is not experienced as surprising but simply as an alternate outcome. We tested that idea further in Experiment 7.

Experiment 6: Expectation, Uncertainty, and Resolution

In Experiment 5, the subjects were asked to generate guesses about terminations for predictive and coherent contexts. Although that procedure was necessary to test how the act of generating a guess affects the subjects' behavior, it raised a few problems. One is the selectively greater accuracy in generating old than new terminations in predictive contexts; another is the great difference in accuracy of generating candidates after consistent versus predictive contexts. Third, guessing trials occurred only in the second phase of test, and so were both more temporally distant from the training phase and also occurred after the subject had already performed a large recognition test. Finally, in the first test phase, the probe occurred 2,500 ms after the onset of the context, but in the second test phase the time between presentation of the context and the recognition probe depended on the speed with which the subject generated a candidate.

These problems were avoided in Experiment 6. In this study, subjects were not asked to generate candidates, the critical conditions were randomly intermixed in a single test, and recognition probes were always shown 2,500 ms after onset of the context. In this study, on half the test trials, subjects saw only a stem and later a probe, as in the first test phase of Experiment 5 (uncompleted condition; e.g., "The cat eagerly awaited the . . . *CREAM*"). On the remaining trials, a plausible termination that was different from the recognition probe word was presented along with the stem (completed condition: e.g., "The cat eagerly awaited the *MILK*") and later replaced with a recognition probe that also completed the context plausibly (*CREAM*). The subjects were made aware that the termination that was first presented with the stem (*MILK*) was never an old word; they were to wait until that word was replaced by another before making a recognition judgment.

Thus, in the uncompleted condition, the subjects encountered probes after predictive and consistent contexts in the same way as in Experiments 1, 2, 4, and the no-guess condition of Experiment 5. We thus expected to observe higher claims of recognition after predictive than consistent contexts in that condition. In contrast, in the completed condition, the subjects always thought about a termination that was different from the actual probe but did not generate it for themselves. If the elimination of the effect of predictive context in the guess condition of Experiment 5 was due to the subjects having negative feelings about producing the wrong answer, that effect should be restored in the completed condition of this study. In contrast, if the guess condition eliminated the context effect because it caused the subjects to think of a possible termination, thus resolving the expectation engendered by the stem, then the effect should also be eliminated when the subjects were shown an alternative ending in the completed condition of this study.

Method

Subjects. Thirty-three Simon Fraser University students participated for course credit.

Procedure. Training was the same as in Experiment 5. Also, as in Experiment 5, the test phase of Experiment 6 presented old and new items in predictive and consistent contexts. However, in this case, the subjects were not asked to guess a plausible completion but instead were shown one on some trials. On half the trials they were shown a stem with three dots for 2,500 ms (e.g., "He secured the gate with a lock and . . . "). At the end of that interval, the dots were replaced by a probe, as in Phase 1 of Experiment 5 (e.g., "He secured the gate with a lock and *CHAIN*"). On the other half of trials, they were shown a stem along with a word that sensibly completed the stem but that was always new and was not the actual probe for that trial (e.g., "He secured the gate with a lock and *KEY*"). After 2,500 ms, that completion was replaced by the actual probe (e.g. "He secured the gate with a lock and *CHAIN*"). The subjects were instructed that the word initially presented with the stem was not a recognition probe and would always be new. They were asked to read the whole sentence as first presented, wait until the terminal word changed, and then perform a recognition judgment on the replacement terminal word.

The same pairs of plausible completions and target items were used in making both predictive and consistent sentences. The set of items used as probes was the same as that used as recognition probes in Experiments 1, 2, 4, and 5; the set of items used as nonprobe completions was newly created for this study. Unlike Experiment 5, trials with and without nonprobe completions were randomly intermixed in the test.

Results and Discussion

As can be seen in Table 6, the subjects discriminated well between old and new items, both when stems were presented without a completion, $F(1, 32) = 217.77$, $MSE = 0.04$, and when stems were presented with a completion, $F(1, 32) = 234.43$, $MSE = 0.03$. More important, when stems were presented without a completion, predictive contexts resulted in an average of 4% greater claims of recognition than did consistent contexts, $F(1, 32) = 5.44$, $MSE = 0.01$. However, when stems were presented with a nonprobe completion, claims of recognition were identical for predictive and consistent contexts, both in hits and false alarms, $F(1, 32) = 0.01$. (In both cases, the interactions were not reliable, $F < 1$.) That is, as in Experiment 5, predictive context had no effect on feelings of familiarity when a plausible completion intervened between stem and probe.

Unlike Experiment 5, the subjects in this study did not think of the nonprobe completions for themselves, so that they lacked whatever feeling accompanies such generation. However, the data support exactly the same conclusion: The illusion of familiarity caused by presenting a predictive rather than a consistent stem was

Table 6
Experiment 6: Effects of Providing Nonprobe Completions; p(Claim "Old")

	Uncompleted stems				Completed stems			
Context	Old	New	d_L	C_L	Old	New	d_L	C_L
Consistent	.73	.21	2.32	.17	.76	.27	2.15	−.08
Predictive	.78	.24	2.41	−.06	.76	.27	2.15	−.08

Note. d_L and C_L are indexes of discrimination and bias, like d' and β (Snodgrass & Corwin, 1988).

eliminated by thinking about a plausible completion of the context before seeing the recognition probe. Once again, that suggests that the subjects did not deliberately or explicitly think of possible terminations for sentence stems when those stems were presented without completions. Instead, it suggests that the expectation on which the illusion is based consists of a latent readiness to integrate any consistent recognition probe with the theme presented by the stem. People can use the stems to generate a specific expectation about the probe when asked to do so, but without that request they do not do so.

Because the subjects did not generate the completions, they were unlikely to think of them as wrong answers. It, therefore, seems unlikely that the elimination of the context effect occurred through negative feelings. Instead, we again suggest that presenting a completion before the probe resolves the expectation aroused by a predictive context. It thus prevents the person from experiencing uncertainty and, consequently, prevents the predictive context from inducing a feeling of surprise on encountering the probe.

Experiment 7: Hindsight, Foresight, Uncertainty, and Surprise

The idea that the effect of predictive context results from inducing a state of uncertainty would be more convincing if that state could be altered through changing the way the subjects thought of their task rather than changing the stimuli. In Experiment 7, we again presented probes after stems, requiring the subjects to guess the probe on some trials but not others. However, we also attempted to change their perception of uncertainty in those tasks, using hindsight and foresight judgments.

In Experiment 5, on trials in which subjects were asked to guess an outcome before seeing the recognition probe, their job was simply to generate a word that sensibly completed the stem. That task was easy to perform, particularly for predictive contexts. After generating a completion that sensibly conformed to the stem, the subjects might feel satisfaction, resolution, or coherence; they would certainly not experience a feeling of uncertainty about the fit of that item with the stem. No illusion of familiarity was observed in that condition. In Experiment 7, we again asked the subjects to generate a completion, but now they were instructed to attempt to predict the probe, not just generate a sensible ending (foresight condition). They were also asked to judge the likelihood that their guess would actually match the probe that was about to be presented. We reasoned that their confidence in guessing the probe would be low, because they would often guess wrongly and find that out immediately, when the probe was presented. Consequently, the generation task should now cause them to feel uncertainty just before making the recognition judgment. If they did so, and if uncertainty is a critical component of the context effect, then we should now observe an illusion of familiarity in this foresight condition, opposite to the results of the guess and completion conditions of the last two studies.

In the no-guess and no-completion conditions of the last two studies, we presented the probe automatically after a pause. In those conditions, we observed an illusion of familiarity caused by predictive contexts. As discussed earlier, we believe that the predictive stems caused the subjects to experience a general expectation, and the pause enabled them to experience uncertainty, causing the probe to be experienced as surprisingly coherent. In the present experiment, we also had a condition in which probes followed stems automatically after a pause. However, before making a recognition judgment on the probe, we now asked the subjects to judge how likely they were to have guessed that probe had we asked them (hindsight condition). In this condition, we expected the subjects to experience relatively higher confidence, based on the hindsight illusion (cf. Tversky et al., 1992). In that case, they would feel less uncertain in that condition than did subjects in the no-guess and no-completion conditions of the previous studies. If they did so, and if uncertainty is a critical component of the context effect, then we should observe a smaller context effect in this condition, the reverse of the effects of the no-guess and no-completion conditions of the earlier studies.

Method

Subjects. Twenty Simon Fraser University students participated for course credit.

Procedure. Training was the same as in Experiment 5. The test was conducted in two phases. One phase, the foresight condition, was similar to the guess condition of Experiment 5. In this phase, old and new items were presented in predictive and consistent contexts, factorially crossed as always. The stem was presented with three dots (e.g., "He secured the gate with a lock and . . . "). Subjects were asked to read the stem aloud and then to guess what word would later be presented as a probe. Next, they were asked to judge the likelihood that the word they guessed would actually be the probe. In making this judgment, the subjects used a 5-point scale ranging from 1 (*very unlikely*) to 5 (*very likely*). The midpoint, 3, was marked "maybe." After the subjects had made this likelihood decision, the actual probe was presented for a recognition judgment. This condition was identical to the guess condition of Experiment 5, except for the requirement to judge the likelihood that the guess was correct.

In the other phase, the hindsight condition, other old and new items were presented in other predictive and consistent contexts. The stem was presented for 2,500 ms; at the end of that interval, the dots were automatically replaced by the probe item for that trial (e.g., "He secured the gate with a lock and *CHAIN*"). The subjects were instructed to read the stems aloud as soon as they appeared. On presentation of the probe, the subjects were asked to judge the likelihood that they would have guessed that item for that stem, had they been asked to do so. They made this judgment using the same 5-point scale, with 1 indicating high confidence that they would not have guessed that word had they been asked to guess, 5 indicating high confidence that they would have guessed that word; and 3 indicating low confidence in judging either way. After that judgment, they judged the probe as new or old.

Half of the subjects engaged in the foresight task for the first half of the test and the hindsight task for the second half. The other half encountered those tasks in the reverse order. Assignment of items to the training phase and assignment of new and old items to hindsight versus foresight and predictive versus consistent contexts was random, subject to the requirements that half of the items were shown in training and that equal numbers of each were presented in each cell of the test design. The order of trials within phases was also randomized.

Results and Discussion

We first discuss the effects of context (predictive-consistent), presentation status (old-new) and task (hindsight-foresight) on the

subjects' confidence that they could or could not guess the probe.[3] Scores on the 5-point scale reflect the subjects' state of certainty or uncertainty. In foresight, high scores indicate a feeling of certainty that one is guessing correctly; low scores reflect a feeling of certainty that one is guessing incorrectly; and moderate scores indicate an inability to decide whether the guess is right or wrong. That is, high and low scores indicate a sense of high confidence in knowing the outcome (that the guess will or will not match the probe), whereas moderate scores indicate a feeling of uncertainty. The scale has the same meaning in the hindsight task: Low scores mean "I am certain I would not have guessed," high scores mean "I am certain I would have guessed," and moderate scores mean "Maybe I would have guessed, maybe not: I'm uncertain." The subjects appear to have used the scale in exactly this fashion. Although in various conditions, their scores departed from the central point, the mean of their judgments in the hindsight condition was 2.93 and the mean in the foresight condition was 3.02, giving a grand mean of 2.99.

The presentation status of the items (new vs. old) had no effect on confidence estimates, $F(1, 19) = 0.01$, nor did it interact with the other factors ($F < 1$ in all interaction tests; see Table 7, upper panel). Table 7 (lower panel) presents the data again, collapsed across the presentation condition. As can be seen in the table, predictive versus consistent context had a strong effect on the subjects' judgments: They generally thought they could predict (or could have predicted) the probe word in predictive contexts (mean claim = 3.75) but thought they could not in consistent contexts (mean = 2.22). The former mean was reliably larger than 3, $F(1, 19) = 124.10$, $MSE = 0.18$, the latter reliably less than 3, $F(1, 19) = 129.87$, $MSE = 0.20$. The only other important effect was an interaction between context and hindsight versus foresight, $F(1, 19) = 49.75$, $MSE = 0.09$. This effect can be seen in the lower panel of Table 7, in the difference between the size of the context effect in the hindsight condition (1.87) and in the foresight condition (1.20). This difference of 0.67 scale points occurred both because hindsight raised the subjects' confidence in their ability to guess the probe in the predictive context by 0.24 scale points, relative to the foresight condition, $F(1, 19) = 5.03$, $MSE = 0.03$, and also increased their confidence that they could not guess the probe in the consistent context by 0.43 scale points, $F(1, 19) = 10.93$, $MSE = 0.32$. In summary, predictive contexts led the subjects to conclude that they could predict (have predicted) the outcome; consistent contexts led the subjects to conclude that they could not predict (have predicted) the outcome; and the hindsight task increased the certainty of both of those judgments relative to the foresight task.

Table 7
Experiment 7: Effects of Hindsight and Foresight on Confidence Judgments

	Full data			
	Hindsight		Foresight	
Context	Old	New	Old	New
Consistent	2.04	1.97	2.48	2.37
Predictive	3.90	3.83	3.68	3.57

	Means, collapsed across old/new		
	Hindsight	Foresight	*M*
Consistent	2.00	2.43	2.22
Predictive	3.87	3.63	3.75
Difference:	1.87	1.20	

Note. Confidence estimates for guesses about the probe were made on a scale of 1 (*very unlikely correct*) to 3 (*maybe*) to 5 (*very likely correct*).

Turning now to the recognition data, in the hindsight condition the subjects claimed to recognize old items about 41% more often than new ones, $F(1, 19) = 76.35$, $MSE = 0.04$ (see Table 8, upper panel). However, the main effect of predictive versus consistent context was only 1% and not reliable, $F(1, 19) = 0.55$.

In the foresight condition, we first examined the subjects' accuracy in guessing the probe (see Table 8, middle panel). They guessed correctly on about 50% of predictive trials and about 7% of trials using the consistent context. As in Experiment 5, we subdivided the recognition data for the foresight condition according to whether the subjects guessed the probe correctly or not. As before, there were too few words correctly guessed in consistent contexts to examine the effect of context on recognition after correct guesses. However, we were again able to perform that comparison for incorrect guesses (Table 8, lower right panel). The subjects claimed to recognize old items about 31% more often than new ones, $F(1, 19) = 41.55$, $MSE = 0.06$. More important, the subjects claimed to recognize items 9% more often in predictive contexts than in consistent contexts, $F(1, 19) = 6.79$, $MSE = 0.02$.

As described earlier, this study was nearly identical to Experiment 5. The only major procedural change was that subjects in Experiment 7 were asked to make likelihood estimates about the match of guesses and actual probes. It is also similar in design to Experiment 6, in which a completion of the stem was or was not presented with the stem before presentation of the probe. The data of all three studies are summarized in Table 9 for easy comparison.

[3] Although not germane to the main argument, we also observed that the order of the test tasks had an influence on the extremeness of confidence estimates but not their valence. To examine the extremeness of scores, we took the absolute value of the difference between each score and 3 (the neutral point of the scale). These absolute deviation scores represent the person's degree of certainty that their guess will match (would have matched) the target; larger deviations indicate greater confidence, without preserving the valence of the decision. When the hindsight condition was performed first, the mean of the deviation scores for the hindsight condition was 1.00 and the mean of the deviation scores for the foresight condition was .61; the difference of .39 scale points was reliable, $F(1, 9) = 28.47$, $MSE = .03$, $p < .001$. When the foresight condition was performed first, the mean of the deviation scores for the hindsight condition was .89 and the mean of the deviation scores for the foresight condition was .68; the difference of .21 scale points was reliable, $F(1, 9) = 14.24$, $MSE = .02$, $p < .004$. In a between-groups comparison across the two task orders, the difference between hindsight and foresight was .18 scale points larger when the hindsight task was performed first, $F(1, 19) = 3.48$, $MSE = .04$, $p < .078$: That difference was due primarily to lower confidence in the hindsight task when that task came second. Although not reliable with only 10 subjects in each group, that comparison suggests that when the foresight task came first the subjects learned from that task how difficult it was to predict the probe, and in consequence behaved more conservatively in the hindsight task when it came second. Taking the two groups together, the mean deviation was .81 in the hindsight task and .65 in the foresight task, the difference being reliable, $F(1, 19) = 60.09$, $MSE = .02$, $p < .001$.

Table 8
Experiment 7: Effects of Hindsight and Foresight on Recognition

Hindsight (no guess): p(judge probe old)

Context	Presentation status: Old	New	d_L	C_L
Consistent	.69	.27	1.79	.10
Predictive	.70	.28	1.79	.06

Foresight (guess): p(guess actual probe)

Context	Presentation status: Old	New
Consistent	.08	.05
Predictive	.51	.50

Foresight: p(judge probe old) after guessing

Context	Correct guesses: Old	New	Incorrect guesses: Old	New	d_L	C_L
Consistent	—	—	.67	.36	1.28	−.07
Predictive	.79	.51	.76	.45	1.35	−.48

Note. Dashes indicate that no recognition data are presented for correctly guessed trials with consistent contexts, because subjects guessed the probe on only 7% of those trials. d_L and C_L are indexes of discrimination and bias, like d' and β (Snodgrass & Corwin, 1988).

As can be seen, the change in procedure in Experiment 7 caused a reversal of the effects of the context manipulation compared to Experiments 5 and 6. A reliable context effect was observed in the no-guess condition in Experiment 5 and the no-completion condition of Experiment 6, but not in the parallel hindsight condition of Experiment 7; no context effect was observed in the guess condition in Experiment 5 and the completion in Experiment 6, but a context effect did occur when subjects were asked to guess in the foresight condition of Experiment 7.

We suggest that this reversal occurred because the act of rating confidence changed the subjects' experience of uncertainty. We now provide a complete interpretation of Experiments 5 to 7, in terms of the constructs of expectation, uncertainty, and surprise. To assist the reader in tracing through the conditions, we provide reference numbers corresponding to the cells shown in Table 9. The interest in this analysis is which manipulations of independent variables cause a difference in the rate of claiming recognition (shown in the bottom row of each panel).

1. When stems were presented without a guess or completion in Experiments 5 and 6, the consistent contexts did not cause any particular expectation, and so the probe produced no surprise.

2. In contrast, when predictive contexts were presented without a guess or completion in Experiments 5 and 6, the subjects developed a general expectation; the pause between stem and probe allowed them to experience uncertainty, causing the probe to be experienced as surprising and consequently familiar. In consequence, the subjects experienced greater familiarity in the predictive condition of both studies than in the consistent condition (bottom row).

3. In the consistent condition of Experiment 7 (hindsight), the low semantic relationship between context and stem caused the subjects to decide they would not have predicted those items from those contexts. However, that was not surprising.

4. In the predictive condition of Experiment 7, stems caused the subjects to experience a general expectation, and the pause created uncertainty. However, the subjects also made a confidence decision about their ability to have guessed that probe had they been asked. In hindsight, observing the strong coherence between the stem and probe, they judged that they could have. Their experience of uncertainty was now replaced by a feeling of inevitability, that of course that word would occur in that context. Given that decision, the probe no longer felt surprising in its predictive context; consequently, the subjects had nothing to attribute to factors external to the event and failed to experience a feeling of familiarity. In consequence, there was no difference in recognition judgments after the two conditions (bottom row of panel).

5. When stems were accompanied by a guess or completion in Experiments 5 and 6, the consistent contexts again did not cause any particular expectation, and so the probe produced no surprise.

6. When predictive contexts were accompanied by a guess or completion in Experiments 5 and 6, the subjects developed a general expectation from the context, but the guess or completion resolved that expectation, so that they were no longer in a state of uncertainty when the probe occurred. Consequently, they experienced no surprise and no familiarity from that source. Claims of recognition were, therefore, no greater in this condition than in the consistent condition (bottom row of panel).

7. In Experiment 7 (foresight), the low semantic constraint of the context in the consistent condition caused the subjects to judge that they probably could not predict those items. The subjects therefore, were likely to experience uncertainty. However, because they had not developed a strong general expectation from the context, that uncertainty did not translate into surprise when the probe was shown.

8. However, in the predictive condition of Experiment 7, the subjects did not just have to generate a guess that was consistent with the stem, but also had to make a confidence decision that their guess corresponded to the probe about to be shown. In this fore-

Table 9
Experiments 5 to 7: Summary, p(Judge Probe Old)

Context	Exp. 5 (no guess)	Exp. 6 (no completion)	Exp. 7 (hindsight)
Consistent	.42 (1)	.47 (1)	.48 (3)
Predictive	.48 (2)	.51 (2)	.49 (4)
Difference	.06*	.04*	.01
	Exp. 5 (guess)	**Exp. 6 (completion)**	**Exp. 7 (foresight)**
Consistent	.40 (5)	.52 (5)	.52 (7)
Predictive	.40 (6)	.52 (6)	.61 (8)
Difference	.00	.00	.09*

Note. The data are the means from Tables 5, 6, and 8, collapsed across new and old cells. Numerals in parentheses correspond to numbered explanations in the text. Exp. = Experiment.
* $p < .05$.

sight judgment, they responded to the semantic constraint of the predictive stems, claiming on average that they could predict; but in foresight, their judgment was less confident than in hindsight. They were more likely to experience uncertainty in making the confidence decision. They had developed a strong general expectation from the context; they now experienced uncertainty in dealing with their guess. Instead of experiencing their guess as a resolution of the expectation, they instead experienced a sense that they did not know what was coming. Consequently, the probe, when it occurred after a pause, appeared to fit surprisingly well into the context. As described earlier, they could not attribute this surprise to any aspect of the processing event; in retrospect, the word was simply semantically congruent with its context. Unable to attribute their surprise to the event, they attributed it externally. In the context of a recognition test, they ascribed it to a previous experience of that word and experienced an illusion of familiarity.

In summary, across Experiments 5 to 7, we observed within each experiment that thinking about a completion (guessing it or having it provided) had different consequences for the occurrence of a context effect than did not thinking about a completion. However, we also observed that the factor had opposite consequences for the context effect when the subjects' task was to think of the completion as fitting the stem versus to think of it as being predictable from the stem. Because the stems were generally, but not definitely, constraining, thinking of a completion in the former way produced a sense of resolution, whereas thinking of them in the latter way produced a sense of uncertainty. This sense of certainty or uncertainty modulated the effect of the expectation that derived from the predictive stems. When the subjects experienced uncertainty after the expectation, the probe seemed to fit surprisingly well into the stem, causing a perception of discrepancy; when instead they experienced a resolution before the probe, the fit of the probe with the stem was experienced as simply coherent.

General Discussion

In these experiments, we explored a causal chain leading to the feeling of familiarity, consisting of (a) the establishment of a general, indefinite expectation, (b) the experience of uncertainty, (c) the surprising affirmation of the expectation, (d) failure to be able to attribute the surprise to the internal characteristics of the processing event, and (e) the consequent attribution of the surprise to a source in the past. Experiments 1 and 2 demonstrated that an illusion of familiarity occurs only if the stem is generally predictive of a class of outcomes; it does not occur if the stem is unpredictive or specifically predictive of the probe. We concluded that surprise depends on expectation. Nonpredictive contexts do not create an expectation: they, therefore, can be neither violated nor validated; instead, they can only produce a perception of coherence of processing. Definite expectations can be both validated and violated, but the former is unsurprising and the latter simply feels wrong (Experiments 2 and 3). By contrast, general, nonspecific expectations provide a basis for incorporating further information in a surprising way.

However, to experience surprise, one must also experience uncertainty. That uncertainty was initially set up by the indefiniteness of the predictive stems. However, to experience surprise, people must also have a chance to experience that indefiniteness. The pause manipulation of Experiment 1 provided that opportunity: When there was no pause after a predictive context, the probe was simply integrated into the developing theme of the stem, but when a pause followed a predictive context, the combination produced an illusion of familiarity. Experiments 5 and 6 demonstrated that the experience of uncertainty occurs because people do not explicitly generate outcomes for the general expectation raised by the stem; when an outcome was generated by the subject or supplied automatically, it canceled the surprise caused by the match between the probe and stem. However, Experiment 7 demonstrated that it is not the semantic relationship between the generated outcome and stem per se that removes the uncertainty, but instead thinking of it as a resolution of the expectation raised by the stem. When instead subjects were led to generate outcomes as candidates for the "real" outcome (the probe), in the foresight condition of Experiment 7, they again experienced a sense of uncertainty, so that the probe, when it occurred, was again surprising.

Finally, to experience a feeling of familiarity, the person must be unable to identify the source of their surprise. When outcomes affirmed definite expectations (Experiments 2 and 3) or violated definite or indefinite expectations in obvious ways (Experiments 3 and 4), the subjects did not experience feelings of familiarity, but instead perceived the event as simply coherent or incongruous. Instead, the perception of discrepancy and consequent feeling of familiarity occurs when indefinite expectations are violated in indefinite ways (the *HENSION* and rhyme-priming studies of the companion article) or validated in a way that normatively should not cause surprise (the occurrence of a consistent probe after a predictive stem and pause). Examining their performance after those events, the subjects could find no evident cause for surprise within the event itself. Consequently, they attributed the cause of the surprise to an external source, a prior experience of the probe item, and experienced a feeling of familiarity.

Varieties of Discrepancy

In previous studies (e.g., Whittlesea & Williams, 1998, 2000, 2001), we have induced false feelings of familiarity in various ways. We have come to believe that there are actually several distinguishable states that a person can experience, all of which involve the perception of discrepancy among some aspects of processing and cause feelings of familiarity but that differ in other important ways. One is the state that a person is in after reading a regular nonword like *HENSION*. This state, which we refer to as "surprising inconsistency," occurs when the person expects to be able to integrate two aspects of a processing experience but cannot do so and, further, cannot immediately identify why those characteristics conflict. As discussed in the companion article, we believe that people experience the fluency of processing orthographically regular nonwords like *HENSION* as incompatible with the fact that they are meaningless. They perceive this inconsistency as surprising but not definitely wrong (i.e., not actually incongruous). They resolve this perception of discrepancy by unconsciously attributing the fluency to a prior experience of the item and consequently experience a feeling of unfamiliarity.

We have also induced illusions of familiarity using stimuli that cause subjects to experience the opposite of what they experience in encountering *HENSION* stimuli. Whittlesea and Williams

(1998, Experiments 1 and 2) presented natural words and nonwords in training and test. We also presented some test words as pseudohomophones (e.g., the word *FROG* spelled as *PHRAUG,* or *TABLE* as *TAYBL*). Such presentations are initially processed with low fluency, because of their unusual orthography. They then cause considerable surprise as a stimulus that is initially thought to be a nonword catastrophically reorganizes itself into the pronunciation code of a well-known word. They also produced very high (false) claims of recognition (43% compared with 19% for words spelled normally). That is, they produce the same subjective outcome as the *HENSION* items (a feeling of familiarity) but clearly for a different reason: In this case, the discrepancy is between initial low fluency of processing and later discovering that the item is a well-known word. We refer to this way of inducing the perception of discrepancy as "surprising redintegration." (We suspect that this is also the state that a person is in after unexpectedly solving a word fragment; see Experiment 3 of the accompanying article; also Lindsay & Kelly, 1996.)

The present studies demonstrate a third variant of the perception of discrepancy: "surprising coherence."[4] Unlike the perceptions of "surprising inconsistency" and "surprising redintegration," the surprise in this case is produced by the affirmation of an expectation rather than a violation. In all three cases, there is a perception of discrepancy, but one is due to a perceived inability to integrate aspects of the processing, the second to a catastrophic change in one's understanding of the stimulus, and the third to surprise that the stimulus fits so well into its context. These differences seem to be accompanied by a difference in affective tone: "Surprising inconsistency" is puzzling, "surprising redintegration" is startling, and "surprising coherence" feels like achievement or unexpected resolution of tension. It is instructive to watch subjects encountering the three types of item: *HENSION* items often cause a frown, *PHRAUG* items often cause raised eyebrows, and an "Oh!" reaction, whereas the sentence-completion items used in these studies cause at most a nod.

There are also other ways in which the perception of discrepancy can be produced, through creating various tensions between one part of a processing experience and another. For example, in studies of the Deese effect (e.g., Roediger & McDermott, 1995), subjects are shown lists of words (e.g., *BED, NIGHT, DREAM, BLANKET*) that are associates of a prototype word (*SLEEP*); the prototype word itself is not presented in training but is shown in test along with associates that were shown earlier. Whittlesea (2001) observed that processing of these prototype words is facilitated considerably by their semantic relationship to the prior list; they can be identified as words in a lexical-decision task about 100 ms faster than nonprototype words not shown earlier. He also observed that these words show a reversed context effect. When presented in predictive contexts, prototype words were no more likely to be judged old than when presented without a context, whereas prototypes presented in inconsistent contexts were about 8% more likely to be judged old than when shown in predictive contexts. Whittlesea interpreted this effect as a consequence of the extreme semantic fluency associated with processing of prototypes prepared by lists of associates in training. Prototypes shown in inconsistent contexts are much more fluently processed than one should expect, given their relationship to the context, and that fact, rather than their semantic mismatch, becomes the primary issue to be explained. The subjects perceive discrepancy between the contextual mismatch and the semantic fluency of the prototype; not realizing that that great fluency occurs because the item is a prototype, the subjects attribute the discrepancy to a source in the past and experience a feeling of familiarity. This case is similar to that of "surprising redintegration" in that it is caused by a burst of processing fluency following a presentation that normatively would not seem to support it, but differs in that the person cannot later reconcile the elements of the event: The prototype remains inconsistent with its context.

This, then, is a fourth way of inducing an illusion of familiarity through the perception of discrepancy, through "surprising incongruity." We suspect that there may be yet more specific ways in which that perception can be brought about. The important commonality of all of them is that the person experiences a sense of surprise in processing an event without being able to identify the source of that surprise. That leads the person to perform an attribution to a source external to the current event; in a recognition study, that attribution is likely to be to the past, causing a feeling of familiarity.

The Nature of Expectations

Throughout this article, we have suggested that expectation is essential to surprise and a feeling of familiarity. However, we used the word "expectation" without carefully defining it. We tried to narrow the definition by the example of the experiments: the demonstration in Experiments 1 and 2 that only "general, indefinite" expectations cause an illusion of familiarity, and the demonstration in Experiments 5 and 6 that the expectation must be implicit, because the illusion did not occur when subjects explicitly thought of examples of what was expected, whether self-generated or suggested to them. We now want to clarify the general nature of the expectations that cause feelings of familiarity.

On some occasions, people may have a conscious and well-formed expectation of what is about to happen, as in anxiously waiting for one's spouse to appear at an appointed time and place. In the case of such deliberate, explicit expectations, we believe that people do not experience discrepancy or a feeling of familiarity from either violation or validation. Because the expectation is detailed and definite, events will either agree with it and be experienced as coherent, or disagree and be experienced as incongruous and wrong. On other occasions, people may have a general notion about the class of events that may be expected to occur in some context without any detailed predictions, as in listening for the first time to a symphony by a well-known composer. This is the state into which we placed our subjects by presenting indefinite but predictive stems in advance of recognition probes. In that case, people can also experience discrepancy and a feeling of familiarity, when a general expectation is affirmed surprisingly by a specific outcome. On yet other occasions, people may have no

[4] To avoid confusion, let us directly contrast our use of the terms "coherence" and "surprising coherence." The difference is parallel to the reaction of a teacher in receiving the same high-quality work from an excellent and a mediocre student. The former is perceived as good, the latter as surprisingly good: The fact that the text is coherent is itself surprising in the latter context. The former requires no external explanation; the latter might prompt the teacher to suspect a different source, such as plagiarism.

conscious thought whatsoever about the event that is about to take place, and yet still be prepared to experience various occurrences as coherent, discrepant, and incongruous. It is this latter case, of unconscious preparedness, that we think is most important in producing feelings of familiarity in ordinary life, certainly strong and spontaneous feelings. This is the case, for example, in watching a film and unexpectedly seeing an actor one cannot immediately identify but who one has seen in many other roles. The prior experiences prepare one to process the face efficiently and may even cause one to experience incomplete, fragmentary ideas unconnected to the present context, causing the perception of discrepancy; but the preparation is entirely implicit. The earlier experiences do not lead one to expect to see that face again, nor does the local context of the film cause one to anticipate any specific set of actors for that role. We have come closest to emulating that case by our use of the *HENSION* stimuli. Although the subjects knew they would be seeing new and old words and nonwords in test, on each trial there was no specific preparation for any item by a lead-up local context.

The preparedness that people have for events that are not consciously anticipated consists of the mass of prior experiences they have had in dealing with events in similar contexts. That information is not organized in advance (cf. the debate over unconscious abstraction of prototypes or schemas in the concept identification and implicit learning literatures; e.g., Neal & Hesketh, 1997); instead, it simply exists as a mass of resources, a readiness to incorporate this but not that feature or stimulus within the current event. The expectation is diffuse, latent, and distributed across processing episodes. Seeing a white crow is startling, not because one has previously worked out the proposition that "all crows are black" but because one can readily integrate the shape, size, and movement of the crow with prior experiences but cannot easily integrate its color with the other information. Ready assimilation of part of the stimulus coupled with inability to assimilate the remainder causes a surprise reaction: Whether that reaction is of incongruity (causing a feeling of wrongness) or of discrepancy (causing a feeling of familiarity) depends on whether or not one can identify what part of the experience is causing the problem.

Norms on the Fly

The idea that feelings of familiarity result from surprise entails that people make some kind of comparison between the actual outcome of a processing event and the outcome that they could normatively expect for that item. This idea was first raised by Jacoby and Dallas (1981). They observed that prior study facilitates naming of low-frequency words more than that of high-frequency words, but that old high-frequency words are still named faster. That is, naming of old high-frequency words is more fluent, but the fluency of naming old low-frequency words is farther from its baseline than is that for high-frequency words. They further observed that old low-frequency words are more likely to be claimed old than are old high-frequency words. That is, in making recognition decisions, people appear to be impressed not by the fluency of processing per se but rather by the difference between their actual fluency and the fluency that could ordinarily be expected for that item. In some way, participants in that study appeared to have a norm for each type of word against which they could compare the actual fluency of items in test.

In that experiment, the norm appeared to depend simply on the class from which test items were taken (i.e., low- vs. high-frequency words). However, later studies demonstrated that the norm people use to make recognition decisions also takes into account the significance of the context in which items are shown in test. This was demonstrated by the Jacoby and Whitehouse (1989) study, described in the companion article (Whittlesea & Williams, 2001). In that study, subjects discounted the likelihood that a test word was old when aware of another potential source of fluency, namely a prior occurrence of that same word immediately before the test presentation. The rhyme-priming study in the companion article (Experiment 3) extended this idea, demonstrating that people discount fluency of processing induced by context when thinking about that context as a source of fluency but fall into an illusion of familiarity when that source is backgrounded. Whittlesea and Williams (2000, Experiments 3 to 5) also showed that the feeling of familiarity caused by the *HENSION* items can be substantially modified by variations in local context. That is, as suggested by Kahneman and Miller (1986) in a different context, the norm is computed on the fly; it is not a precomputed, stable standard, applied in a consistent way across classes of stimuli. Instead, the evaluation that produces feelings of familiarity is better characterized as part of a constructive process, an attempt to integrate the various aspects of the current event, based on what aspects of the current situation are salient to them, given the current task and context (Marcel, 1983).

The experiments of the present article push that idea further, showing that a feeling of familiarity for a test stimulus depends critically on the predictiveness of the context, the opportunity to experience uncertainty, the semantic fit of the probe with the stem, and the ability or inability to identify why the occurrence of the probe in that context is surprising. That is, the norm that sets up the basis for surprise, and hence the basis for the feeling of familiarity, is a complex sequence of events occurring in the immediate context of the test item.

Norms on the Fly Versus Distribution-Based Norms

The conclusion that familiarity results from a norm computed on the fly seems to be in conflict with the assumptions of a common way of understanding recognition performance: the signal-detection paradigm. Very briefly, in a standard signal-detection theory (SDT) account of recognition, each test item is assumed to cause a mental event, which forms the evidence for the recognition decision. This evidence is assumed to differ in strength from item to item, forming a distribution for each class of item (old and new). On average, the mean of the distribution for old items (targets) is greater than that for new ones (lures). People are assumed to set a criterion somewhere along the strength-of-evidence dimension, claiming to recognize any item whose strength exceeds the criterion.

SDT accounts can also explain performance in more complex situations, such as the "mirror effect." That effect consists of the observation that when two classes of items (e.g., low- and high-frequency words) are intermixed in a recognition study, items of one class, which, when old, attract more hits than items of another class, also attract fewer false alarms when new (e.g., Glanzer & Adams, 1985). Investigators of the effect have proposed two general solutions to this effect (cf. Stretch & Wixted, 1998). By

one general form of account, the distribution of evidence for novel items of each type is the same, whereas the distributions for old items differ between the classes. This type of account invokes a shifting criterion, set to be relatively liberal or conservative by the subject, depending on their appreciation of the class of item from which a particular test item comes (e.g., Gillund & Shiffrin, 1984; Hintzman, 1994). By the second kind of account, the subjects use a single criterion for all items: The reversal shown by the "mirror effect" occurs because the relative amount of evidence associated with members of the two classes is also different when new and is actually reversed to the relationship when old (e.g., Hintzman, 1988; Hirshman & Arndt, 1997).

Without getting into much detail, we wish to emphasize two points about SDT accounts in general. First, their starting assumption (usually undiscussed) is that the subjects' experience of test items differs only on some strength-of-evidence dimension, usually called "familiarity." This strength of evidence is treated as a primitive: Having experienced some amount of evidence, the subject goes on to make a decision about whether it is strong enough to call the item "old" by comparing it with the criterion. By such accounts, the important issue to be investigated is the shape and location of distributions of this evidence along the underlying dimension. The second point, which is consequent to the first, is that familiarity is understood as a property of classes of items. Although there is variation of "familiarity" within a class, forming a distribution, the standard that is applied to any item in judging recognition is based on the average properties of classes of stimuli, such as low versus high frequency. Further, the amount of "familiarity" that any stimulus causes in test depends on classwise properties, such as how much familiarity members of that class produce when new, how members of that class are affected by study, and so on.

We suggest that such classwise properties are the wrong level of analysis. First, the feeling of familiarity is not a primitive. It is instead the by-product of an unconscious evaluation process that takes into account what is known about that item as well as the quality of current processing. From the SDT perspective, this objection could be answered by saying "but that is just a restatement of the SDT account: simply assume that fluency of processing is the strength of evidence that people experience, and you have an account of an inferential process that takes into account what is known about that item (its class and the implications of that for processing ease) as well as its actual current quality." That kind of account could indeed explain the *HENSION* effect documented in the companion article, suggesting that people apply an inappropriately conservative criterion to the regular nonwords because of their overt class. However, as demonstrated by the rhyme-priming studies in that article and the experiments in this article, the evaluation process can go well beyond a simple appreciation of the fluency of processing an item and the implications of its nominal class. The average, statistical properties of classes of stimuli can be used to predict performance on those stimuli when that is all that is manipulated in a study, as in the typical SDT experiment. However, when more extensive and idiosyncratic properties of experience are manipulated, as in the present studies, it becomes clear that people can take into account much more individual and detailed information about stimulus processing than is suggested by SDT accounts. This, of course, does not mean that SDT accounts are wrong, only that, in using the statistical properties of classes of stimuli to predict performance, they wrongly imply that that is the level of analysis at which people come to experience a feeling of familiarity. An SDT account could simulate our data; however, to do so, it would have to establish distributions of processing experience based not only on the frequency and presentation status of the words but also the predictiveness of the local context, the availability of an interval to experience uncertainty, the effects of completing the context before seeing the probe, and the ability or inability to identify the source of surprise within the event.

Thus, we suggest that the proximal mechanism leading to a feeling of familiarity is not comparison of processing fluency to a criterion established by the average properties of classes of stimuli. Instead, we suggest that, in arriving at that feeling, people use a standard created on the fly, incorporating all of the aspects of the stimulus and context that are salient given the current task and making what sense they can of that information, given their intuitive theories of cause and effect (Marcel, 1983; Ross, 1977; Nisbett & Ross, 1980).

Summary

This article and its companion are about bias: not simple response bias, but bias in constructing percepts and cognitions and in evaluating that processing. We have demonstrated a number of stimulus and contextual factors that shift people's willingness to claim test items to be old, independent of the true distribution of that property across stimuli. The readiness with which people fall for such manipulations might seem to suggest a weakness in the system, a proneness to error. However, we argue that the same mechanism that makes the human mind susceptible to such systematic illusions of familiarity is also the mechanism that permits it to discriminate real differences in prior occurrence. The veridical feeling of familiarity resulting from an actual prior experience is caused by a perception of discrepancy between the current experience of the stimulus and some expectation, just as is the illusory feeling.

We argue that people are never in direct contact with the stimulus properties that they perceive. They are always in the position of constructing attitudes toward external stimuli, evaluating the variety of sensory input, imposing organization on it, and attributing identity and meaning to it. This is the production function of mind, controlled by a complex interaction among memory, stimulus structure, task, and context (cf. Whittlesea, 1997; Whittlesea & Leboe, 2000). At the same time, people construct attitudes toward the quality of their processing of external stimuli, evaluating the coherence of the stream of perception, cognition, and behavioral response. That evaluation feeds back on the production function, the perception of coherence causing it to continue, the perceptions of discrepancy and incongruity causing it to halt and change course. These perceptions are attributed to a source in the present or past, giving rise to subjective feelings of familiarity or goodness, just as the attribution of meaning to the external stimulus gives rise to a subjective experience of knowing what that stimulus is. The same constructive, evaluative, and attributive mechanism that allows people to know that "a rose is a rose" also permits them to know they have heard that phrase before.

References

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition, 13,* 8–20.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory-model. *Psychological Review, 95,* 528–551.

Hintzman, D. L. (1994). On explaining the mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 201–205.

Hirshman, E., & Arndt, J. (1997). Discriminating alternative conceptions of false recognition: The cases of word concreteness and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1306–1323.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306–340.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General, 118,* 126–135.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93,* 136–153.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

Lindsay, D. S., & Kelley, C. M. (1996). Creating illusions of familiarity in a cued recall remember/know paradigm. *Journal of Memory and Language, 35,* 197–211.

Lindsay, D. S., & Read, J. D. (1994). Psychotherapy and memory for childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology, 8,* 281–337.

Marcel, A. J. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology, 15,* 238–300.

MRC psycholinguistic data base. (1998). Nedlands: University of Western Australia. Department of Psychology. Retrieved January 12, 1998, from the World Wide Web: *http://www.speech.cs.cmu.edu/comp.speech/Section1/Lexical/mrc.html*

Neal, A., & Hesketh, B. (1997). Episodic knowledge and implicit learning. *Psychonomic Bulletin & Review, 4,* 24–37.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice Hall.

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.

Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 10, pp. 21–54). New York: Academic Press.

Schachter, S., & Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review, 69,* 379–399.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50.

Stretch, V., & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1397–1410.

Tversky, A., Kahneman, D., Gentner, D., Collins, A., Fischoff, B., Hoch, S. J., & Loewenstein, G. F. (1992). Inferential aspects and judgment under uncertainty. In T. O. Nelson (Ed.), *Metacognition: Core readings* (pp. 377–436). Boston: Allyn & Bacon.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1235–1253.

Whittlesea, B. W. A. (1997). Production, evaluation, and preservation of experiences: Constructive processing in remembering and performance tasks. *Psychology of Learning and Motivation, 37,* 211–264.

Whittlesea, B. W. A. (2001). *Illusions of familiarity and the discrepancy-attribution hypothesis: The Deese effect in context.* Manuscript submitted for publication.

Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language, 29,* 716–732.

Whittlesea, B. W. A., & Leboe, J. P. (2000). The heuristic basis of remembering and classification: Fluency, generation, and resemblance. *Journal of Experimental Psychology: General, 129,* 84–106.

Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? The unexpected basis of feelings of familiarity. *Acta Psychologica, 98,* 141–166.

Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 547–565.

Whittlesea, B. W. A., & Williams, L. D. (2001). The discrepancy-attribution hypothesis: I. The heuristic basis of feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 3–13.

Appendix

Predictive and Consistent Sentence Stems

The first word of each pair was used as a recognition target in Experiments 1 to 7. The second word was used as an alternate completion in Experiment 6. The first of the pair of stems was used in predictive conditions, the second stem in consistent conditions.

AISLE, HALLWAY
The bride and groom walked slowly down the
In the middle of the restaurant she found the
BADGE, CARD
The policeman identified himself with his
Every elected official should carry a
BEACH, LAKE
They swam and played at the
Later that day they went to the
BENCH, CHAIR
They sat in the garden on an old oak
She couldn't help liking the charming
BLOOD, GLASS
After the accident he was covered in
On the corner table there was a spot of
BRAIN, MIND
The difficult problem strained his
What she admired most was his
BREEZE, HUM
The mighty wind died down to a
He predicted there might be a
BRIDGE, PLATFORM
The train roared out of the tunnel and onto the
He got a well-paid job working on the new
BROOM, MOP
She cleaned the kitchen floor with a
She couldn't find a place to put the
CHAIN, KEY
He secured the gate with a lock and
He found a new way to use the old
CHALK, PENCIL
The teacher wrote her name in
They are as different as corn and
CHART, MAP
The sailor marked their position on the
The clerk spilled coffee all over the
CHEEK, EAR
She kissed her son tenderly on the
The south wind blew gently on his
CHEESE, YOGHURT
She likes most dairy food but not
That company is famous for their
CHEST, FACE
The wrestler had a very hairy
The young singer worried about his
CLIFF, SPIRE
The mountain climber clung to the rocky
Behind his house there is an interesting
CLOCK, CALENDAR
As the time drew near they looked at the
She didn't have any desire to look at the
CLOUD, FAULT
The sky was blue without a
There in the middle was a
CLOWN, TRAPEZE
The best act at the circus was the
He's really not suitable to be a
COACH, RINK
The hockey team hired a new
She went for dinner with the
COUCH, SOFA
After dinner he watched TV on the
They felt a need to replace their old
COURT, TRIAL
The judge and jury walked into the
In the center of the building is the
CREAM, MILK
The cat eagerly licked the last drop of
He went out to buy sausages and
CRIME, THEFT
The police were worried about the rate of
She had seen his name in a book about
CROOK, SWINDLER
He seems honest but he's really a
She didn't think that he acted like a
CROWD, RUSH
The little boy got lost in the
There was a lot of noise in the
CROWN, LAND
The prince became king when he got his
At the end of the game they gave him a
CRUMB, SCRAP
They ate everything down to the last
You can't do very much with just a
DANCE, WALTZ
When the music started he asked her to
She developed a severe headache at the
DITCH, HOLE
The reeling drunk fell into the
It was his job to clean out the
DRESS, SCARF
At the fashion show she bought a red
She didn't have the money to buy a
DRINK, BEER
They went into the tavern to get a
He ran by the spot without wanting a
FEAST, DINNER
They celebrated the harvest with a
They were excited about going to the
FENCE, WALL
The dog chased the yellow cat over the
They didn't know what to do with the
FLEET, SHIP
The new admiral took command of the
He had always enjoyed watching the
FLOOD, POND
After forty days of rain there was a
It all happened six days after the

(*Appendix continues*)

FRAME, GALLERY
She put the new picture into a
She went downtown to see the new
FROST, COLD
The farmer lost his crop to the
In this case you don't expect a
GEESE, SWANS
The pond was alive with ducks and
He spent a pleasant day watching the
GLASS, JUG
They sell wine by the bottle or the
Next to the precious tureen was a
GRAPH, LINE
To explain the equation the teacher drew a
It's easy to see the mark she made on the
GUARD, JUDGE
The prisoner screamed insults at the
He didn't really need to have another
HEDGE, LAWN
The gardener spent the day trimming the
Her husband thought it would be nice to have a
HORSE, ROPE
The cowboy drove the cattle with his
When she retired she bought a nice
HOUSE, FARM
They moved from the apartment into a
They used her inheritance to buy a new
KNIFE, SHOTGUN
The assassin killed his victim with a
The only thing he still requires is a
KNIGHT, CAVALIER
The dragon in the story killed a brave
He thought it must be great fun to be a
MATCH, LAMP
In the darkness of the cellar she lit a
She thought it was a good idea to use a
MOOSE, DEER
Deep in the forest the hunter shot a
It was the first time she had ever seen a
MOUSE, BIRD
The cat waited patiently to catch the
He went into the attic to try to see the
NURSE, SURGEON
The doctor couldn't operate without a
You need a lot of education to be a
PRIEST, TEACHER
The nuns gave their blessing to the old
He recognized the lawyer but not the
PROOF, EVIDENCE
She doubted his story but had no real
She didn't really think she had the
PURSE, BAG
She carried all her money in her
In the accident she lost her new
QUART, PINT
He loves apple juice and drinks it by the
She asked him to stop and pick up another
SCHOOL, CHURCH
The children played in the yard at the
The new building over there will be a
SHARK, LION
The smell of blood excited the
The largest of them all was the
SHEET, BLANKET
She made the bed and replaced the
She needed a new place to store the
SHORE, SEASIDE
They searched for seashells on the sandy
It was after midnight when they reached the
SLEEP, EAT
After working all night he wanted to
He thought it had come to him in his
SKATE, GLOVE
The hockey player adjusted his
His father had taught him to like a
SKETCH, PICTURE
The artist used his pencil to draw a
It's always a pleasure to see a good
SKIRT, BLOUSE
She wore a red sweater and a matching
She asked the clerk to show her a nice
SKULL, VICTIM
The cannibal drank blood from the
She was sure the next one would be a
SKUNK, RAT
Nothing smells worse than a dead
He was always afraid he might meet a
SLEEVE, JACKET
The tailor sewed buttons on the
She held on tight to his woolen
SLEIGH, WAGON
On a snowy day he saw a horse pulling a
In the museum they showed him an old
SLOPE, HILL
He panted heavily as he ran up the
She searched for him on the far side of the
SLUSH, WATER
After the storm the roads were covered with
He ruined his best suit by stepping into a mess of
SMILE, SIGH
When she winked at the baby he responded with a
The famous actress was best known for her
SMOKE, SMOG
The great fire caused clouds of ash and
He couldn't find his way through the
SNAKE, SPIDER
She was terrified of being bitten by a
They had told him earlier about the
SNEEZE, COUGH
He had a cold that made him sniffle and
While eating supper he suddenly had to
SOUND, NOISE
She tiptoed across the room without making a
She was afraid she was about to make a
SPADE, SHOVEL
To plant the tree he dug a deep hole with his
He wasn't very interested in her square
SPICE, GARLIC
The French chef cooked with plenty of
The doctor told him he was allergic to
SPONGE, TOWEL
She soaked up the spilled milk with a
She bought some paper and a yellow
SPRING, GARDEN
The flowers and trees all bloomed in the
He always felt more comfortable in the
SQUARE, CENTER
The festival took place in the village
To help explain the idea he drew a
STAIN, SMUDGE
She took her shirt to the cleaners to remove the
White carpet is very elegant but you risk a

STAFF, CABINET
The president hired a new secretary to complete his
She had a secret unknown to anyone else in the
STAIR, TREAD
Coming down to breakfast she tripped on the
He had a novel idea about how to make a good
STAKE, POST
He planted the little tree and supported it with a
While cutting the lawn she ran the mower over a
STEAM, SIMMER
The hot chicken soup began to bubble and
She couldn't remember whether she had seen it
STEEL, WOOD
The bridge was built with beams of iron and
He bought shares in a company that makes
STONE, MARBLE
The sculptor carved a beautiful statue out of
The child sat on the lawn and picked up a
STORE, SHOP
She settled for a job as a cashier in a
She decided to invest her money in a big
STORM, GALE
The electricity went off during the violent
She had to delay her appointment because of the
STOVE, BURNER
She cooked spaghetti and meatballs on the
For his birthday she bought a shiny new
STRAW, HAY
The floor of the stable was covered with
At one time he dominated the market in
STREET, ROAD
He rode his bike up and down the
He met the love of his life in the
STRING, TWINE
She tied the parcel together with
She had saved up miles of red
STUMP, LOG
The logger cut down the tree and sat on the
To finish the garden they had to remove the
SUITE, ROOM
At the hotel she stayed in the most expensive
The business conference took place in the
SWAMP, MARSH
There were snakes and alligators in the gloomy
Her neighbour's backyard is as messy as a
SWEAT, GRIME
The coal miners were covered with dust and
After waiting for so long he was covered with
SWORD, CLEAVER
The executioner cut off his head with a
In the antique shop he found a rusty old
THIEF, ROBBER
After the robbery the policeman chased the
He usually tried not to think like a
THREAD, NEEDLES
She sewed the wedding dress with silver
She had to quit working when she ran out of
THROAT, MOUTH
He had a bad headache and a very sore
That kind of medicine is good for the
THUMB, FINGER
When the scissors broke she cut her
He took the tape and wrapped it around her
TOAST, BREAD
For breakfast he likes honey and jam on his
She thought it was a good time to make fresh
TONGUE, LIP
When the car hit a bump he bit his
His research was all on the cells of the
TOOTH, MOLAR
He bit a hard piece of popcorn and broke a
She thought it was a lot of money to fix a
TRACK, LIST
The gray horse is the fastest on the
He never thought about the state of the
TRAIL, PATH
The hikers got lost when they left the
He helped the company build a new
TRAIN, BUS
The conductor took their tickets on the
There are few things as relaxing as the
TRASH, GARBAGE
He threw the orange peel into the
She thought the TV series was complete
TRIBE, NATION
The old Indian was the head of his
He didn't recognize her as one of his
TRICK, MANEUVER
The magician showed them a neat
She was afraid to admit that it was a
TRUNK, HOOD
He locked the car doors and opened the
She often thought about the condition of her
VERSE, LINE
The poet got stuck writing the second
Later that evening she told him the second
WAIST, NOSE
Her hips are tiny and so is her
He took precautions to protect his
WHALE, BEACHBALL
If he gets any fatter he'll look like a
The last thing he expected on that day was a
WHEAT, GRAIN
That summer the farmer harvested tons of
After four years of study he was an expert on
WITCH, SORCERESS
She told the children a scary story about an old
They were sure the old lady had once been a
WORLD, GLOBE
The explorer set out on a voyage around the
She wrote him a letter telling about the
WOUND, SCRATCH
He escaped from the plane crash without a
Many people go through life without a
WRECK, WRITE-OFF
The car that hit the telephone pole was a total
In his business it was often necessary to buy a
YACHT, SLOOP
They sailed across the Pacific in a small
The old man couldn't find a buyer for his

Received June 15, 1999
Revision received August 10, 2000
Accepted August 10, 2000 ■