Information-Based and Experience-Based Metacognitive Judgments:
Evidence from Subjective Confidence

Asher Koriat, Ravit Nussinson, Herbert Bless, and Nira Shaked

Introduction

Dual-process theories have been very influential in social psychology and cognitive psychology. These theories postulate a distinction between two modes of thought that underlie judgment and behavior (see Chaiken & Trope, 1999; Kahneman & Frederick, 2005). Different labels have been proposed to describe the two modes (see Koriat, Bjork, Sheffer, & Bar, 2004): nonanalytic versus analytic (Jacoby & Brooks, 1984), associative versus rule based (Sloman, 1996), impulsive versus reflective (Strack & Deutsch, 2004), experiential versus rational (Epstein & Pacini, 1999), and heuristic versus systematic (Chaiken, Liberman, & Eagly, 1989; Johnson, Hashtroudi, & Lindsay, 1993). Although each of these labels emphasizes different aspects of the distinction, there is a general agreement that one mode of thought is fast, automatic, effortless, and implicit, whereas the other is slow, deliberate, effortful, and consciously monitored. Several researchers preferred to use the labels proposed by Stanovich and West (2000), System 1 versus System 2, which are more neutral.

A similar dual-process framework has been proposed for the analysis of metacognitive monitoring, focusing on the question of how people know that they know. The distinction is between experience-based (EB) and information-based (IB) metacognitive judgments (Koriat, 2007; Koriat & Levy-Sadot, 1999; Strack, 1992). The conceptualization of this distinction brings to the fore specific features that may have some bearing for dual-process views in general. In the rest of the introduction, we first describe this distinction and then illustrate how it was applied in research on judgments of learning (JOLs) and feelings of knowing (FOKs). In the experimental part of the chapter, we show how reliance on experience-driven and information-driven processes can yield diametrically opposed effects.

Information-Based and Experience-Based Processes in Metacognition

What is the basis of metacognitive judgments? Assuming that these judgments are inferential in nature, what are the cues on which they are based? Cue utilization views assume a distinction between two possible bases of metacognitive judgments. On the one hand, such judgments may be based on a deliberate use of beliefs and

memories to reach an educated guess about one's competence and cognitions. On the other hand, they may rely on the automatic application of heuristics that take advantage of various mnemonic cues and result in a sheer subjective feeling. Possibly, both processes may contribute in each case to metacognitive judgments, sometimes operating in collaboration and sometimes acting in opposition (see Kelley & Jacoby, 2000). However, for the sake of exposition, we sharpen the distinction between them as if they represent alternative cognitive processes.

Let us consider IB (or theory-based) judgments first. Clearly, judgments about one's knowledge and competence may be based on similar processes as those underlying many judgments and predictions that people make in everyday life. Thus, when students are asked to judge how well they have done on an exam, their judgments may be based on such data as their preconceived notions about their competence in the domain tested, the amount of time they had spent studying for the exam, their assessment of the difficulty of the exam, and so on. For example, Dunning, Johnson, Ehrlinger, and Kruger (2003) found such retrospective assessments to greatly overestimate performance, partly because people tend to base their assessments on their preconceived, inflated beliefs about their skills rather than on their specific experience with taking the test. Also, retrospective assessments of one's performance in a test have been found to depend on people's beliefs about what the test measures, irrespective of their actual performance on that test (Ehrlinger & Dunning, 2003). The study of "metacognitive knowledge" has figured prominently among developmental psychologists: Children's beliefs about their own memory capacities and limitations, and about the factors that affect memory performance have been found to affect both learning strategies and recall predictions (A. L. Brown, 1987; Flavell, 1999; Schneider & Pressley, 1997).

The FOK judgments may also be based on deliberate inferences from one's own beliefs and knowledge. Consider a person who fails to retrieve the answer to a question and is then asked to assess how likely he or she is to "know" the answer to the extent of being able to choose it among distracters. The person may base this assessment on such beliefs as how much expertise he or she has on the topic, whether he or she recalls having used that information in the past, and so on. In that case, the assessment has the quality of an educated guess, and the person may prefer to phrase his or her judgment as "I ought to know the answer" rather than "I feel that I know the answer" (see Costermans, Lories, & Ansay, 1992).

The EB judgments, in contrast, actually involve a two-stage process (Koriat, 2000), first a process that gives rise to a sheer subjective feeling and second a process that uses that feeling as a basis for memory predictions. Thus, when the person in the previous example searches his or her memory for a solicited target, the person may have the experience of directly detecting the presence of the target, as occurs in the tip-of-the-tongue (TOT) state (see R. Brown & McNeill, 1966). The person may even sense that recall is imminent and may experience frustration for failing to retrieve the elusive target. These feelings may serve as the basis for the reported FOK judgments.

What is the process that gives rise to such metacognitive feelings? It has been proposed that metacognitive feelings are formed on the basis of mnemonic cues that give rise directly to these feelings. For example, JOLs made during study have been assumed to rely on the ease with which to-be-remembered items are encoded

or retrieved during learning (Benjamin & Bjork, 1996; Dunlosky & Nelson, 1992; Koriat & Ma'ayan, 2005). Indeed, Hertzog, Dunlosky, Robinson, and Kidder (2003) found JOLs to increase with the success and speed of forming an interactive image between the cue and the target during paired-associate learning. Benjamin, Bjork, and Schwartz (1998) had participants answer general information questions and predict the likelihood of recalling their answers at a later free-recall test. Recall predictions were found to correlate positively with the speed of retrieving an answer, although actual recall exhibited the opposite effect. Also, when participants studied paired associates under self-paced instructions, JOLs were found to decrease with the amount of time invested in the study of each item. These results suggest that learners' JOLs are based on a memorizing effort heuristic that easily learned items are more likely to be remembered than items that require more effort to learn (Koriat, Ma'ayan, & Nussinson, 2006). This heuristic has been found to have some degree of validity because ease of learning is generally diagnostic of recall likelihood (Koriat, in press).

The EB FOK judgments have been assumed to rely on such mnemonic cues as the familiarity of the pointer that serves to probe memory (Metcalfe, Schwartz, & Joaquim, 1993; Reder & Ritter, 1992; Reder & Schunn, 1996) and on the accessibility of pertinent partial information about the solicited memory target (Dunlosky & Nelson, 1992; Koriat, 1993). Indeed, advance priming of the terms of a question (assumed to increase the familiarity of the question) was found to enhance speeded FOK judgments without correspondingly raising the probability of recall or recognition of the answer (Reder, 1988; B. L. Schwartz & Metcalfe, 1992). Other studies support the view that FOK judgments are influenced by the overall accessibility of pertinent information regarding the solicited target (Koriat, 1993; Koriat & Levy-Sadot, 2001). The assumption is that even when recall fails, people may still access a variety of partial clues about the target, and these partial clues may produce the feeling that the target is stored in memory and will be recalled or recognized in the future.

Basic Differences Between Experience-Based and Information-Based Judgments

The foregoing brief review illustrates some of the basic differences between IB and EB metacognitive judgments. The first difference concerns the nature of the cues that are used as the basis of these judgments. IB judgments draw on the *declarative content* of domain-specific beliefs that are retrieved from long-term memory (e.g., "memory declines over time," "I am not very good in geography"). In contrast, EB judgments rely on mnemonic cues that are devoid of declarative content. These cues derive from the very experience of learning, remembering, and deciding rather than from the content of thought. Hence, such cues as the fluency with which information is encoded or retrieved have been referred to as "structural" or "contentless" cues (Koriat & Levy-Sadot, 1999) because they relate to the very quality of processing, that is, to the feedback that one obtains online from one's own processing and performance.

The second difference concerns the quality of the underlying process. In the case of IB judgments, the inferential process is an *explicit*, deliberate process that yields an educated, reasoned assessment. In the case of EB judgments, in contrast, the process

that gives rise to a subjective feeling is *implicit* and largely unconscious: Various mnemonic cues act en masse to give rise to a sheer intuitive feeling.

Third, the process that gives rise to IB judgments is a *dedicated* process that is initiated and compiled ad hoc with the goal of producing a metacognitive judgment. In contrast, EB metacognitive judgments are *by-products* of the ordinary processes of learning, remembering, and thinking. Thus, when learners study a new item of information, their immediate intention is normally to master that item rather than to monitor the degree with which it is studied. However, when attempting to study the item, they also detect its encoding fluency, which then gives rise to the feeling of mastery (Koriat, Ma'ayan, et al., 2006). In a similar manner, when people attempt to retrieve an item from memory, their normal intention is that of remembering rather than of judging its ease of access. However, when retrieval fails, the accessibility of partial clues about the elusive item can serve to support FOK judgments (Koriat, 1993). Thus, the processes that give rise to EB judgments can be said to be *parasitic* on the normal cognitive operations and to arise as a fringe benefit from the performance of these operations.

Finally, the accuracy of IB judgments depends on the validity of the beliefs on which they rest. Inflated beliefs about one's competence may lead to unwarranted overconfidence (Metcalfe, 1998). The accuracy of EB judgments, in contrast, depends on the validity of the mnemonic cues utilized. Indeed, in paired-associate learning, delayed JOLs, when cued by the stimulus term, tend to be markedly more accurate in predicting recall than immediate JOLs (Dunlosky & Nelson, 1992; Nelson & Dunlosky, 1991). Presumably, in making delayed JOLs, learners rely heavily on the accessibility of the target, which is an effective predictor of subsequent recall (Nelson, Narens, & Dunlosky, 2004). When JOLs are solicited immediately after study, the target is practically always retrievable, and hence its accessibility has little diagnostic value.

The Distinction Between Information-Based and Experience-Based Judgments in Previous Research

We cite here only a couple of studies to illustrate the usefulness of the distinction between IB and EB metacognitive judgments. Several studies examined the question of how people know that they do not know the answer to a question. The results of Glucksberg and McCloskey (1981; see also Klin, Guzman, & Levine, 1997) suggest that lack of familiarity with the question normally serves as a basis for an EB "don't know" response. When participants were told in an earlier phase of the experiment that the answer to particular questions is not known, this was found to *increase* the latency of a don't know response to these questions when presented later, possibly because now the response tended to be based on information rather than on sheer subjective experience. Presumably, EB judgments are made faster and more automatically than IB judgments.

The remaining examples concern JOLs made during study. Koriat and Bjork (2005) examined the illusion of competence that often arises in studying new information. They proposed that this illusion derives in part from the inherent discrepancy between the learning and testing conditions: On a typical memory test, people

are presented with a question and are asked to produce the answer, whereas in the corresponding learning condition both the question and the answer generally appear in conjunction. A failure to discount the answer during learning has the potential of creating a *foresight bias* — an unduly strong feeling of competence. This bias is particularly strong in paired-associate learning when the target (present during study) brings to the fore aspects of the cue that are less apparent when the cue is later presented alone (at test). For example, the pair baby–cradle (in Hebrew) tends to produce inflated JOLs during learning (Koriat & Bjork, 2006a) because the association in the backward direction (cradle–baby) is much stronger than that in the forward direction (baby–cradle): In a word association task, the likelihood of *cradle* eliciting *baby* as the first response is .88, whereas that of *baby* eliciting *cradle* is .00. However, participants estimated that 54% of the people who are presented with the word *baby* would be likely to respond with the word *cradle* as the first word that comes to mind (Koriat, Fiedler, & Bjork, 2006).

Koriat and Bjork (2006b) compared the effectiveness of two procedures in alleviating the foresight bias, a mnemonic-based procedure and a theory-based (or IB) procedure. The mnemonic-based procedure, which involved a repeated presentation of the same list, was based on previous findings suggesting that study–test experience, and particularly test experience, enhances learners' sensitivity to mnemonic cues that are diagnostic of memory performance. The theory-based procedure, in contrast, induced participants to resort to theory-based judgments as a basis for JOLs. Both procedures proved effective in mending the foresight bias. Importantly, however, they yielded differential effects with regard to the transfer of improved monitoring to the study of new items. Only the theory-based procedure exhibited transfer, as reflected in JOLs and self-regulation of study time. Thus, subjective experience can be educated through metacognitive training, but the effect of this training on the accuracy of EB judgments is item specific. In contrast, an effective theory that helps mend IB judgments can ensure generalization to new situations.

Another study that illustrates the importance of distinguishing between EB and IB judgments was based on the idea that EB JOLs should be insensitive to the anticipated retention interval because the processing fluency of an item at the time of encoding should not be affected by when testing is expected (Koriat et al., 2004). Indeed, JOLs were entirely indifferent to the expected retention interval, although actual recall exhibited a typical forgetting function. As a result, participants predicted about a 50% recall after a week, whereas actual recall was less than 20%.

This result is surprising because forgetting is a central part of everyone's naïve beliefs about memory. However, several manipulations that were intended to induce participants to apply their theory about forgetting failed to yield a forgetting curve for JOLs. The only procedures that were successful were when retention interval was manipulated within individuals and when recall predictions were framed in terms of forgetting rather than in terms of remembering. These and other results suggest that participants do not spontaneously apply their theories about memory in making JOLs. Rather, they can access their knowledge about forgetting only when theory-based predictions are solicited and the notion of forgetting is accentuated.

Kornell and Bjork (2006) produced even more dramatic results in comparing subjective and objective learning curves. Participants were presented with one, two,

three, or four study–test cycles of a list of paired associates, and during the initial study cycle they were asked to predict their recall performance on the last test in the series. Although actual recall exhibited the typical learning curve, predicted learning curves were essentially flat. In a second experiment, participants made predictions for each of the tests during the initial study cycle. Despite the within-participant manipulation, predicted learning curves hardly increased with study cycle. These results underscore the idea that learners do not spontaneously apply their theories in making recall predictions.

The few studies described above demonstrate the usefulness of the distinction between IB and EB metacognitive judgments and bring to the fore the critical role that experience-driven processes play in influencing these judgments. Whereas the foregoing discussion focused on JOLs made during learning and on FOK judgments made during remembering, the rest of the chapter applies the distinction between IB-driven and EB-driven processes to the analysis of retrospective subjective confidence. The results are intended to show that the two types of processes may sometimes yield diametrically opposed patterns of results. We conclude with several questions that deserve further research.

Information-Based and Experience-Based Confidence Judgments

In the experiments to be reported, we examined the distinction between EB and IB metacognitive judgments with regard to subjective confidence. Some discussions assume that confidence in the answer to a general information question is based on the weight of the evidence that is marshaled in favor of that answer relative to the evidence in support of the alternative answers (e.g., Griffin & Tversky, 1992; Koriat, Lichtenstein, & Fischhoff, 1980; McKenzie, 1997; Yates, Lee, Sieck, Choi, & Price, 2002). These discussions would seem to stress information-driven processes. Other discussions, in contrast, focus on experience-driven processes, emphasizing the contribution of mnemonic cues such as the ease with which the answer is retrieved or selected (Nelson & Narens, 1990). Indeed, confidence in an answer has been found to increase with the speed of reaching that answer. Furthermore, response latency has been found to be generally diagnostic of the correctness of the answer (e.g., Kelley & Lindsay, 1993; Koriat, Ma'ayan, et al., 2006; Robinson, Johnson, & Herndon, 1997).

In the experiments to be reported, we contrast the two hypothesized bases of confidence judgments, borrowing the ease-of-retrieval paradigm introduced by N. Schwarz et al. (1991; see N. Schwarz, 2004, for a review). In that paradigm, participants are required to retrieve many instances or few instances favoring a particular proposition and then make a judgment about that proposition. The requirement to list many instances is assumed to produce a conflict between two potential cues — the content of the information retrieved and the ease of retrieving it: Retrieving many instances provides stronger content-based evidence but is also associated with the experience of greater effort. In a large number of studies, the effects of ease of retrieval on judgment were found to win over the effects of content in affecting judgment (e.g., Aarts & Dijksterhuis, 1999; Haddock, 2002; Wänke & Bless, 2000; Wänke, Bohner, & Jurkowitsch, 1997; Winkielman, Schwarz, & Belli, 1998). For example, participants

who were asked to recall many past episodes demonstrating self-assertiveness later reported lower self-ratings of assertiveness than those who were asked to recall fewer such episodes, presumably because of the greater difficulty experienced in recalling many episodes (N. Schwarz et al., 1991).

In our experiments, we examined the relative contribution of informational content and ease of retrieval to confidence judgments by comparing two conditions that differed in report option: In both conditions, participants answered general knowledge questions by choosing one of two alternative answers. They then listed reasons in support of that answer and finally indicated their confidence in that answer. In the free-report condition, participants listed as many reasons as they could, whereas in the forced-report condition they were asked to provide a specified number of reasons. In the free-report condition, we expected confidence to increase with number of reasons. This is because the strength of the supporting evidence can be assumed to increase with number of reasons retrieved and because in the free-report condition, we expect ease of retrieval to increase with the number of reasons listed. This expectation is based on the finding of Koriat (1993) with regard to FOK judgments. Koriat observed that the number of letters that people retrieved (spontaneously) about a memorized target correlated positively with the speed of retrieving the *first* reported letter, and that both number of letters and speed of retrieval contributed to FOK judgments.

In the forced-report condition of our experiments, in contrast, the retrieval of many reasons should be associated with a stronger experience of effort than the retrieval of few reasons. The effects of ease of retrieval are expected to counteract those of the content of the information retrieved to the extent of reversing the relationship between number of reasons and confidence.

Experiment 1

In Experiment 1, each forced-report participant was yoked to a participant in the free-report condition and was required to provide the same number of reasons that the matched free-report participant had provided for each question. Report option was expected to moderate the effects of number of reasons on confidence judgments.

Method

Participants   Eighty 11th- and 12th-grade high school students participated in the experiment as volunteers.

Materials and Procedure   A set of 16 general knowledge questions in Hebrew, each with two alternative answers, was used. The questions covered a wide range of topics (e.g., "How old was Abraham when his son Isaac was born? (a) 100, (b) 75"). All instructions and materials were compiled in booklets, each question appearing at the top of a separate page. Participants were instructed to choose an answer and then list reasons in support of their choice. For the free-report condition, the instruction, "Write down all supporting reasons you can think of," appeared below the question, followed by five slots. For the forced-report condition, participants were asked to

TABLE 1   The Frequency Distribution of Number of Reasons Across All Participants and Questions and the Number of Participants Who Reported Each Number of Reasons for the Free-Report and Forced-Report Conditions (Experiment 1)
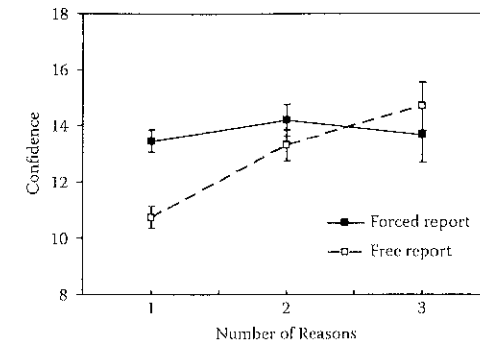
	Free Report					
	Number of Reasons					
	0	1	2	3	4	5
Number of observations	13	375	182	38	6	1
Number of participants	6	40	39	22	4	1

	Forced Report				
	Number of Reasons				
	1	2	3	4	5
Number of observations	388	182	38	6	1
Number of participants	40	39	22	4	1

Figure 1   Mean confidence as a function of number of reasons plotted separately for the forced-report and free-report conditions. Error bars represent + 1 standard error of the mean (SEM) (Experiment 1).

provide for each question the exact number of reasons as their free-yoked participants gave to that question. The instruction was, "Write down X supporting reasons:" and the number of slots differed from one question to another accordingly. For both conditions, a 19-point confidence scale appeared at the bottom of each page, with one end (1) labeled, "There is a very low chance that the answer I chose is correct," and the other (19) labeled, "There is a very high chance that the answer I chose is correct."

There were 13 instances (of 618) in which free-report participants failed to provide any reason. In these cases, the yoked participants were required to give one reason for the respective items.

Results   Table 1 shows the distribution of number of reasons for the free- and forced-report conditions. The distribution is skewed: Free-report participants provided one reason in about 60% of the cases. In only 7% of the cases did participants provide three or more reasons.

Figure 1 presents mean confidence judgments as a function of number of supporting reasons for each of the two conditions. For this figure, we treated three or more reasons as three reasons. A Condition × Number of Reasons analysis of variance (ANOVA) was conducted to evaluate the interaction suggested in this figure, using only 21 participants who provided one, two, and three reasons at least once. Because of the yoking procedure, we treated report option as a repeated factor, so that the effective number of "participants" was 21. The analysis yielded a nonsignificant effect for report option $F(1, 40) = 1.35$, MSE (mean square error) = 16.70, but significant effects for number of reasons, $F(2, 40) = 6.88$, $MSE = 8.87$, $p < .005$, and for the interaction, $F(2, 40) = 5.69$, $MSE = 5.71$, $p < .01$. Separate one-way ANOVAs indicated that confidence increased significantly with number of reasons for the free-report condition (the means were 10.5, 13.4, and 14.5, respectively, for one, two, and three reasons, for the 21 participants), $F(2, 40) = 11.89$, $MSE = 7.53$, $p < .0001$, but not for the forced-report condition, $F < 1$.

Discussion   As expected, report option moderated the effects of number of reasons on confidence. The free-report condition yielded the expected increase in confidence with number of reasons, whereas the forced-report yielded no such increase. The pattern observed for the forced-report condition suggests that the effects of ease-of-retrieval counteracted those of the amount of supporting evidence but failed to reverse this effect. One possible reason for this failure is the yoking procedure used. We found that the questions differed reliably in the number of supportive reasons they elicited: When the free-report participants were divided randomly into two groups, mean number of reasons provided by one group to each question correlated .42 ($p < .11$) across the 16 questions with the number of reasons provided by the other group. Assuming that amount (number of reasons) and ease are correlated positively in the free-report condition (see Koriat, 1993), then the questions for which forced-report participants were required to produce many reasons may not induce a sufficiently strong experience of retrieval effort. If so, the item-by-item yoking feature of Experiment 1 underestimates the effects of ease of retrieval in the forced-report condition. To evaluate this possibility, in Experiment 2 we imposed a predetermined number of reasons on forced-report participants independent of the number of reasons provided by the free-report participants. The number of reasons imposed in the forced-report condition was either 1 or 4. We speculated that perhaps retrieving two or three reasons would not produce a sufficiently strong feeling of difficulty that would reverse the impact of amount of evidence. Indeed, in previous studies that contrasted the effects of amount versus ease, the number of reasons (or statements) imposed in the many-reasons condition was sometimes 10 or more (e.g., Tormala, Petty, & Briñol, 2002; Wänke et al., 1997; Winkielman & Schwarz, 2001).

Experiment 2

In Experiment 2, forced-report participants were required to list 1 reason for 8 of the 16 questions and 4 reasons for the remaining questions. We ran twice as many free-report participants as forced-report participants to obtain a sufficient number of free-report participants who provided both one and four reasons. We hypothesized that if indeed amount and ease correlated positively in the case of the free-report condition, then the positive effect of number of reasons on confidence judgments in this condition should be stronger than the respective negative effect in the forced-report condition.

Method

Participants   Sixty University of Haifa undergraduates (43 women and 17 men) participated in the experiment. Participants were assigned randomly to the 2 conditions with the constraint that there were 40 participants in the free-report condition and 20 in the forced-report condition.

Materials and Procedure   The materials were the same as in Experiment 1. The instructions were similar with two exceptions. First, forced-report participants were asked to list either one or four reasons, with number of reasons alternating between questions, and the assignment of number of reasons to questions was counterbalanced across participants. The order of the questions was the same for all participants. Second, participants were specifically instructed that even when they were uncertain, they should avoid such reasons as "just a guess" or "it seems likely."

Results   For the free-report condition (see Table 2), confidence generally increased with number of reasons. Because the means for each category are based on different participants, we compared confidence judgments for 1 and 2 reasons using only 30 participants who provided both 1 and 2 reasons. The respective means were 10.7 and 13.7, $t(29) = 5.74$, $p < .0001$. There were only 10 participants who provided 1, 2, and 3 reasons (the respective means were 8.5, 11.1, and 13.5), yielding $F(2, 18) = 5.92$, $MSE = 10.62$, $p < .05$.

Turning next to the free-forced comparison, only six participants gave both one and four reasons to some of the questions (see Table 2). Figure 2 (top panel) depicts mean confidence as a function of number of reasons for these participants as well as for the 20 forced-report participants. A two-way ANOVA on these means yielded $F <$

TABLE 2   **Mean Confidence as a Function of Number of Reasons for the Free-Report Option and the Number of Observations and Participants on Which Each Mean Was Based (Experiment 2)**

| | Number of Reasons | | | |
	1	2	3	4
Confidence	10.8	13.7	13.5	18.4
Number of observations	310	139	45	11
Number of participants with nonzero observations	40	30	11	6
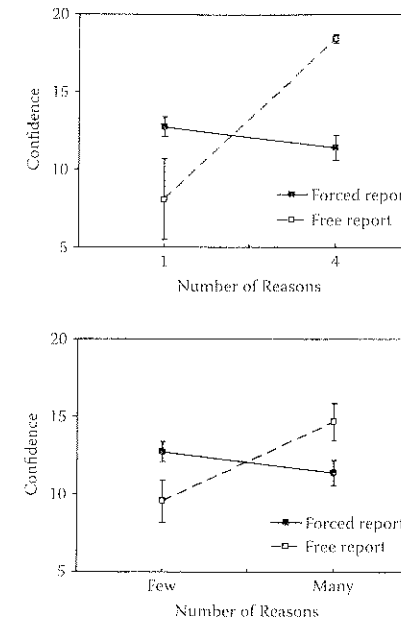
Figure 2   Mean confidence as a function of number of reasons plotted separately for the forced-report and free-report conditions. The free-report means are for participants who gave both 1 and 4 reasons (top panel) and for participants who gave both few (1 or 2) and many (3 or more) reasons (bottom panel). Error bars represent + 1 standard error of the mean (SEM) (Experiment 2).

1 for report option, but number of reasons and the interaction were both significant, $F(1, 24) = 21.07$, $MSE = 8.15$, $p < .0001$, and $F(1, 24) = 38.73$, $MSE = 8.15$, $p < .0001$, respectively. For the free-report condition, confidence increased significantly from one reason to four reasons, $t(5) = 3.63$, $p < .05$, whereas for the forced-report condition, it decreased, $t(19) = 2.16$, $p < .05$.

To ascertain that the results for the free-report participants were not specific to the six participants included in the analysis, we enlarged the sample of free-report participants by combining one and two reasons, treating them as few reasons, and combining three and four reasons, treating them as many reasons. In this manner, we could include 13 free-report participants. Figure 2 (bottom panel) compares the results for these participants with those of the forced-report participants. A two-way ANOVA yielded $F(1, 31) = 0.00$, $MSE = 21.03$, ns (not significant), for report option, but again the effects of number of reasons and the interaction were significant, $F(1, 31) = 6.45$, $MSE = 8.31$, $p < .05$, $F(1, 31) = 19.71$, $MSE = 8.31$, $p < .0001$, respectively.

Here again, confidence increased significantly with number of reasons for the free-report participants, $t(12) = 3.32$, $p < .01$.

Figure 2 also suggests that, indeed, the positive effect of number of reasons on confidence in the free-report condition is stronger than the respective negative effect in the forced-report condition. The mean increase in confidence from one to four reasons in the free-report condition (Figure 2, top panel) was significantly larger than the respective mean decrease in the forced-report condition, $t(24) = 4.79$, $p < .0001$. A similar pattern was observed for the results presented in the bottom panel of Figure 2, $t(31) = 2.59$, $p < .05$.

Discussion   Experiment 2 yielded the expected crossover interaction: Confidence increased significantly with number of reasons under free reporting and decreased significantly under forced reporting. A comparison of these results with those of Experiment 1 supports our suggestion that the extent to which report option moderates the effect of number of reasons on confidence depends on the experienced effort associated with listing many reasons under forced reporting.

The observation that confidence increased more strongly with number of reasons in the free-report condition than it decreased in the forced-report condition is consistent with the idea that whereas amount and ease correlate negatively in the forced-report condition, they correlate positively in the free-report condition. This idea is explored in the next experiment.
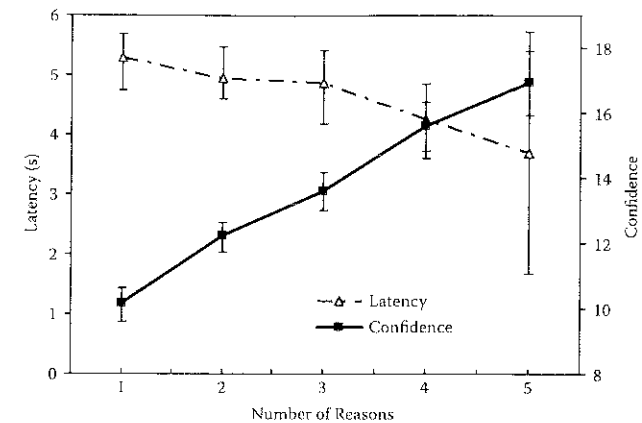
Experiment 3

Experiment 3 attempted to obtain support for the hypothesized positive link between amount and ease in the free-report condition. Participants listed reasons in support of their answer, and the time to initiate report of the *first* reason was measured. We examined whether response latency was indeed shorter when more reasons rather than fewer reasons were produced.

Method   Participants were 60 undergraduates (32 women). The materials and procedure were similar to those of the previous experiments except that the experiment was conducted on a personal computer. On each trial, the question and its two alternative answers appeared on the screen. Participants chose an answer by clicking on it with the mouse and then typed in as many supporting reasons as they could, one in each of five blank windows. The latency to type in the first reason — the interval between clicking the chosen answer and starting to type in the first reason — was recorded. After typing in reasons, participants rated their confidence on the 19-point scale, which appeared on the screen.

Results   Across all participants and questions, there were 418, 351, 148, 36, and 7 instances in which participants provided 1, 2, 3, 4, and 5 reasons, respectively.

Figure 3 presents mean latency of providing the *first* reason. It can be seen that latency decreased monotonically with number of reasons, yielding a Spearman rank correlation of 1.00, $p < .05$. We compared the means of ease of retrieval for one or two

Figure 3   Mean latency and confidence as a function of number of reasons. Error bars represent + 1 standard error of the mean (SEM) (Experiment 3).

reasons versus three or more reasons. Of 42 participants for whom both means were available, 27 exhibited shorter latencies for the many-reasons than for the few-reasons category, $p < .05$, by a binomial test. These results suggest that reasons are more easily retrieved the more of them are available for free reporting.

As in the previous experiments, confidence increased with the number of reasons provided (Figure 3). The rank order correlation (1.00) between confidence and number of reasons was significant at the .05 level. When the analysis was confined to 1, 2, and 3 reasons, using only 39 participants who provided 1, 2, and 3 reasons, mean confidence judgments were 9.8, 12.2, and 13.5, respectively, $F(2, 76) = 21.49$, $MSE = 6.49$, $p < .0001$.

We also examined whether ease of retrieval affected confidence judgments over and above the effects of number of reasons. This examination could be carried out only for the one-reason category for which there was a sufficient number of observations. Using 53 participants who provided 1 reason for at least 2 questions, confidence for slow (above-median) and fast (below-median) responses averaged 10.0 and 11.1, respectively, $t(52) = 1.93$, $p < .06$. Thus, the trend was in the expected direction: A faster retrieval of reasons was associated with higher confidence ratings even when the number of reasons was held constant.

Discussion   The results of Experiment 3 exhibited two trends that are consistent with our expectations. First, ease of retrieval correlated positively with number of reasons; second, ease of retrieval appeared to enhance confidence even when the number of reasons was held constant. These results suggest that the positive correlation observed in all three experiments between number of reasons and confidence in the free-report condition may reflect the joint effects of amount and ease. This may explain in part why the positive effect of number of reasons on confidence was stronger in Experiment 2 than the respective negative effect in the forced-report condition.

General Discussion

The results of this study are consistent with the distinction between IB and EB metacognitive judgments. These results suggest that confidence judgments are affected conjointly by the content of declarative information retrieved from long-term memory and by the ease or effort with which that information is retrieved. When reasons in support of an answer are retrieved spontaneously, confidence increases with number of reasons, possibly because of the increased supportive evidence as well as the greater ease of retrieval. In contrast, when number of reasons is experimentally imposed, the two cues conflict, and the greater effort required to retrieve many reasons may tip the balance, producing a negative relationship between number of reasons and confidence.

Studies using the ease-of-retrieval paradigm in social cognition (see N. Schwarz, 2004) have stressed the idea that the two cues — amount and ease — exert conflicting effects in the case of forced reporting. We showed that the two cues go hand in hand in the case of free reporting, consistent with Koriat's (1993) observation in the context of FOK judgments.

We should note, however, that in Koriat's accessibility model (Koriat, 1993) both amount and ease are conceived as nonanalytic mnemonic cues (see Kelley & Jacoby, 1996): They were assumed to enhance immediate FOK regardless of the content and accuracy of the information retrieved and regardless of the compatibility between the various pieces of partial clues retrieved. According to Koriat (1998), only when the computation of FOK judgments becomes more deliberate does the content of the information enter into consideration so that additional clues may sometimes reduce rather than enhance FOK judgments (see also Vernon & Usher, 2003). This assumption differs from that underlying the studies of the ease-of-retrieval paradigm, in which "amount" and "content" are used interchangeably to describe the strength of declarative arguments in favor of a particular judgment. This is understandable because in that paradigm participants are induced to selectively access arguments that have a specific valence (e.g., arguments in support of buying a certain car).

Nevertheless, because the accessibility model has been applied to confidence judgments as well (e.g., Brewer, Sampaio, & Barlow, 2005; Swann & Gill, 1998), it is important to inquire whether the sheer number of arguments retrieved might contribute to the immediate sense of confidence independent of the content of these arguments. If confidence is affected by accessibility, then three cues may act collaboratively to enhance confidence in the free-report condition: amount, ease (both as nonanalytic, mnemonic cues that feed into EB judgments), and content (as a cue for analytic, IB confidence judgments). All three cues may also be operative in the forced-report condition, except that now amount and ease would operate in opposite directions. These speculations deserve further investigation.

Concluding Remarks

This chapter reviewed evidence demonstrating the usefulness of the distinction between IB and EB processes. This distinction has been applied to the study of JOLs, FOK, and confidence judgments, but its ramifications extend beyond the realm of

metacognitive judgments. Possibly, the analysis of the distinction between the two processes in metacognition can contribute to the refinement and specification of dual-process theories in general.

In concluding this chapter, we should mention several directions for future research. Throughout this chapter, we treated information-driven and experience-driven processes as if they represent alternative routes to metacognitive judgments. Both processes, however, would seem to operate conjointly, contributing in different degrees to these judgments. The results that we presented on confidence judgments underscore the need to examine the complex interactions that exist between the two processes when they operate in tandem. Future work should examine in greater detail the dynamics of the interaction between these processes as it may vary between different conditions (e.g., free reporting vs. forced reporting) and across time (see Koriat, 1998; Vernon & Usher, 2003).

Research on social cognition suggests several additional directions in which the distinction between IB and EB metacognitive processes may be explored. In reviewing the work on the effects of metacognitive experience on judgments, N. Schwarz (2004) emphasized the point that the effects of metacognitive experiences (e.g., the ease with which ideas come to mind) depend on the naïve theory of mental processes that people use in interpreting these experiences. Indeed, it has been observed that participants can be induced to discount the effects of mnemonic cues by attributing them to irrelevant sources (Jacoby & Whitehouse, 1989; N. Schwarz & Clore, 1983; Strack, 1992). A question of interest is whether this is also true for the effects of mnemonic cues on metacognitive judgments such as JOLs and FOK. Can people be induced to discount the effects of cue familiarity and accessibility on FOK judgments by attributing these effects to a different source? Also, there has been increasing evidence suggesting that the naïve theories underlying the effects of metacognitive experiences are highly malleable to the extent that theories with opposite implications can be successfully induced (Unkelbach, 2006; Winkielman & Schwarz, 2001). Can learners be induced to apply a naïve theory that states that fluently processed items are *less* likely to be remembered than those requiring greater encoding effort (see Koriat, in press)? These are some of the questions that await further research.

Acknowledgment

References

Aarts, H., & Dijksterhuis, A. (1999). How often did I do it? Experienced ease of retrieval and frequency estimates of past behavior. *Acta Psychologica, 103,* 77–89.

Benjamin, A. S., & Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. Reder (Ed.), *Implicit memory and metacognition* (pp. 309–338). Hillsdale, NJ: Erlbaum.

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

Brewer, W. F., Sampaio, C., & Barlow, M. R. (2005). Confidence and accuracy in the recall of deceptive and nondeceptive sentences. *Journal of Memory and Language, 52,* 618–627.

Brown, A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 95–116). Hillsdale, NJ: Erlbaum.

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

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology.* New York: Guilford Press.

Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 142–150.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20,* 374–380.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12,* 83–87.

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *Journal of Personality and Social Psychology, 84,* 5–17.

Epstein, S., & Pacini, R. (1999). Some basic issues regarding dual-process theories from the perspective of cognitive-experiential self-theory. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 462–482). New York: Guilford Press.

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology, 50,* 21–45.

Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 311–325.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology, 24,* 411–435.

Haddock, G. (2002). It's easy to like or dislike Tony Blair: Accessibility experiences and the favourability of attitude judgments. *British Journal of Psychology, 93,* 257–267.

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 22–34.

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 1–47). New York: Academic Press.

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., & Frederick, S. (2005). A model of heuristic judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267–293). Cambridge, UK: Cambridge University Press.

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language, 35,* 157–175.

Kelley, C. M., & Jacoby, L. L. (2000). Recollection and familiarity: Process dissociation. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 215–228). Oxford, UK: Oxford University Press.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32,* 1–24.

Klin, C. M., Guzman, A. E., & Levine, W. H. (1997). Knowing that you don't know: Metamemory and discourse processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1378–1393.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling-of-knowing. *Psychological Review, 100,* 609–639.

Koriat, A. (1998). Metamemory: The feeling of knowing and its vagaries. In M. Sabourin, F. Craik, & M. Robert (Eds.), *Advances in psychological science* (Vol. 2, pp. 461–469). Hove, UK: Psychology Press.

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition, 9,* 149–171.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–325). Cambridge, UK: Cambridge University Press.

Koriat, A. (in press). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition.*

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31,* 187–194.

Koriat, A., & Bjork, R. A. (2006a). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition, 34,* 959–972.

Koriat, A., & Bjork, R. A. (2006b). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32,* 1133–1145.

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133,* 643–656.

Koriat, A., Fiedler, K., & Bjork, R. A. (2006). Inflation of conditional predictions. *Journal of Experimental Psychology: General, 135,* 429–447.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 483–502). New York: Guilford Press.

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and the accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27,* 34–53.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 107–118.

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language, 52,* 478–492.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135,* 36–69.

Kornell, N., & Bjork, R. A. (2006). *Objective and subjective learning curves.* Talk presented at the 47th annual meeting of the Psychonomic Society, November 2006, Houston, TX.

McKenzie, C. R. M. (1997). Underweighting alternatives and overconfidence. *Organizational Behavior and Human Decision Processes, 71,* 141–160.

Metcalfe, J. (1998). Cognitive optimism: Self-deception or memory-based processing heuristics? *Personality and Social Psychology Review, 2,* 100–110.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 851–864.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science, 2,* 267–270.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125–123). San Diego, CA: Academic Press.

Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: pre-judgment recall and monitoring (PRAM). *Psychological Methods, 9,* 53–69.

Reder, L. M. (1988). Strategic control of retrieval strategies. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 227–259). San Diego, CA: Academic Press.

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 435–451.

Reder, L. M., & Schunn, C. D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 45–77). Mahwah, NJ: Erlbaum.

Robinson, M. D., Johnson, J. T., & Herndon, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *Journal of Applied Psychology, 82,* 416–425.

Schneider, W., & Pressley, M. (1997). *Memory development between 2 and 20* (2nd ed.). Mahwah, NJ: Erlbaum.

Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1074–1083.

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology, 14,* 332–348.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61,* 195–202.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45,* 513–523.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3–22.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences, 23,* 645–665.

Strack, F. (1992). The different routes to social judgments: Experimental versus informational strategies. In I. I. Martin & A. Tesser (Eds.), *The constructions of social judgments* (pp. 249–275). Hillsdale, NJ: Erlbaum.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247.

Swann, W. B., Jr., & Gill, M. (1998). Beliefs, confidence, and the Widows Ademosky: On knowing what we know about others? In V.Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Cognitive and social dimensions* (pp. 107–125). London: Sage.

Tormala, Z. L., Petty, R. E., & Briñol, P. (2002). Ease of retrieval effects in persuasion: A self-validation analysis. *Personality and Social Psychology Bulletin, 28,* 1700–1712.

Unkelbach, C. (2006). The learned interpretation of cognitive fluency. *Psychological Science, 17,* 339–345.

Vernon, D., & Usher, M., (2003). Dynamics of metacognitive judgments: Pre- and postretrieval mechanisms. *Experimental Psychology: Learning, Memory, and Cognition, 29,* 339–346.

Wänke, M., & Bless, H. (2000). The effects of subjective ease of retrieval on attitudinal judgments: The moderating role of processing motivation. In H. Bless & J. P. Forgas, (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 143–161). Philadelphia: Psychology Press.

Wänke, M., Bohner, G., & Jurkowitsch, A. (1997). There are many reasons to drive a BMW: Does imagined ease of argument generation influence attitudes? *Journal of Consumer Research, 24,* 170–177.

Winkielman, P., & Schwarz, N. (2001). How pleasant was your childhood? Beliefs about memory shape inferences from experienced difficulty of recall. *Psychological Science, 12,* 176–179.

Winkielman, P., Schwarz, N., & Belli, R. F. (1998). The role of ease of retrieval and attribution in memory judgments: Judging your memory as worse despite recalling more events. *Psychological Science, 9,* 124–126.

Yates, J. F., Lee, J. W., Sieck, W. R., Choi, I., & Price, P. C. (2002). Probability judgment across cultures. In T. Gilovich & D. Griffin (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 271–291). New York: Cambridge University Press.