Developmental Science

Developmental Science 13:3 (2010), pp 441–453 DOI: 10.1111/j.1467-7687.2009.00907.x

PAPER

Choice latency as a cue for children's subjective confidence in the correctness of their answers

Asher Koriat and Rakefet Ackerman

Department of Psychology, University of Haifa, Haifa, Israel

Abstract

Research with adults indicates that confidence in the correctness of an answer decreases as a function of the amount of time it takes to reach that answer, suggesting that people use response latency as a mnemonic cue for subjective confidence. Experiment 1 extended investigation to 2nd, 3rd and 5th graders. When children chose the answer to general knowledge questions, their confidence in the answer was inversely related to choice latency. However, the strength of the relationship increased with grade, suggesting increased reliance with age on the feedback from task performance. The validity of latency as a cue for the accuracy of the answer also increased with age, possibly contributing to the observed age increase in the extent to which confidence judgment discriminated between correct and wrong answers. Whereas these results illustrate the dependence of metacognitive monitoring on the feedback from control operations, Experiments 2 and 3 examined the idea that control-based monitoring affects subsequent control operations. When children were free to choose which answers to volunteer under a payoff schedule that emphasized accuracy, they tended to volunteer high-confidence answers more than low-confidence answers (Experiment 2) and more short-latency answers than long-latency answers (Experiment 3). The latter tendency was again stronger for older than for younger children. The results are discussed in terms of the intricate relationships between monitoring and control processes.

Introduction

A central question in metacognition concerns the bases of metacognitive judgments: how do people monitor their knowledge and competence during learning and remembering? This question has been addressed with regard to judgments of learning (JOL) made during the study of new material (Benjamin & Bjork, 1996; Koriat, 1997), with regard to feeling of knowing (FOK) judgments made during the retrieval of information from memory (Koriat, 1993; Metcalfe, Schwartz & Joaquim, 1993; Reder & Ritter, 1992), and with regard to confidence judgments in the correctness of one's answers (Kelley & Lindsay, 1993). Whereas early conceptualizations endorsed a direct-access view according to which metacognitive judgments are based on detecting the presence and/or the strength of memory traces (e.g. Cohen, Sandler & Keglevich, 1991; Hart, 1965; see Schwartz, 1994), much recent work assumes that such judgments are inferential in nature: they are based on cues and heuristics that have a certain degree of validity in predicting memory performance.

A distinction has been generally drawn within cue-utilization approaches between information-based and experience-based metacognitive judgments (Koriat & Levy-Sadot, 1999; Koriat, Nussinson, Bless & Shaked, 2008). Information-based judgments rely on an analytic, deliberate inference, in which various considerations retrieved from memory are consulted and weighed to reach an educated judgment (Koriat, Lichtenstein & Fischhoff, 1980; Yates, Lee, Sieck, Choi & Price, 2002). These judgments have the quality of reasoned assessments. Experience-based judgments, in contrast, rest on an unconscious and automatic inference from a variety of mnemonic cues that reside in the immediate feedback from task performance. These cues, such as fluency or ease of access, give rise directly to a sheer feeling of knowing, which is experienced as an intuitive hunch rather than as an educated inference (see Koriat, 2000).

Traditionally, developmental research on metacognition has placed a greater emphasis on information-driven processes than on mnemonic-based processes (see Flavell, 1999; Schneider & Bjorklund, 1998; Schneider & Lockl, 2008; Schneider & Pressley, 1997). Much of that work is predicated on the assumption that children's beliefs about their own memory capacities and limitations and about the factors that contribute to memory performance affect both their assessment of their competence (monitoring) and their learning strategies (control; Lockl & Schneider, 2002; Schneider, 2008). Thus, research on declarative metacognition examined the development of beliefs such as the belief that memory decays over time (Lyon & Flavell, 1993), that related paired associates are easier to remember than unrelated pairs (Dufresne & Kobasigawa, 1989; O'Sullivan, 1997), that

Address for correspondence: Asher Koriat, Department of Psychology, University of Haifa, Mount Carmel, Haifa, Israel, 31905; e-mail: akoriat@research.haifa.ac.il

memory for studied materials increases with the amount of effort invested in study (O'Sullivan, 1993), and so on. The assumption, then, is that declarative metacognitive knowledge affects both metacognitive monitoring and metacognitive control.

Researchers of adult metacognition, in contrast, have placed a greater emphasis on experience-based judgments (see Koriat, 2007). Although they acknowledge the influence of declarative knowledge (e.g. Koriat, 1997; Ehrlinger & Dunning, 2003; Mazzoni, Loftus & Kirsch, 2001; Odegard & Lampinen, 2006), they stress the idea that metacognitive judgments rest primarily on the immediate feedback that people gain when performing cognitive operations. Indeed, many results on adult metacognition support the notion of experience-based judgments that are based on mnemonic cues. Thus, evidence suggests that JOLs are based on the fluency with which the to-be-remembered items are encoded or retrieved during learning (Benjamin & Bjork, 1996; Benjamin, Bjork & Schwartz, 1998; Koriat & Ma'ayan, 2005; Matvey, Dunlosky & Guttentag, 2001). For example, when participants studied paired associates under self-paced instructions, their JOLs were found to decrease with the amount of time invested in the study of each item, suggesting that JOLs are influenced by the ease with which items are committed to memory (Koriat, Ma'ayan & Nussinson, 2006).

FOK judgments, too, appear to rest on mnemonic cues that are accessed on-line during remembering. Results suggest that they are enhanced by manipulations that increase the familiarity of the question that is used to probe memory (Metcalfe *et al.*, 1993; Reder & Ritter, 1992; Reder & Schunn,1996). Other results indicate that FOK judgments are also affected by the amount of partial clues that come to mind during the search for the memory target, and by the ease with which these clues come to mind (Koriat, 1993, 1995; Schwartz, 2002).

It would seem somewhat odd that research on adult metacognition has placed a greater emphasis on the contribution of the immediate feedback from task performance, whereas research on metacognition in children has traditionally stressed the contribution of declarative knowledge retrieved from long-term memory. In fact, recent findings suggest that even college students hardly apply their declarative knowledge and theories in making metacognitive judgments (Koriat, Bjork, Sheffer & Bar, 2004; Kornell & Bjork, 2006). These findings indicate that college students do not apply spontaneously some of the most basic theories about learning and remembering in making recall predictions, and do so only under specific conditions (see Finn, 2008).

Nevertheless, recent developmental work on procedural metacognition has paid greater attention to the contribution of experiential, mnemonic cues to metacognitive judgments. For example, Lockl and Schneider (2002) examined children's FOK judgments and found evidence suggesting that children, like adults, rely on the accessibility heuristic (Koriat, 1993): they base their FOK judgments on the amount of information that can be accessed regardless of its correctness. Similarly, results by Roebers, von der Linden, Schneider and Howie (2007) suggest that children's JOLs also rely in part on the amount of information that comes to mind and its ease of access. More relevant to the present study is the recent work of Koriat, Ackerman, Lockl and Schneider (2009b), which examined the development of the memorizing-effort heuristic as a basis of JOLs. When 3rd- to 6th-grade children studied a list of paired associates under self-paced instructions, their JOLs decreased with the amount of time they invested in studying each item, similar to what had been found for young adults (Koriat *et al.*, 2006). The results for 1st-grade and 2nd-grade children, in contrast, did not yield this relationship, suggesting that young children do not use the feedback from study experience as a cue for recall predictions (see also Koriat, Ackerman, Lockl & Schneider, 2009a).

The present study extends this line of developmental research to confidence judgments in the correctness of one's answers to general knowledge questions. Numerous studies with adults have indicated that confidence judgments are generally correlated with the accuracy of the answer. In attempting to explain the confidence–accuracy correlation, it has been proposed that confidence judgments are based on the ease with which an answer is retrieved or selected and that response latency is generally diagnostic of the accuracy of the answer. Indeed, several studies have documented an inverse relationship between the time it takes to retrieve or select an answer and the confidence expressed in that answer (Koriat *et al.*, 2006; Nelson & Narens, 1990; Zakay & Tuvia, 1998). Kelley and Lindsay (1993) found that priming participants with answers to general knowledge questions reduced the latency with which the primed answers were recalled, and also enhanced confidence in these answers. In parallel, it was found that response latency is also inversely related to accuracy (e.g. Koriat, 2008b; Robinson, Johnson & Herndon, 1997). These results are consistent with the view that confidence judgments are based on the internal feedback that participants gain from attempting to select or retrieve an answer, and that this feedback is usually diagnostic of the correctness of the answer.

The experiments to be reported below capitalized on these studies, extending investigation to 2nd, 3rd and 5th graders. As noted earlier, a previous study indicated a development in this age range in the degree to which children relied on the mnemonic cue of memorizing effort as a basis of JOLs (Koriat *et al.*, 2009b). In Experiment 1 of the present study we examined the extent to which children rely on response latency as a basis of their confidence judgments in their answers. In Experiments 2 and 3, in turn, we examined one possible consequence of subjective confidence – the tendency to commit oneself to the answer when accuracy is at stake.

Figure 1 helps to clarify the theoretical framework underlying the experiments to be presented. These experiments may be seen to reconcile two contrasting metatheoretical positions regarding the causal relationship between metacognitive monitoring and metacognitive control (see Koriat *et al.*, 2006). Underlying much of the work in metacognition is the Monitoring → Control (MC) model, which assumes that metacognitive judgments (monitoring) drive and guide the strategic regulation of information processing and behaviour (control; Barnes, Nelson, Dunlosky, Mazzoni & Narens, 1999; Nelson & Narens, 1990; Son & Schwartz, 2002). Consistent with this model are findings suggesting that JOLs during self-paced learning affect the allocation of study time (Dunlosky & Hertzog, 1998; Lockl & Schneider, 2003; Nelson & Leonesio, 1988), FOK judgments determine the amount of time spent searching for an elusive memory target before giving up, and confidence judgments in a belief or attitude determine whether people translate that belief or attitude into behaviour (Fischhoff, Slovic & Lichtenstein, 1977; Gill, Swann & Silvera, 1998; Goldsmith & Koriat, 2008; Koriat, 2008c). These findings suggest that metacognitive judgments are not mere epiphenomena but actually influence behaviour.

It has not always been noticed, however, that some of the findings regarding the relationship between monitoring and control imply a Control → Monitoring (CM) model, in which monitoring is assumed to rest on the feedback from control operations and hence must follow rather than precede control. This model is implicit in the theorizing that JOLs are based on study time (Koriat *et al.*, 2009a,b), that FOK judgments are based on the accessibility of partial clues regarding the target (Koriat, 1995; Lockl & Schneider, 2002; Roebers *et al.*, 2007), and that confidence is based on the amount of time it takes to reach an answer or a solution (Kelley & Lindsay, 1993). Thus, Experiment 1 of this study is predicated on the hypothesis that it is by trying to answer a question that one can appreciate the likelihood that the answer is correct (see arrow A in Figure 1).

The question that emerges then is this. What is the function of monitoring if it merely reflects the feedback from control operations? Koriat *et al.* (2006) proposed that, even when monitoring is based on the feedback from control operations, it can affect and drive subsequent control operations (Figure 1, arrow B). They argued that monitoring and control may alternate in a cascaded pattern so that control-based monitoring can guide subsequent control operations, and the feedback from these operations can inform further monitoring, and so on (for an example, see Koriat & Levy-Sadot, 2001). Thus, although confidence judgments may be based on the feedback from the operations involved in retrieving and choosing an answer, they may then affect further control operations.

Whereas Experiment 1 focused on control-based monitoring (arrow A in Figure 1), Experiments 2 and 3 focused on monitoring-based control (arrow B in Figure 1). In these experiments we examined the idea that although confidence judgments are based on the feedback from control operations (response latency) they can affect subsequent strategic choices, specifically, the decision to volunteer or withhold the answer when report accuracy is rewarded. Experiment 2 was conducted on 2nd graders and examined the full chain Latency–Confidence–Volunteering. Experiment 3 examined the relationship between response latency and volunteering behaviour without the possible interference that might ensue from the collection of confidence ratings. This experiment used 2nd and 5th graders to examine age changes in the strength of this relationship.

Experiment 1

In Experiment 1, children of three age groups were presented with forced-choice, two-alternative general knowledge questions, and were asked to choose an answer, and to indicate their confidence in the answer. Choice latency – the time it took to choose the answer – was measured.

The results were analysed in a manner similar to that proposed by Brunswik (1956) for perception, to yield information about cue utilization, cue validity, and achievement. (For a similar analysis, see Koriat, 2008a; Koriat *et al.*, 2006.) With regard to *cue utilization*, we examined whether children's confidence judgments exhibit an inverse relationship with response latency, suggesting reliance on the *choice–latency heuristic*. We also examined whether reliance on this heuristic increases with age, as was found to be the case for the memorizing-effort heuristic underlying JOLs (Koriat *et al.*, 2009b).

As far as *cue validity* is concerned, we examined whether choice latency is indeed a diagnostic cue for the

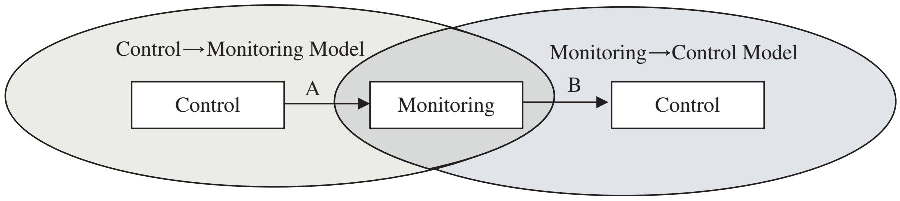

Figure 1 *A conceptual scheme depicting the postulated causal links between metacognitive monitoring and metacognitive control. Experiment 1 examined the relationship marked by arrow A, Experiment 2 examined the entire chain, and Experiment 3 focused on the relationship marked by arrow B.*

correctness of the answer. Previous studies with adults have indicated that by and large faster responses are more likely to be correct than slower responses. The latency–correctness relationship has also been observed in forensic research: the speed with which a witnessed event was reported was found to be diagnostic of the accuracy of the report (Robinson *et al*., 1997; Sporer, 1993; Weber, Brewer, Wells, Semmler & Keast, 2004). Previous studies that extended investigation of this relationship to children yielded somewhat inconsistent results. Whereas Weber *et al*. (2004) found better accuracy for faster than for slower lineup identifications among 11-year-old children, Brewer and Day (2005) failed to find such a relationship among children aged 8–11.

Finally, *achievement* was indexed by the confidence–accuracy correlation: to the extent that confidence judgments rest on choice latency and that choice latency is diagnostic of accuracy, the confidence–accuracy correlation should be positive. Indeed, results obtained with adults suggest that the confidence–accuracy correlation is mediated in part by response latency (Koriat *et al*., 2006; Robinson *et al*., 1997).

Method

Participants

Participants were 60 children from elementary schools in Israel, mostly of middle-class and upper-middle-class socioeconomic background. The sample included 20 2nd graders (mean age 7.9), 20 3rd graders (mean age 9.0), and 20 5th graders (mean age 11.1).

Materials

A 42-item general knowledge test (in Hebrew) was developed, with questions covering a broad range of topics; 40 questions were used for the experiment phase, and two questions were used for the preceding practice phase. Each question was followed by two answers, each consisting of one or two words. The questions were pretested to ensure their appropriateness to the age levels included in the study.

Procedure

The consent of the parents and of the school was obtained before beginning the study. Children were tested individually in a quiet room in the school, using a PC-compatible laptop computer. They were told that they would be asked to answer general knowledge questions displayed on the screen one at a time. For each question, two alternative answers would be shown and they would have to choose the correct answer. To initiate each trial, they were to click with the mouse on a box labelled 'display question', at which time the question was exposed. They were told that as soon as they finished reading the question (and before attempting to figure out the answer) they should click on a box labelled 'answers'. The two alternative answers then appeared on the screen side by side, and children indicated their answer by clicking on a circle below the chosen answer. When the child encountered even a slight difficulty in reading the question, the experimenter read aloud the question and the two answers. Children were allowed to change their choice before clicking a 'continue' box. Choice latency was defined as the interval between the presentation of the alternative answers and the clicking of the 'continue' box. After clicking the 'continue' box, a red frame was added to mark the chosen answer, and the following question appeared. 'How sure are you that the answer is correct?' The measurement of confidence capitalized on the hot–cold game familiar to children, using a thermometer procedure (adapted from a procedure used by Koriat *et al*., 2009b; Koriat & Shitzer-Reichert, 2002). Children made their ratings by sliding a pointer on a coloured scale using the mouse. The position on the scale was transformed into a confidence percentage score (0%–100%). After marking the confidence rating, children clicked on a box labelled 'next question' to initiate the next trial. The first two questions were the practice questions. The order of the remaining 40 questions was randomly determined for each child. Children went through the process at their own pace. All children received a small gift at the conclusion of the experiment.

Results and discussion

Table 1 presents mean choice latency, confidence, and percentage correct for each grade. It can be seen that whereas choice latency decreased with grade, percentage correct increased with grade. A one-way analysis of variance (ANOVA) of the effect of age group on choice latency yielded $F(2, 57) = 34.50$, $MSE = 4556.1$, $p < .0001$. A similar one-way ANOVA on percentage correct yielded $F(2, 57) = 10.13$, $MSE = 64.4$, $p < .0001$.

Cue utilization: choice latency as a cue for confidence

Preliminary analyses on the confidence data yielded little difference between the groups in the use of the confidence scale. For example, the standard deviations of confidence ratings averaged 24.3, 24.3 and 23.7, for 2nd, 3rd and 5th graders, respectively.

Table 1 *Mean response latency, confidence, and percentage correct (standard deviations in parentheses) for children in the various grades (Experiment 1)*

Grade	Choice latency	Confidence	Percentage correct
2nd	10.1[a] (2.5)	79.3[a] (9.3)	73.8[a] (8.9)
3rd	7.9[b] (2.5)	77.1[a] (11.7)	78.8[a] (8.3)
5th	4.5[c] (1.2)	82.2[a] (9.7)	85.2[b] (6.7)

Note: Column means with distinct superscripts are significantly different by Tukey's post-hoc test.

To examine the within-participant relationship between confidence and choice latency, all choice latencies were split at the median for each participant, and average confidence judgments were calculated for below-median items and above-median items, as was done by Koriat *et al.* (2006). The means of these averages are depicted in Figure 2(a) for each grade. It can be seen that for each grade confidence judgments decreased as a function of increasing choice latency, consistent with the idea that confidence is influenced by the time it takes to reach an answer (e.g. Kelley & Lindsay, 1993). However, the slope of the function relating confidence to choice latency increased with grade. A two-way ANOVA, Choice Latency (below median vs. above median) × Grade on confidence yielded $F(1, 57) = 155.10$, $MSE = 67.06$, $p < .0001$, for latency; $F(2, 57) = 1.18$, $MSE = 212.74$, $p = .31$, for grade; and $F(2, 57) = 5.76$, $MSE = 67.06$, $p < .01$, for the interaction. Confidence was significantly higher for below-median than for above-median choice latencies for 2nd graders, $t(19) = 6.43$, $p < .0001$, for 3rd graders, $t(19) = 7.03$, $p < .0001$, and for 5th graders, $t(19) = 8.31$, $p < .0001$.

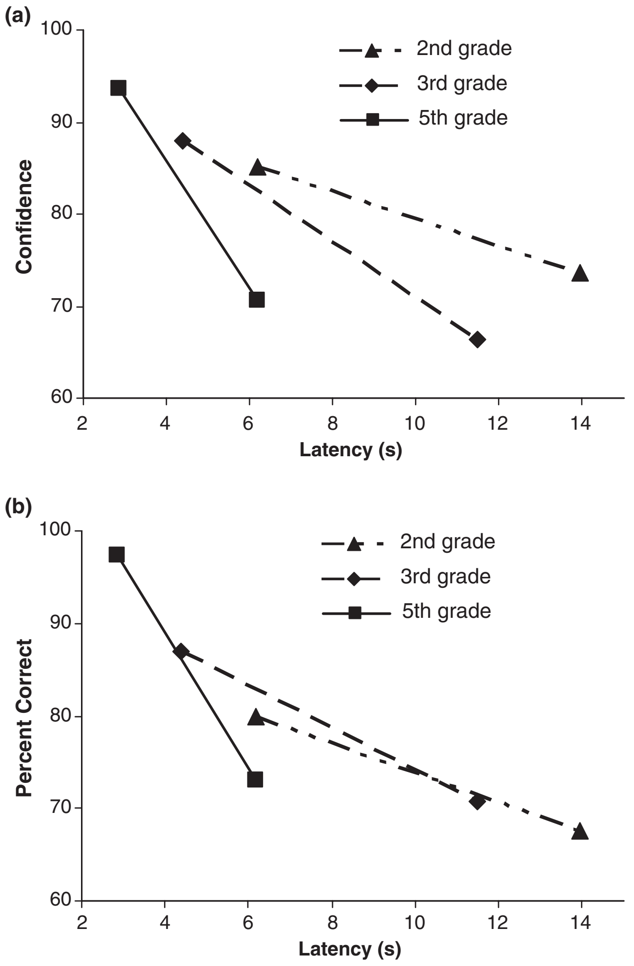

Figure 2 *Mean confidence (a) and mean percentage correct (b) for below-median (short) and above-median (long) latency for each of the three grades in Experiment 1.*

The interaction reflects the observation that the slope of the linear regression relating confidence to choice latency increased with grade. Thus, when confidence (the full range) was regressed on choice latency for each participant, the slope of the regression line averaged –1.54, for 2nd graders; –2.57, for 3rd graders; and –5.97, for 5th graders, all significantly different from zero ($p < .0001$). A one-way ANOVA, however, revealed a significant difference between the groups, $F(2, 57) = 15.58$, $MSE = 6.89$, $p < .0001$. A Tukey post-hoc test suggested that the slope for 5th graders was steeper than the slopes for the 2nd and 3rd graders, which did not differ from each other.

These results were also confirmed by an analysis of the within-participant gamma correlations between choice latency and confidence (taking into account the full range of latencies) for each participant. These correlations averaged −.30, −.43 and −.54 for 2nd, 3rd and 5th graders, respectively. Each of these correlations was significantly different from 0, $p < .0001$. A post-hoc test indicated that the correlation was significantly lower for the 2nd graders than for the two higher grades, which did not differ significantly from each other. Note that the gamma correlation was negative for each and every child except for one 2nd grader.

In conclusion, the results are consistent with the cue-utilization view, suggesting that even 2nd graders base their retrospective confidence judgments on the feedback from task performance, as has been found for adults (Koriat *et al.*, 2006; Robinson *et al.*, 1997). The results, however, also disclose a developmental trend, suggesting an age increase in reliance on the mnemonic cues that derive from control operations (see also Koriat *et al.*, 2009b).

Cue validity: choice latency as a predictor of accuracy

Is choice latency indeed a diagnostic cue for the correctness of the answer? The pertinent results appear in Figure 2(b), which presents mean percentage correct for below-median and above-median choice latencies for each of the grades. A two-way ANOVA on percentage correct, as above, yielded $F(1, 57) = 141.22$, $MSE = 66.24$, $p < .0001$, for latency; $F(2, 57) = 10.13$, $MSE = 128.81$, $p < .0001$, for grade; and $F(2,57) = 5.38$, $MSE = 66.24$, $p < .01$, for the interaction. The effects of latency were significant for each of the grades: $t(19) = 3.66$, $p < .01$, for 2nd graders; $t(19) = 9.34$, $p < .0001$, for 3rd graders; and $t(19) = 10.52$, $p < .0001$, for 5th graders. Thus, for all grades, choice latency was indeed diagnostic of the correctness of the answer. The interaction, however, suggests that the validity of choice latency increased with age: the slope of the function relating percentage correct to choice latency averaged −1.16 for 2nd graders; −1.81 for 3rd graders; and −6.20 for 5th graders, all significantly different from zero ($p < .01$ for 2nd graders, and $p < .0001$ for 3rd and 5th graders). A one-way ANOVA on these slopes, however, yielded F(2,

57) = 12.37, MSE = 12.11, $p < .0001$, for grade, and a post-hoc test indicated a steeper slope for 5th graders than for the 2nd and 3rd graders, which did not differ from each other.

These trends were also confirmed by within-participant gamma correlations between latency and accuracy. As expected, these correlations were negative, averaging −.24 for 2nd graders, −.38 for 3rd graders, and −.65 for 5th graders. Each of these correlations was significantly different from zero, $p < .0001$. A one-way ANOVA indicated a significant difference between the grades, $F(2, 57) = 9.64$, MSE = 0.03, $p < .0001$, and a post-hoc test indicated a significant difference between the 5th graders and the other two groups, which did not differ from each other.

In sum, choice latency appears to be a relatively valid cue for the correctness of the answers, as was found to be the case for adults (Kelley & Lindsay, 1993; Koriat *et al.*, 2006; Nelson & Narens, 1990). The developmental trend observed, however, suggests that the validity of choice latency as cue for accuracy increases with age.

Achievement: the confidence–accuracy relationship

Given that children rely on choice latency as a cue for confidence and that latency is a good predictor of accuracy, we should expect confidence to exhibit some degree of validity in predicting accuracy. Indeed, monitoring accuracy, as indexed by the within-participant confidence–accuracy gamma correlation (Nelson, 1984), averaged .48, .58 and .74 for 2nd, 3rd and 5th graders, respectively. Each of these correlations was significantly greater than zero ($p < .0001$), indicating that the children were successful in discriminating between correct and wrong answers. However, a one-way ANOVA revealed a significant difference between the groups, $F(2,57) = 8.45$, MSE = 0.04, $p = .001$, indicating an age-related increase in the ability to monitor the correctness of one's own answers. A post-hoc test indicated that the confidence–accuracy correlation was significantly higher for 5th graders than for the two younger age groups, which did not differ from each other.

We also assessed the degree to which the age differences in monitoring accuracy were mediated by differential sensitivity to choice latency as a cue for confidence. To do so, we calculated for each participant the Pearson correlation between them both before and after partialling out the effects of choice latency. The raw Pearson correlations averaged .32, .39 and .45, for 2nd, 3rd and 5th graders, respectively. When latency was partialled out, the respective correlations were lower but still significant, averaging .29, .34 and .33. A two-way ANOVA on these correlations, Correlation Type (overall vs. partialled out) × Grade, yielded a main effect for correlation type, $F(1, 57) = 38.63$, MSE = 0.003, $p < .0001$, but the interaction was also significant, $F(2, 57) = 6.87$, MSE = 0.003, $p < .01$. These results suggest that monitoring accuracy was partly accounted for by reliance on choice latency, and that the improved accuracy with age was also largely a result of the age-increase in the use of choice latency as a cue for confidence.

Experiment 2

Whereas Experiment 1 focused on a determinant of children's confidence judgments, Experiments 2 and 3 explored a consequence of these judgments. The methodology used was similar to that developed by Koriat and Goldsmith (1996) to simulate the situation of a person on a witness stand who is sworn to tell the whole truth and nothing but the truth. That methodology assumes that the decision of a witness to volunteer or withhold a piece of information that comes to mind is based heavily on the confidence in that information. In their Experiment 1, Koriat and Goldsmith had participants (college students) answer general knowledge questions under forced-report instructions and indicate their confidence in their answers. When the participants later took the same test under a free-report condition, with the instructions that they would receive a monetary bonus for each correct answer but would pay a penalty for each wrong answer, their tendency to volunteer an answer correlated strongly (over .90) with the confidence associated with that answer. Similar results were obtained with elementary-school children who were asked to report about a witnessed event (Koriat, Goldsmith, Schneider & Nakash-Dura, 2001). These observations accord with the idea that metacognitive judgments guide control processes (Barnes *et al.*, 1999; Metcalfe & Finn, 2008; Nelson & Narens, 1990; Son & Schwartz, 2002; Thiede, Anderson & Therriault, 2003).

Experiment 2 was conducted only with 2nd graders to examine whether evidence for the causal chain depicted in Figure 1 can be detected even in this age group. It traced the full chain from choice latency, through confidence, to volunteering behaviour. Apart from confirming the link between choice latency and confidence, it tested the idea that confidence judgments affect the decision whether to volunteer or withhold the answer. The results were expected to yield evidence for the idea (see Figure 1) that although monitoring (confidence) is based on the feedback from control operation (speed of reaching an answer) it can affect subsequent control operations (volunteering behaviour). To foreshadow, Experiment 3 included both 2nd graders and 5th graders, with the intention to examine age differences in the strategic control of volunteering behaviour.

Method

Participants

Participants were 20 2nd graders (mean age 7.2) drawn from the same population as in Experiment 1.

Materials and procedure

The items were the same as those used in Experiment 1. The experiment was introduced as a game: children were told that they would win one point for each correct answer but would lose one point for each incorrect answer. However, they had the option to withhold an answer, in which case they would neither win nor lose a point. Thus, the children were instructed that for each of the answers they should indicate whether that answer would be taken into account in calculating the total number of points that they would win, and that they should try to maximize their winnings. It was emphasized that they should deliberate whether to risk volunteering an answer or not. The 'game' was illustrated using a set of four questions. It was made sure that the children understood the implications of volunteering all four answers, just some of them, or none.

The procedure was identical to that of Experiment 1, but after marking their confidence on the hot–cold scale, a volunteering question was added. The children were asked to indicate whether they wished the answer to be taken into account in calculating their winnings, and to respond by clicking one of two boxes marked 'Yes' or 'No.' No feedback about the correctness of the answer was given. The entire procedure was demonstrated using two practice questions. Choice latency was measured as in Experiment 1.

Results and discussion

Control-based monitoring

The results that pertain to the CM model largely replicate the corresponding results from Experiment 1. First, confidence judgments were found to increase with decreasing choice latency. Thus, when choice latencies were split at the median for each participant, confidence ratings averaged 88.8 for short-latency answers and 75.2 for long-latency answers, $t(19) = 5.30$, $p < .0001$. Second, as in Experiment 1, choice latency was generally diagnostic of accuracy: percentage correct averaged 79.3 for short-latency answers and 64.8 for long-latency answers, $t(19) = 3.88$, $p = .001$. The diagnostic value of choice latency was expected to contribute to the accuracy of confidence judgments in monitoring the correctness of the answer. Indeed, the gamma correlations between confidence and accuracy averaged .43 ($p < .0001$), consistent with the results of Experiment 1.

Monitoring-based control

We turn next to the results pertaining to the MC model as reflected in the confidence–volunteering relationship. To examine the control function of confidence, we divided confidence ratings at the median of each participant, and compared volunteering rate for below-median and above-median confidence. Overall, volunteering rate was very high, averaging 87.1%. However, it was still higher for above-median confidence (96.4%) than for below-median confidence (75.2%), $t(19) = 3.82$, $p = .001$. It should be noted that 6 children volunteered all of their answers. For the remaining 14 participants, the within-participant gamma correlation between confidence and volunteering averaged .90.

The benefit that ensues from using one's confidence as a basis for volunteering behaviour can be seen in comparing the accuracy of the volunteered and withheld answers. Focusing only on the 14 participants for whom volunteering rate was less than 100%, on average, 76.7% of each child's volunteered answers were correct, whereas only 58.5% of the withheld answers were correct, $t(13) = 2.19$, $p < .05$. Thus, the children were effective in withholding answers that were liable to be wrong.

Although some of the children volunteered all of their answers, those who withheld some answers benefitted somewhat from the option of free-report (see Koriat & Goldsmith, 1996; Koriat *et al*., 2001). Had they reported all of their answers, as is the case with forced-report memory tests, their output-bound accuracy (see Koriat & Goldsmith, 1996) would have averaged 71.1. Under free-report conditions, their output-bound accuracy was slightly higher: 76.7.

In sum, these results are consistent with both the CM and MC models. Confidence judgments were found to depend on choice latency, consistent with the CM model. In turn, however, they affected the strategic regulation of volunteering behaviour, consistent with the MC model. Because confidence ratings were generally accurate, reliance on confidence as a basis for volunteering decisions helped children avoid answers that were likely to be incorrect, thereby increasing the overall accuracy of the answers that they volunteered.

Experiment 3

In Experiment 3 we included both 2nd graders and 5th graders, but eliminated the collection of confidence judgments. This was done in order to avoid an artificial dependence of volunteering responses on the confidence reported by the child. Assuming that confidence judgments are based on choice latency and that they affect the tendency to volunteer or withhold an answer, we may expect volunteering behaviour to correlate with choice latency even without the overt elicitation of confidence ratings. Experiment 3 also examined age differences in the latency–volunteering correlation: the results of Experiment 1 lead us to expect a stronger volunteering–latency relationship for 5th graders than for 2nd graders.

Method

Participants

Participants were 24 2nd graders (mean age 7.4) and 24 5th graders (mean age 10.5) drawn from the same population as in Experiments 1 and 2.

Materials

The questions were the same as those used in Experiment 1. However, the distracters were modified for some of the questions that yielded too high accuracy and volunteering rates in Experiment 2.

Procedure

The procedure was the same as that of Experiment 2 except that the confidence solicitation phase was eliminated. Thus, after choosing an answer, participants were asked to decide whether to volunteer that answer or to withhold it as in the 'game' procedure used in Experiment 2.

Results and discussion

As in Experiment 1, choice latencies were shorter for 5th graders (6.9 s) than for 2nd graders (10.4 s), $t(46) = 3.89$, $p < .0001$, and accuracy was better for 5th graders (80.9%) than for 2nd graders (68.0%), $t(46) = 5.02$, $p < .0001$. Volunteering rate was somewhat higher for 5th graders (87.5%) than for 2nd graders (79.9%), $t(46) = 2.21$, $p < .05$.

Choice latency as a diagnostic cue for accuracy

We first examine the validity of choice latency as a predictor of accuracy. As in Experiment 1, choice latencies were split at the median for each participant, and percentage correct was calculated for below-median and above-median items. The respective means are presented in Figure 3(a) for each of the two grades. The results generally replicate those of Experiment 1. A two-way ANOVA, Choice latency × Grade, yielded $F(1, 46) = 66.87$, $MSE = 90.94$, $p < .0001$, for latency; $F(1, 46) = 25.20$, $MSE = 158.67$, $p < .0001$, for grade; and $F(1, 46) = 5.96$, $MSE = 90.94$, $p < .05$, for the interaction. As in Experiment 1, the difference between short and long latencies was significant for 2nd graders, $t(23) = 3.33$, $p < .01$, and for 5th graders, $t(23) = 10.44$, $p < .0001$, but the slope was steeper for 5th graders (−3.05) than for 2nd graders (−1.21), $t(46) = 2.66$, $p = .01$. The gamma correlation between latency and accuracy was negative and significant ($p < .0001$) for both grades, −.26 for 2nd graders, and −.51 for 5th graders, with a significant difference between them, $t(46) = 3.91$, $p < .0001$.

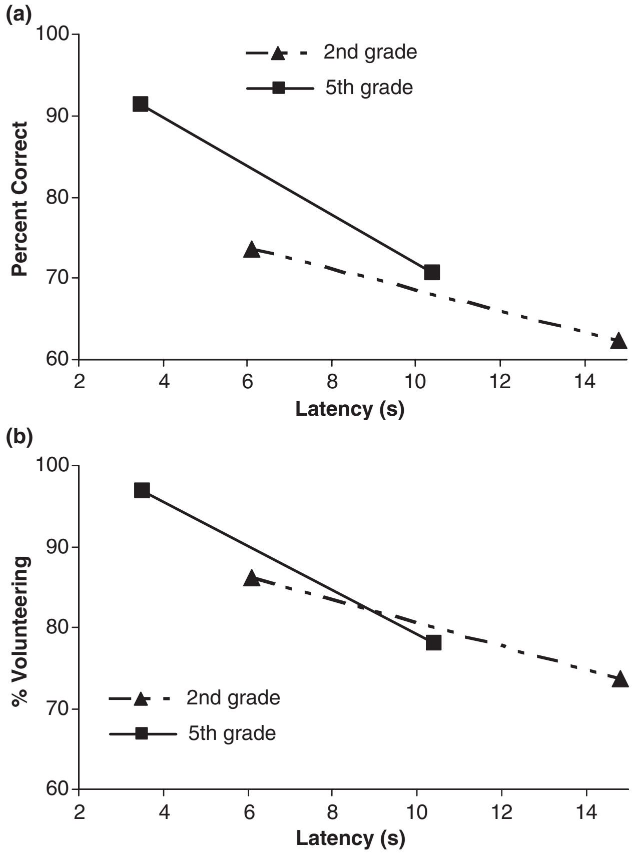

Figure 3 *Mean percentage correct (a) and mean volunteering percentage (b) for below-median (short) and above-median (long) latency of 2nd and of 5th graders in Experiment 3.*

The relationship between choice latency and volunteering

We next examine volunteering behaviour. Figure 3(b) presents mean volunteering percentage for below-median and above-median choice latencies for 2nd and 5th graders. A two-way ANOVA, as before, yielded $F(1, 46) = 95.33$, $MSE = 60.90$, $p < .0001$ for latency, $F(1, 46) = 4.87$, $MSE = 283.29$, $p < .05$ for grade, and $F(1, 46) = 4.14$, $MSE = 60.90$, $p < .05$, for the interaction. The results mimic those obtained for confidence judgments in Experiment 1. First, children in both age groups volunteered short-latency items more than long-latency items, $t(23) = 4.80$, $p < .0001$ for 2nd graders, and $t(23) = 9.96$, $p < .0001$ for 5th graders. Second, the effect of choice latency on volunteering rate was stronger for 5th graders than for 2nd graders.

This pattern of results was also confirmed by an analysis of the within-participant linear regressions of volunteering on latency. The slope of these regressions was significantly higher for 5th graders (−3.61) than for 2nd graders (−1.27), $t(46) = 2.40$, $p < .05$, but each of them was significantly different from zero ($p < .0001$). Similar trends were observed for the within-participant gamma correlations between choice latency and the decision to volunteer the answer: these correlations averaged −.29 for 2nd graders and −.68 for 5th graders (both significant at $p < .01$ and $p < .0001$, respectively), and the correlations were higher for 5th graders than for 2nd graders, $t(46) = 4.60$, $p < .0001$.

In sum, the results suggest that response latency is a good predictor of children's volunteering behaviour, but better so for the older children than for the younger children.

The effectiveness of the volunteering policy

The effectiveness of volunteering behaviour is reflected in the number of points won. Whereas 5th graders gained a mean of 25.1 points, 2nd graders gained only 15.6 points, $t(46) = 4.75$, $p < .0001$. This difference possibly derives from the better memory performance exhibited by 5th graders. However, it may also derive from their better monitoring effectiveness (see Koriat & Goldsmith, 1996): although confidence judgments were not collected in Experiment 3, the results of Experiment 1 suggest that the confidence–accuracy correlation should have been also higher for 5th graders than for 2nd graders. Indeed, correct answers were more likely to be volunteered by 5th graders (92.6%) than by 2nd graders (86.9%) and the answers volunteered by 5th graders (85.8%) were more likely to be correct than those volunteered by 2nd graders (74.6%).

The results also suggest that the higher payoff achieved by the 5th graders may have stemmed in part from their greater reliance on choice latency. To assess the contribution of choice latency to the relationship between volunteering and accuracy, we calculated for each participant the Pearson correlation between them both before and after partialling out the effects of latency. The raw Pearson correlations averaged .27 for 2nd graders and .34 for 5th graders, both significantly higher than zero, $p < .0001$. When latency was partialled out, the respective correlations were lower but still significant, averaging .23 and .26, $p < .0001$. A two-way ANOVA on these correlations, Correlation Type (overall vs. partialled out) × Grade yielded a main effect for correlation type, $F(1, 46) = 24.24$, $MSE = 0.003$, $p < .0001$, but the interaction was also significant, $F(1, 46) = 4.90$, $MSE = 0.003$, $p < .05$. We might speculate that the age difference observed derives from a stronger latency–confidence correlation among 5th graders. This higher correlation may have contributed to the higher payoff achieved by the 5th graders.

General discussion

In this article we attempted to bring into the study of metacognitive development a theoretical approach that has been influential in the study of adult metacognition (e.g. Benjamin & Bjork, 1996; see Koriat, 2007). In that approach, metacognitive monitoring is seen to rely primarily on the immediate feedback from task performance. This approach has far-reaching metatheoretical implications regarding the cause-and-effect relationship between monitoring and control (or between subjective experience and behaviour in general, see Kelley & Jacoby, 1998; Koriat *et al.*, 2006). Unlike the dominant view in metacognition research, that metacognitive judgments drive and guide strategic control (the MC model, Koriat *et al.*, 2006), this approach implies that monitoring actually follows control operations (the CM model). It also implies that metacognitive monitoring is based on processes that are parasitic on the ordinary cognitive operations (Koriat *et al.*, 2008): it is by studying a piece of information that we know whether we will remember it in the future; it is by attempting to retrieve an item from memory that we appreciate whether we 'know' it; and it is by attempting to solve a problem that we judge the likelihood that the solution is correct.

A previous study (Koriat *et al.*, 2009b) indicated that even 3rd graders rely on the feedback from study time or study effort in making recall predictions, and that the reliance on study effort as a cue for JOLs develops with age. The present study examined whether a similar developmental trend would be found for the relationship between response latency and subjective confidence. Previous results with adults established that confidence in the answer to general knowledge questions decreases with choice latency – the amount of time it takes to settle on the answer. This pattern was found in the present study for children in all grades. The latency–confidence relationship is consistent with the CM model, suggesting that children rely on response latency in judging the correctness of their choice. A developmental trend was observed, however, suggesting increased reliance with age on the feedback gained from task performance.

In parallel to these results, which speak for cue utilization, a developmental trend was also observed with regard to cue validity: choice latency was a valid predictor of the accuracy of the answer. That is, the more time it took to answer a question the less likely it was to be correct. However, the validity of response latency as a mnemonic cue for correctness also increased with age.

The combined results on cue utilization and cue validity would lead us to expect improved monitoring effectiveness with age. The results generally supported this expectation: the confidence–accuracy correlation was found to increase with age, and this increase was partly accounted for by the increased reliance on response latency.

Overall, the results of Experiment 1 suggest that even 2nd graders base their confidence judgments on the mnemonic cues that they gain ad hoc from the cognitive operations involved in choosing the answer. It is not clear, however, whether the effective cue is choice latency per se or the effort and deliberation experienced. In either case, these cues are structural or content-less in nature, unlike the kind of declarative metacognitive knowledge that is assumed to underlie information-based judgments. Such cues as fluency, familiarity, accessibility, and experienced effort, give rise directly to immediate feelings of knowing, feelings of mastery, and confidence, through a process that operates below full consciousness (Koriat & Levy-Sadot, 1999). Information-based metacognitive judgments, in contrast, draw upon the content of declarative, domain-specific knowledge retrieved from memory. Our results confirm previous findings indicating that reliance on mnemonic cues that derive from control operations is beneficial because these cues have a certain degree of validity.

Whereas Experiment 1 was driven by the CM model, Experiments 2 and 3 attempted to bring to the fore the implications of the MC model as well. As argued by Koriat *et al*. (2006), the two models are not mutually exclusive. Thus, monitoring that is based on the feedback from control operations may in turn guide subsequent control operations. Experiment 2 attempted to trace the link between the antecedents of monitoring and its consequents. It examined the idea that confidence judgments are based on the feedback from the process of answering a question, but then affect the strategic regulation of memory reporting when a premium is placed on accurate reporting. The results, which were obtained for 2nd graders, yielded clear support for the conceptual scheme depicted in Figure 1. Consistent with the CM model, an inverse relationship was observed between choice latency and confidence. Furthermore, choice latency was diagnostic of accuracy, as was also true of confidence ratings. These results replicate those obtained in Experiment 1. In addition, however, the results suggested that confidence judgments affect the decision whether to volunteer or withhold an answer, consistent with the MC model. Because confidence judgments were generally accurate in discriminating between correct and incorrect answers, reliance on confidence allowed children to screen out answers that were likely to be wrong. These results illustrate the functional value of monitoring-based regulation of behaviour (Koriat, 2000; Son & Schwartz, 2002).

In Experiment 3 we examined possible age changes in control-based monitoring. We deliberately avoided soliciting confidence judgments, but assumed that confidence in an answer mediates the link between choice latency and volunteering decisions, as suggested by the results of Experiment 2. The results confirmed the validity of choice latency as a cue for accuracy and also replicated the finding from Experiment 1 that cue validity increases with age. In addition, the results for volunteering behaviour mimicked those obtained for confidence judgments in Experiment 1. First, the likelihood of volunteering was inversely related to choice latency for both 2nd and 5th graders. Second, this relationship was stronger for the older children than for the younger children.

The results of Experiments 2 and 3 that pertain to the MC model are consistent with the idea that metacognitive judgments are not mere epiphenomena but actually guide and affect strategic control (Nelson & Narens, 1990). This is true even when metacognitive judgments are based on the feedback from control processes (Koriat, 2000). The results suggest that the option that children were given to volunteer or withhold an answer was exploited more effectively by the older children than by the younger children. The improved strategic regulation of the older children seems to derive in part from their better monitoring, which in turn stems from their greater reliance on response latency as a cue for confidence.

The present study raises several theoretical questions. What are the developmental changes that underlie the age differences observed in Experiments 1 and 3? Specifically, what is the process responsible for the increased reliance with age on the feedback from task performance and for the increased validity of that feedback as a cue for accuracy? Several results with adults suggest that people interpret mnemonic cues according to their ecological validity, and may change and adapt their interpretation of these cues depending on the feedback that they receive. Unkelbach (2006), for example, reported results suggesting that the interpretation of fluency cues is learnable and can even be reversed as a result of training. Possibly, children gradually internalize the probabilistic associations between mnemonic cues and various performance criteria according to the feedback that they gain as they carry out different tasks (see Koriat *et al*., 2009a). This development may be similar to that underlying spatial perception, in which infants develop sensitivity to various depth cues that help them perceive the world correctly (e.g. Yonas, Elieff & Arterberry, 2002).

It would seem that the cue-learning process involved is implicit rather than explicit. In fact, children as well as adults are hardly aware of reliance on mnemonic cues. In previous work (Koriat *et al*., 2009a) it was found that when children were presented with repeated study-test blocks of a series of items, their recall and JOLs for different items decreased as a function of number of trials to acquisition (TTA). Nevertheless, when asked immediately after study, the children did not disclose the belief that items that require more TTAs are less likely to be recalled than those that require fewer TTAs. These and other results suggest that the relationships between mnemonic cues and memory performance are established through implicit learning, and that the utilization of these cues is also implicit and largely unconscious. Furthermore, findings with adults suggest that attempts to enhance awareness of the reasoning processes involved in decisions may impair the confidence–accuracy relation, and that this correlation can sometimes be enhanced by increasing the salience of automatic processes (Robinson & Johnson, 1998).

In addition to the age increase in cue utilization, however, there was also an age increase in cue validity: the diagnostic validity of choice latency as a cue for accuracy improved with age. Koriat *et al*. (2009b), who observed an age increase in the effectiveness with which study time was allocated to easy and difficult items, suggested that this difference may derive from a developmental change in the very regulation of study time: the allocation of study time by the older children is better tuned to features of the studied materials that are critical for learning and remembering. Perhaps in a similar manner, as children develop, response latency also becomes increasingly sensitive to the cues that disclose the correctness of the answer. This implies a developmental change not only in monitoring but also in control operations – the amount of time that children invest in trying to answer each question.

In conclusion, the present study joins with the previous work on JOLs in bringing to the fore the contribution of subtle mnemonic cues and heuristics that influence children's monitoring of their knowledge during learning and remembering. Such cues may ultimately influence strategic cognitive processes and behaviour.

Acknowledgements

This study was conducted at the Institute of Information Processing and Decision Making, University of Haifa, Haifa, Israel. It was supported by European Commission Grant FP6-NEST: EYEWITMEM; 43460. We are grateful to the Chief Scientist of the Israeli Ministry of Education and Culture for approval of the project. We thank Dana Gershon, Galia Wieder and Tamar Zamet for their help in conducting the experiments.

References

Barnes, A.E., Nelson, T.O., Dunlosky, J., Mazzoni, G., & Narens, L. (1999). An integrative system of metamemory components involved in retrieval. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 287–313). Cambridge, MA: MIT Press.

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, N., & Day, K. (2005). The confidence–accuracy and decision latency–accuracy relationships in children's eyewitness identification. *Psychiatry, Psychology and Law*, **12**, 119–128.

Brunswik, E. (1956). *Perception and representative design in psychological experiments*. Berkeley, CA: University of California Press.

Cohen, R.L., Sandler, S.P., & Keglevich, L. (1991). The failure of memory monitoring in a free recall task. *Canadian Journal of Psychology*, **45**, 523–538.

Dufresne, A., & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: differential and sufficient aspects. *Journal of Experimental Child Psychology*, **47**, 274–296.

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve learning in later adulthood: helping older adults educate themselves. In D.J. Hacker (Ed.), *Metacognition in educational theory and practice* (pp. 249–275). Mahwah, NJ: Erlbaum.

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.

Finn., B. (2008). Framing effects on metacognitive monitoring and control. *Memory Cognition*, **36**, 813–821.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: the appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 552–564.

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

Gill, M.J., Swann, B., Jr, & Silvera, D.H. (1998). On the genesis of confidence. *Journal of Personality and Social Psychology*, **75**, 1101–1114.

Goldsmith, M., & Koriat, A. (2008). The strategic regulation of memory accuracy and informativeness. In A. Benjamin & B. Ross (Eds.), *Psychology of learning and motivation, Vol. 48: Memory use as skilled cognition* (pp. 1–60). San Diego, CA: Elsevier.

Hart, J.T. (1965). Memory and the feeling-of-knowing experiments. *Journal of Educational Psychology*, **57**, 347–349.

Hoffrage, U. (2004). Overconfidence. In R.F. Pohl (Ed.), *Cognitive illusions: A handbook on fallacies and biases in thinking, judgment and memory* (pp. 235–254). Hove, UK: Psychology Press.

Kelley, C.M., & Jacoby, L.L. (1998). Subjective reports and process dissociation: fluency, knowing, and feeling. *Acta Psychologica*, **98**, 127–140.

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.

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. (1995). Dissociating knowing and the feeling of knowing: further evidence for the accessibility model. *Journal of Experimental Psychology: General*, **124**, 311–333.

Koriat, A. (1997). Monitoring one's own knowledge during study: a cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, **126**, 349–370.

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: Cambridge University Press.

Koriat, A. (2008a). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition*, **36**, 416–428.

Koriat, A. (2008b). Subjective confidence in one's answers: the consensuality principle. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **34**, 945–959.

Koriat, A. (2008c). When confidence in a choice is independent of which choice is made. *Psychonomic Bulletin and Review*, **15**, 997–1001.

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

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., & 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., & Shitzer-Reichert, R. (2002). Metacognitive judgments and their accuracy: insights from the processes underlying judgments of learning in children. In P. Chambres, M. Izaute & P.–J. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 1–17). New York, NY: Kluwer.

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

Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology*, **79**, 405–437.

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., 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.

Koriat, A., Nussinson, R., Bless, H., & Shaked, N. (2008). Information-based and experience-based metacognitive judgments: evidence from subjective confidence. In J. Dunlosky & R.A. Bjork (Eds.), *A handbook of memory and metamemory* (pp. 117–135). Mahwah, NJ: Erlbaum.

Koriat, A., Ackerman, R., Lockl, K, & Schneider, W. (2009a). The easily learned, easily-remembered heuristic in children. *Cognitive Development*, **24**, 169–182.

Koriat, A., Ackerman, R., Lockl, K, & Schneider, W. (2009b). The memorizing-effort heuristic in judgments of learning: a developmental perspective. *Journal of Experimental Child Psychology*, **102**, 265–279.

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

Lockl, K., & Schneider, W. (2002). Developmental trends in children's feeling-of-knowing judgements. *International Journal of Behavioral Development*, **26**, 327–333.

Lockl, K., & Schneider, W. (2003). Metakognitive Überwachungs- und Kontrollprozesse bei der Lernzeiteinteilung von Kindern [Monitoring and control processes in children's allocation of study time]. *Zeitschrift für Pädagogische Psychologie*, **17**, 173–183.

Lyon, T.D., & Flavell, J.H. (1993). Young children's understanding of forgetting over time. *Child Development*, **64**, 789–800.

Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): an analytic or nonanalytical basis for JOLs? *Memory & Cognition*, **29**, 222–233.

Mazzoni, G, Loftus, E.F., & Kirsch, I. (2001). Changing beliefs about implausible autobiographical events: a little plausibility goes a long way. *Journal of Experimental Psychology: Applied*, **7**, 51–59.

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, **15**, 174–179.

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

Nelson, T.O., & Leonesio, R.J. (1988). Allocation of self-paced study time and the 'labor-in-vain effect'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 676–686.

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* (pp. 125–173). New York: Academic Press.

Odegard, T.N., & Lampinen, J.M. (2006). Memory editing: knowledge, criteria, and alignment. *Memory*, **14**, 777–787.

O'Sullivan, J.T. (1993). Preschooler's beliefs about effort, incentives, and recall. *Journal of Experimental Child Psychology*, **55**, 396–414.

O'Sullivan, J.T. (1997). Effort, interest, and recall: beliefs and behaviors of preschoolers. *Journal of Experimental Child Psychology*, **65**, 43–67.

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. (1998). How not to enhance the confidence–accuracy relation: the detrimental effects of attention on the identification process. *Law and Human Behavior*, **22**, 409–428.

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.

Roebers, C., von der Linden, N., Schneider, W., & Howie, P. (2007). Children's metamemorial judgments in an event recall task. *Journal of Experimental Child Psychology*, **97**, 117–137.

Schneider, W. (2008). The development of metacognitive knowledge in children and adolescents: major trends and implications for education. *Mind, Brain, and Education*, **2**, 114–121.

Schneider, W., & Bjorklund, D.F. (1998). Memory. In W. Damon, R.S. Siegler & D. Kuhn (Eds.), *Handbook of child psychology* (Vol. 2, pp. 467–521). New York: Wiley.

Schneider, W., & Lockl, K. (2008). Procedural metacognition in children: evidence for developmental trends. In J. Dunlosky & R.A. Bjork (Eds.), *A handbook of memory and metamemory* (pp. 391–410). Mahwah, NJ: Erlbaum.

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

Schwartz, B.L. (1994). Sources of information in metamemory: judgments of learning and feeling of knowing. *Psychonomic Bulletin & Review*, **1**, 357–375.

Schwartz, B.L. (2002). The phenomenology of naturally-occurring tip-of-the-tongue states: a diary study. In S. Shohov (Ed.), *Advances in Psychology Research*. (pp. 71–84). New York: Nova Science Publishers.

Son, L.K., & Schwartz, B.L. (2002). The relation between metacognitive monitoring and control. In T.J. Perfect & B.S. Schwartz (Eds.), *Applied metacognition* (pp. 15–38). Cambridge: Cambridge University Press.

Sporer, S.L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential line-ups. *Journal of Applied Psychology*, **78**, 22–33.

Thiede, K.W., Anderson, M.C.M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, **95**, 66–73.

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

Weber, N., Brewer, N., Wells, G.L., Semmler, C., & Keast, A. (2004). Eyewitness identification accuracy and response latency: the unruly 10–12 second rule. *Journal of Experimental Psychology: Applied*, **10**, 139–147.

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.

Yonas, A., Elieff, C.A., & Arterberry, M.E. (2002). Emergence of sensitivity to pictorial depth cues: charting development in individual infants. *Infants Behavior & Development*, **25**, 495–514.

Zakay, D., & Tuvia, R. (1998). Choice latency times as determinants of post-decisional confidence. *Acta Psychologica*, **98**, 103–115.

Received: 10 November 2008
Accepted: 4 March 2009