METACOGNITION
Process, Function and Use

edited by

Patrick Chambres
Marie Izaute
Pierre-Jean Marescaux

Université Blaise Pascal
Clermont-Ferrand, France

KLUWER ACADEMIC PUBLISHERS
Boston / Dordrecht / New York / London

Distributors for North, Central and South America:
Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA
Telephone (781) 871-6600
Fax (781) 681-9045
E-Mail: kluwer@wkap.com

Distributors for all other countries:
Kluwer Academic Publishers Group
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS
Telephone 31 786 576 000
Fax 31 786 576 474
E-Mail: services@wkap.nl

Electronic Services <http://www.wkap.nl>

Library of Congress Cataloging-in-Publication Data

Metacognition: process, function, and use / edited by Patrick Chambres, Marie Izaute, Pierre-Jean Marescaux.
p.cm.
Includes bibliographical references and indexes.
ISBN 1-40207-134-5 (alk. paper)
1. Metacognition. I. Chambres, Patrick, 1955- II. Izaute, Marie, 1965- III. Marescaux, Pierre-Jean, 1958-

BF311 .M44892 2002
153—dc21

2002075420

Copyright © 2002 by Kluwer Academic Publishers

All rights reserved. No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without the written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Permission for books published in Europe: permissions@wkap.nl
Permission for books published in the United States of America: permissions@wkap.com

Chapter 1

Metacognitive Judgments and their Accuracy

Insights from the Processes Underlying Judgments of Learning in Children

Asher Koriat and Rachel Shitzer-Reichert
Department of Psychology, University of Haifa, Israel

Key themes: Devl / Proc / Theor

Key words: Judgments of learning / Monitoring accuracy / Metacognitive development / Mnemonic cues / Delayed JOLs

Abstract: In this chapter we begin by examining the processes underlying metacognitive judgments, contrasting the two major approaches to the study of metacognition –the developmental and cognitive-experimental approaches. Focusing then on the monitoring of one's own knowledge during study, we point out the benefits of applying insights from cognitive psychology to the study of the determinants of monitoring accuracy in children. The results of two experiments suggest that similar processes underlie judgments of learning (JOLs) and their accuracy in adults and children.

A commonly held assumption among students of metacognition is that metacognitive judgments exert a causal effect on information processing and behavior. This assumption has been formulated in terms of the effects of monitoring on control (Nelson & Narens, 1990). Monitoring refers to one's subjective assessment of one's own knowledge whereas control refers to the regulation of behavior that is presumably based on the output of the monitoring system. According to this formulation, there is benefit in investigating the accuracy of metacognitive judgments because it has important consequences for the effective adaptation to reality.

1. DEVELOPMENTAL AND COGNITIVE PERSPECTIVES ON METACOGNITION

Historically, the investigation of metacognitive processes has proceeded along two almost entirely separate lines. On the one hand, there has been extensive research in developmental psychology, spurred mainly by the work of Flavell (1979) and his associates, which emphasized the critical role of metacognitive processes in the development of memory functioning in children. On the other hand, there has been a line of investigation in cognitive psychology that has focused narrowly on several questions concerning the determinants and consequences of the monitoring of one's own knowledge.

Developmental work on metacognition has focused primarily on specifying the components of metacognitive abilities as they develop with age, and on their possible effects on memory functioning. The definition of metacognition is much broader than that which seems to underlie much of the cognitive work on metacognition. Thus, in Flavell's conceptualization metacognition is seen to encompass metacognitive knowledge, metacognitive experiences, goals and actions (see Flavell, 1979; 1999; Flavell & Wellman, 1977). Indeed, developmental research has addressed such questions as what children know about the strengths and limitations of memory in general and of their memory in particular, and what they know about task variables that affect memory performance (e.g., Kreutzer, Leonard, & Flavell, 1975). Such metacognitive knowledge is certainly critical in guiding the effective management of learning and remembering. Developmental work has also placed a heavy emphasis on strategies of learning and remembering, including knowledge about the benefits and costs of using strategies in general, the potential value of specific strategies, the choice of strategies, the ability to take advantage of a strategy following instructions to use it, and so on (Bjorklund & Douglas, 1997; Pressley, Borkowski, & Schneider, 1987).

The assumption underlying much of this work is that memory performance depends heavily on monitoring and regulatory proficiency. Indeed, developmental psychologists investigated the relationship between metamemory and memory skills and how both of these develop with age (see Schneider, 1985). Much of that work is correlational in nature, and some of it is primarily descriptive. Furthermore, some of the work on metacognitive knowledge has relied heavily on self-report techniques such as interviews or questionnaires (e.g., Kreutzer et al., 1975).

In contrast, the study of metacognition by experimental cognitive psychologists has been more narrowly confined to several basic issues concerning the mechanisms of monitoring and control processes in memory (for reviews see Nelson & Narens, 1990; Koriat & Levy-Sardot, 1999; Schwartz, 1994). A great deal of the work has focused on within-individual variation to reveal the dynamics of metacognitive processes. Thus, within-subject correlations have been typically used to examine the accuracy of metacognitive feelings as well as the effects of metamemory on memory (e.g., Nelson, 1984; Koriat & Goldsmith, 1996). This research has given rise to the establishment of several experimental paradigms for examining the monitoring and control processes that occur during learning, during the attempt to retrieve information from memory and following the retrieval of candidate answers (e.g., Hart, 1965; Nelson & Leonesio, 1988; Koriat, Lichtenstein, & Fischhoff, 1980; Reder, 1987).

Several questions have been at the focus of investigation. First how accurate are metacognitive judgments, and what are the factors that affect their accuracy? (e.g., Schwartz & Metcalfe, 1994; Weaver & Kelemen, this volume). Second, what are the bases of metacognitive judgments, that is, how do people monitor their own knowledge? (e.g., Cary & Reder, this volume; Koriat & Levy-Sardot, 2001). Third, what are the processes that are responsible for the accuracy and inaccuracy of metacognitive judgments? For example, what are the processes that lead to illusions of knowing, that is, to situations in which people have strong, unwarranted confidence in their knowledge? (e.g., Benjamin & Bjork, 1996; Bjork, 1999; Koriat, 1998). Fourth, how do metacognitive judgments control and guide information processing and action? (e.g., Barnes, Nelson, Dunlosky, Mazzoni, & Narens, 1999; Son & Metcalfe, 2000). This question is predicated on the assumption that monitoring processes play a causal role in regulating cognitive processes and behavior (see Koriat, 2000). Finally, how do the metacognitive processes of monitoring and control affect learning and remembering? (e.g., Barnes et al., 1999; Koriat & Goldsmith, 1996).

Additional questions, of course, emerge in different contexts. For example, assuming that metacognitive processes are not activated routinely, the question then is what are the conditions that induce people to engage in metacognitive processes? (see Chambres, Bonin, Izaute, & Marescaux, this volume). Do metacognitive skills represent a stable and reliable dimension of individual differences? (see Weaver & Kelemen, this volume). Can metacognition be trained, that is, can procedures be devised that improve monitoring accuracy? (e.g., Dunlosky & Nelson, 1994; Koriat, et al., 1980). These are but a few of the questions addressed by experimental students in their attempt to clarify the processes underlying metacognitive monitoring and control.

In contrast to the focus on process and on within-subject variation, developmental psychologists exhibit a tendency to treat metacognition as a series of skills. Hence the interest in individual differences and age differences, as well as in questions concerning the generality or task-specificity of metacognitive skills, and the extent to which such skills correlate with IQ or predict school achievement. This treatment of metacognition has also led to attempts to specify "deficiencies" that are characteristic of children at different ages, and to seek ways to remedy them.

There is certainly benefit in combining insights from the developmental and cognitive approaches to metacognition. The developmental approach provides breadth (see Paris, this volume): It offers a more comprehensive framework for the analysis of metacognition, and brings to the fore questions that have not attracted sufficient interest among cognitive psychologists. Apart from its emphasis on developmental issues, it has stressed the consequences of metacognitive processes, particularly as far as memory performance is concerned. The cognitive approach, on the other hand, provides depth: A more detailed, theoretically-driven analysis of the working of metacognition. It has also resulted in the development of several standard experimental paradigms that offer many opportunities for the study of various basic processes in metacognition in both children and adults. Although these paradigms are rather restricted, they can provide some insight into the internal dynamics of metacognitive monitoring and control (Barnes et al., 1999).

All of the chapters included in this section can be seen to combine some aspects of the cognitive and developmental approaches in attempting to elucidate the processes underlying metacognition. Weaver and Kelemen's chapter is relevant to the conception of metacognition as a set of skills. As noted earlier, this conception underlies many of the studies that examined age-related differences in metacognition and the relationships between memory and metamemory. What is unique about the studies reported by Weaver and Kelemen is the inclusion of measures of individual differences in metacognitive accuracy that are based on **within**-person correlations. The results reviewed in that chapter question the possibility of a stable and reliable dimension of individual differences in monitoring proficiency. An important challenge is how to reconcile these findings with the systematic age-related differences observed in some aspects of metacognition.

The chapter by Moulin, Perfect, and Fitch, is representative of the recent attempts by developmental psychologists, cognitive psychologists and neuropsychologists to seek an explanation of memory deficits in terms of deficient metacognitive abilities. Not only do such attempts help clarify the nature of metacognitive deficiencies, but they can also contribute a great deal to our understanding of the processes underlying metacognitive judgments in general. An important feature of the results described by Moulin et al. on Alzheimer patients is that they suggest a dissociation in these patients between monitoring and control processes during the study of new materials. This dissociation runs counter the commonly held "monitoring-affects-control" hypothesis of self-paced learning (Nelson & Leonesio, 1988).

Efklides' chapter builds on the distinction advanced by Flavell between metacognitive knowledge and metacognitive experience (ME; see Efklides, in press). Whereas the former refers to long-term beliefs concerned with memory functioning, the latter refers to conscious affective or cognitive experiences that normally accompany on-line the monitoring and self-regulatory processes that take place during encoding and remembering (Brown, Bransford, Ferrara, & Campione, 1983; Paris & Lindauer, 1982). It is MEs that received greater emphasis among experimental cognitive psychologists (Koriat & Levy-Sardot, 1999). However, Efklides' study, which was carried out on 7th to 9th graders, emphasizes the richness of MEs, and their interrelations. Although the study embodies certain aspects of the developmental approach to metacognition, its main focus is to clarify the mechanisms underlying metacognitive feelings and the function of these feelings in cognitive processing.

The experimental work presented in this chapter attempts to import insights from cognitive psychology to the study of developmental aspects of metacognition. It concerns the monitoring of knowledge during learning, focusing on the processes underlying the accuracy of JOLs in children of two age groups. This work, like some of the recent studies referred to below, is intended to promote a greater crosstalk between developmental and cognitive students of metacognition.

2. THE BASIS OF JUDGMENTS OF LEARNING AND THEIR ACCURACY

When studying new material, people normally monitor the extent to which they have mastered different parts of that material and control the allocation of learning resources accordingly. Memory performance, then, should depend not only on "memory" but also on "metamemory", that is on the extent to which a person is successful in monitoring the degree of knowledge of different items and regulating study resources accordingly. An important question in developmental research, then, concerns the extent to which the age-related improvement in memory performance might be mediated by improvement in the monitoring of one's own memory during learning. Several studies that examined this question have yielded inconsistent results (see Schneider, Visé, Lockl, & Nelson, in press). In this study we also investigate developmental trends in monitoring accuracy during learning, but our primary focus is on the processes underlying the accuracy of JOLs elicited during study. We wish to examine the bases for children's accurate monitoring and whether these bases are similar to those that have been found for adults.

Most of the developmental studies on monitoring have concerned calibration or absolute metacognitive accuracy (see Weaver & Kelemen, this volume) that is, the match between the predicted and actual overall memory performance. These studies have generally indicated that preschoolers and kindergarten children tend to overestimate their future memory performance, whereas schoolchildren's predictions tend to be more realistic (see Schneider & Pressley, 1997). In the present study instead we focus on resolution or relative accuracy, that is, the accuracy of JOLs in monitoring the relative recallability of different items, as indexed, for example, by a within-subject Goodman-Kruskal gamma correlation between JOLs and recall (see Nelson, 1984). Resolution, or relative accuracy, is critical for the efficient allocation of time and effort between different items in self-paced learning.

What are the determinants of JOLs and their accuracy? According to the cue utilization model of JOLs proposed by Koriat (1997), JOLs are inferential in nature, and rest on a variety of cues. Three classes of cues for JOLs were distinguished, intrinsic, extrinsic, and mnemonic. Intrinsic cues refer to inherent characteristics of the study items that disclose their a-priori difficulty. For example, in paired associates learning, the judged degree of associative relatedness between the members of the pairs is an important contributor to JOLs. Extrinsic cues pertain to the conditions of learning (e.g., number of presentations), or to the encoding operations applied by the learner (e.g., level of processing). Finally, mnemonic cues are internal, subjective indicators that signal to the person the extent to which an item has been mastered. Several types of mnemonic cues have been discussed as possible determinants of metacognitive judgments: the fluency of processing of a presented item (Benjamin & Bjork, 1996; Koriat, 1997), the familiarity of the cue that serves to probe memory (Cary & Reder, this volume; Metcalfe, Schwartz, & Joaquim, 1993), the accessibility of pertinent partial information about a memory target (Dunlosky & Nelson, 1992; Koriat, 1993), and the ease with which information is retrieved (Kelley & Lindsay, 1993; Koriat, 1993).

Further, Koriat (1997) proposed that intrinsic and extrinsic cues can affect JOLs directly, through the explicit application of a particular rule or theory. For example, a person may hold the belief that the same item is more likely to be remembered if it is presented several times than if it is presented only once, or that semantically related pairs are easier to learn and remember than unrelated pairs in paired-associates learning. Such beliefs may be applied directly in making a theory-based inference. However, intrinsic or extrinsic cues may also affect JOLs through their influence on internal, mnemonic cues. For example, an item seen previously may be processed more fluently than a new item (Jacoby & Kelley, 1987). Processing fluency, then, can serve as the immediate cue for JOLs.

The direct effects of intrinsic and extrinsic cues are assumed to involve an analytic, deliberate inference based on the person's a-priori theory about the memorial consequences of various factors. The effects of mnemonic cues, in contrast, are assumed to rest on the implicit use of a nonanalytic, unconscious inference rather than on a deliberate theory-based deduction (see Koriat, 2000; Koriat & Levy-Sardot, 1999). It is proposed that when intrinsic and extrinsic cues are directly consulted in making JOLs, the result is an information- (or theory-) based judgment of knowing. Mnemonic cues, on the other hand, give rise to a feeling of knowing, which can then serve as a basis for a judgment. Thus, the distinction between information-based and experience-based metacognitive judgments (see Koriat & Levy-Sardot, 1999) parallels in part Flavell's (1979) distinction between the effects of metacognitive knowledge and those of metacognitive experiences (see also Efklides, this volume).

The distinction between the analytic and nonanalytic inferential processes mediating JOLs has important implications for JOL accuracy. When JOLs are based on the explicit application of a belief or theory (e.g., "I have poor memory for names", "associatively-related pairs are better remembered than unrelated pairs"), their accuracy should depend greatly on the validity of the underlying theories or beliefs. It is these theories or beliefs that have received a great deal of attention in the context of developmental studies of declarative metamemory (or metacognitive knowledge, Flavell, 1979). These studies suggest an age-related increase in the accuracy of children's beliefs about memory, and this increase should, of course, contribute to enhanced monitoring accuracy of theory-based JOLs.

The accuracy of heuristic-driven JOLs, in contrast, depends on the validity of the underlying cues. Although such cues can sometimes be misleading as predictors of memory performance (e.g., Benjamin, Bjork, & Schwartz, 1998), they are generally dependable because they are influenced by both intrinsic and extrinsic cues that affect learning and remembering (Benjamin & Bjork, 1996; Jacoby & Kelley, 1987). Therefore the accuracy of JOLs should generally increase as a function of the extent to which they are based on internal, mnemonic cues.

Koriat (1997) proposed that with repeated study of the same material the basis of JOLs changes from reliance on intrinsic cues towards increased reliance on internal, mnemonic cues. In support of this proposition, two changes were observed with practice studying the same list of paired associates. First, the accuracy of JOLs in predicting subsequent recall increased gradually from one study-test cycle to the next (King, Zechmeister, & Shaughnessy, 1980; Leonesio & Nelson, 1990; Mazzoni, Cornoldi, & Marchitelli, 1990; see also Weaver & Kelemen, this volume). Thus the within-participant cross-item correlation between JOL and recall increased from .66 on the first study-test cycle to .89 on the 4th cycle. This improvement in JOL accuracy was attributed to the increased reliance on mnemonic cues under the assumption that mnemonic cues closely reflect the cognitive processing of the items. Second, the correlation between JOLs and the a-priori difficulty of the paired associates, as rated by a different group of participants, decreased gradually with practice, averaging .93 and .73, respectively, for the 1st and 4th blocks. Judged item difficulty represents an intrinsic cue that can affect JOLs, and thus the changes in the JOL-difficulty correlation were seen to reflect decreased reliance on intrinsic cues with practice studying the same items.

In Experiment 1 we tested whether this presumed dynamics of JOLs also occurs in young school children. The question is whether children also reveal the assumed shift from reliance on intrinsic factors towards greater reliance on mnemonic factors with increased practice studying the same list of items. As noted earlier, developmental research have invested little attention in the nonanalytic processes underlying metacognitive judgments, and Experiment 1 may help remedy this

situation by importing insights from the experimental-cognitive study of the dynamics of monitoring processes.

2.1 Experiment 1

Experiment 1 was modeled after Experiment 2 of Koriat (1997) but it was carried out on 2nd grade and 4th grade children. A list of paired associates, composed of hard and easy pairs was presented for four study-test cycles. Thus, apart from age, the experiment included an intrinsic factor (item difficulty) and an extrinsic factor (practice). Feedback about the correctness of the answers was also manipulated between participants.

2.1.1 Method

Participants. Participants were 32 second graders (mean age = 7.2 years) and 32 fourth graders (mean age = 9.7 years) from predominantly middle class homes. In each group participants were assigned randomly to the Feedback and No-Feedback conditions with the constraint that there was an equal number of boys and girls in each Age X Feedback condition.

Materials. The items were 24 pairs of Hebrew words that were selected on the basis of a preliminary study. In that study, 30 2nd graders and 30 4th graders were asked to rate 50 Hebrew pairs in terms of memorability. Specifically, they were asked to imagine that 100 children had studied these pairs, and to estimate how many of them would recall each pair, that is, would recall the response word when presented with the stimulus word. The median estimates were used to order the pairs in terms of judged a-priori difficulty. Twenty-four words were selected for which there was generally an agreement between 2nd and 4th graders, and were divided at the median into two sets of 12 easy and 12 hard pairs.

Procedure. Children were tested individually in a quiet room in the school, using a PC compatible laptop. They were instructed that they would have to study pairs appearing on the computer screen so that, during the test phase, they would be able to recall the response word when cued with the stimulus word. They were also told that at the end of each study trial they would have to estimate the likelihood of recalling the response word during the test phase. The elicitation of JOLs capitalized on the hot-cold game familiar to children, using a thermometer procedure devised by Koriat, Goldsmith, and Schneider (1999). The rules of the hot-cold game were explained, and participants were required to rate their JOLs on a 5-point scale depicted as a color drawing of a thermometer ranging from deep blue ("very cold", i.e., "no chance to recall the response word") to deep red ("very hot", i.e., "completely certain to recall the response word"). A large drawing of the thermometer was placed on the table in front of the child.

During the study phase the intact pair remained on the screen for 5 s, and was replaced by the statement "how sure are you that you will recall the second word later when you see the first word?" The child indicated his/her answer by placing a cube on one of the five colored segments of the thermometer drawing. When all the pairs had been presented for study, the test phase began: Each of the stimulus words was presented in turn, and the child had to speak aloud the answer. The stimulus word remained on the screen until the child responded, or until 10 s have elapsed.

The procedure was the same for the no-feedback and feedback conditions except that in the latter condition a sound was presented for 30 ms when the response provided was incorrect. (The instructions for the feedback condition included an explanation of the significance of the sound).

The study-test phase cycle was repeated three more times. The presentation of the items was random during all study and test phases.

2.1.2 Results.

The feedback manipulation had little effect and will not be discussed further. Considering first the results for the first block, the intrinsic factor of item difficulty had a strong effect on both recall and JOL. Recall for judged easy and hard items averaged 73.0% and 11.8%, respectively, and there was no effect of age. There was an Age X Difficulty interaction, however, with regard to JOLs: Whereas for easy items there was little difference between 2nd and 4th graders (the respective means were 4.06 and 3.92), for the hard items 2nd graders gave higher JOLs (the respective means were 3.51 and 2.99). Note that a rating of 3 was described in the instructions as "I may recall or I may not". If that rating is assumed to be roughly equivalent to a .5 probability, then the results would seem to suggest very inflated JOLs, with degree of overconfidence being stronger for the younger children (see also Schneider et al., in press).

We shall turn now to the effects of the extrinsic factor of repeated presentation. Figure 1 depicts mean recall (top panel) and JOL (bottom panel) as a function of presentation and item difficulty for each of the age groups. It can be seen that both age groups exhibited strong improvement in recall from 42% on the first presentation to 73% for the 4th presentation. JOLs also increased with presentation for both age groups, indicating that children's JOLs are also sensitive to the extrinsic factor of practice (cf. Moulin et al., this volume). This increase, however, was monotonic for easy items, whereas for hard items there was, in fact, a drop from the first to the second presentation. This pattern suggests that children in both age groups corrected their inflated JOLs in response to their low actual memory performance after the first presentation. It can also be seen that 2nd graders continued to provide higher JOLs than the 4th graders throughout the 4 presentations, but this age effect was entirely confined to the hard items.

We examine now the accuracy of JOLs in predicting inter-item differences in recall. For each child, a gamma correlation was calculated between JOL and recall across all 24 items (Nelson, 1984). Figure 2 (top panel) presents the means of these correlations as a function of presentation for each of the two age groups. An Age X Presentation X Feedback ANOVA also failed to yield any effect of feedback and therefore the results in Figure 2 are pooled across the two feedback conditions. However, there were significant effects for both age, $F(1, 55) = 13.91, p<.0005$, and presentation, $F(3, 165) = 13.99, p<.0001$.

The effect of age reflects the observation that the older children's predictions were more accurate than the younger children's predictions. This difference was significant ($p<.002$) even on the first presentation: Gamma correlation averaged .40 for the younger group and .66 for the older group. Note, however, that even the younger group's resolution was relatively high and significant ($p<.0001$). Thus, we have an indication of a developmental trend in monitoring skill but also for efficient relative monitoring even among 2nd graders.

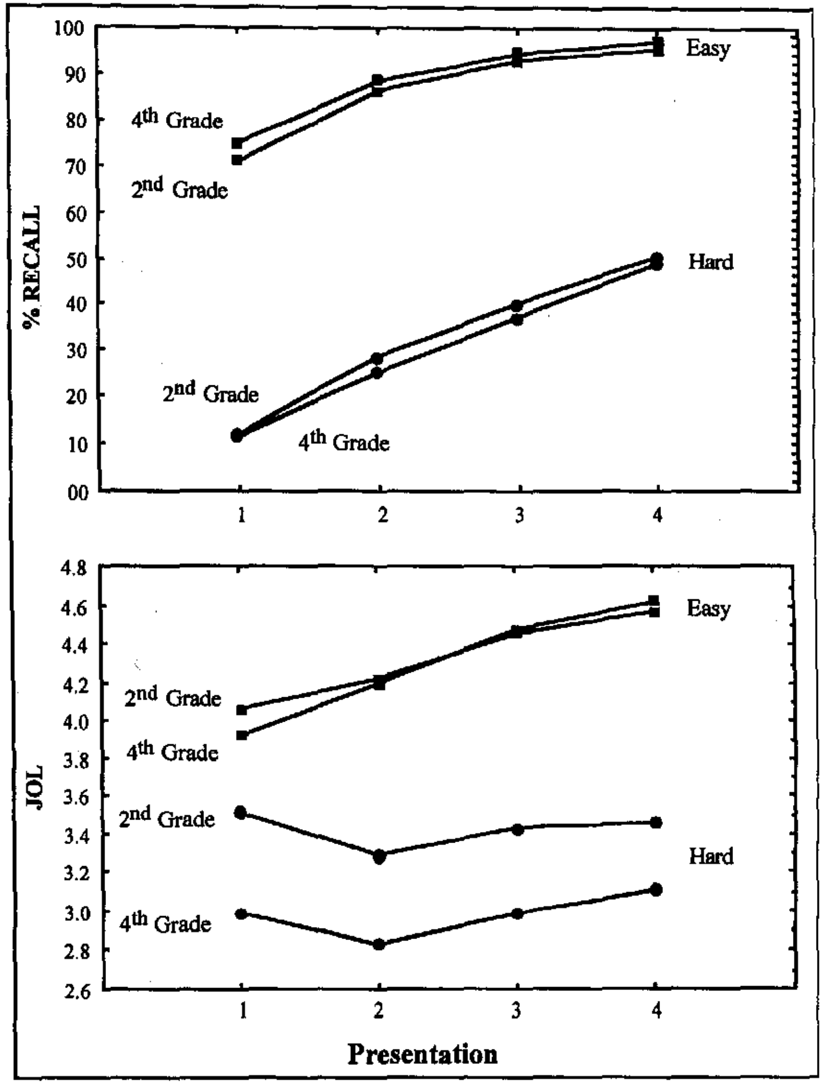

Figure 1. Mean recall (top panel) and JOL (bottom panel) as a function of presentation and item difficulty for 2nd and 4th graders (Experiment 1).

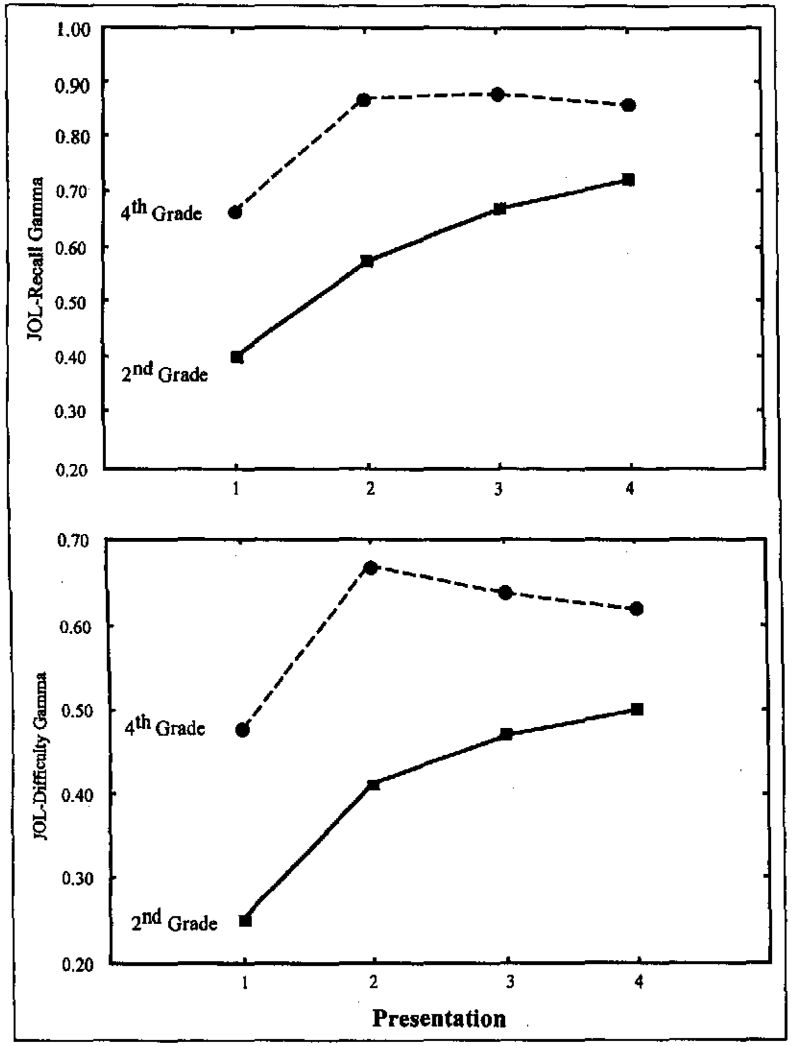

Figure 2. Mean within-subject gamma correlations between JOLs and recall (top panel) and between JOL and item difficulty (bottom panel), plotted as a function of presentation (Experiment 1).

As for the effects of presentation, the results replicate the pattern that has been observed previously with adults. The improvement in JOL accuracy with practice was more clearly seen in the younger children, possibly because of a ceiling effect that occurred for the older group. In fact, practice seemed to close the age gap so that the gamma correlation exhibited by the younger group on the 4th presentation was the same as that exhibited by the older group on the 1st presentation.

What are the implications of these results? According to Koriat (1997), the improvement in resolution occurs because of the increased reliance on internal, mnemonic cues that disclose degree of learning. Further unpublished work by Koriat and his associates with adults suggest that in fact two changes occur with practice studying the same list of items. First, mnemonic cues become increasingly more valid as predictive cues for recall, and second, the reliance on these cues increases with practice. If so, then we have here evidence that a similar reliance on internal cues occurs even in 2nd graders.

The second proposition of Koriat (1997), however, was not supported for children. According to that proposition, the increased reliance on mnemonic cues is accompanied by a decreased reliance on intrinsic cues. As can be seen in Figure 2 (bottom panel), the JOL-difficulty correlation (with difficulty scored dichotomously) actually increased rather than decreased with practice. This increase is evident for 4th graders between the 1st and 2nd presentations only, whereas for the 2nd graders there was a monotonic increase from the 1st to the 4th presentation. In fact, the similarity between the patterns depicted in the two panels of Figure 2 suggests that much of the improvement in the predictive validity of JOLs with practice was mediated by intrinsic cues.

The discrepancy between the children and adult results is interesting. At present we cannot tell whether it discloses a qualitative difference in the bases of JOLs for children and adults. As noted earlier, intrinsic (as well as extrinsic) cues can affect JOLs directly (through an analytic, theory-based process), but they can also affect JOLs indirectly, through their effects on mnemonic cues. Thus, it is possible that in children, much of the inter-item variance in mnemonic cues (e.g., processing fluency) is determined by intrinsic properties such as those captured by the judged a-priori difficulty of the items. In that case the results would be seen to accord with the proposition that practice does result in increased reliance on mnemonic cues even among children. It should be noted that in a post-experimental interview about the strategy used to memorize the pairs, 23% of the 4th graders and 10% of the 2nd graders mentioned reliance on the associative link between the two members of a pair. However, even those who did not mention such strategy indicated that some pairs were easier than others, and chose the strongly-related pairs as an example of the easier pairs. Thus, it is possible that children explicitly used that kind of declarative knowledge in making JOLs on the first presentation of the list. However, the observation that the JOL-difficulty correlation increased with practice suggests that even when young children do not take advantage deliberately of their a-priori knowledge that some items are easier to learn and remember than others, they can appreciate inter-item differences between the items after attempting to learn and remember them, and can then use mnemonic cues in making subsequent JOLs.

Some evidence for reliance on mnemonic cues also comes from the observation that for presentations 2-4, JOLs for different items on one presentation were highly correlated with the recall of these items on the previous test. This correlation was higher for the 4th graders and did not increase with presentation. Thus, for example, JOLs for presentation 2 correlated .76 and .92 for 2nd and 4th graders, respectively, with recall success on the previous test.

2.2 Experiment 2

Experiment 1 yielded some evidence that children's JOLs too exhibit improved predictive accuracy as a result of practice studying the same list of items. Experiment 2 explored one additional factor that has been found to affect JOL accuracy: The elicitation of JOLs immediately after study vs. its elicitation some time after study.

A robust finding that has been repeatedly observed by Nelson and Dunlosky (Nelson & Dunlosky, 1991; Dunlosky & Nelson, 1994, 1997) is the "delayed JOL effect": The accuracy of JOLs in predicting subsequent memory performance is substantially higher when JOLs are solicited some time after study than when they are solicited immediately after study. This effect was only observed when the JOLs were elicited by the stimulus word in the pair; not by the intact stimulus-response pair (Dunlosky & Nelson, 1992).

According to the monitoring-retrieval hypothesis of JOLs (Dunlosky & Nelson , 1997) this is because when JOLs are elicited immediately at the end of the study trial, the item is still in short-term memory and therefore the mnemonic cue associated with attempted retrieval has limited validity in predicting future recall. On the other hand, when JOL is delayed, the mnemonic cues associated with attempted recall tap the kind of retrieval from long-term memory that would be required during testing.

Experiment 2, then, had two aims. The first was to examine whether children's JOLs also exhibit sensitivity to the time at which JOLs are elicited. This possibility has been confirmed recently by Schneider et al. (in press). In their study, children (2nd graders, 4th graders and kindergarteners) made immediate or delayed dichotomous JOLs. Delayed JOLs were found to yield higher JOL-recall gamma correlations than immediate JOLs (.83 and .53, respectively in Study 1, and .75 and .18 in Study 2). These results not only indicate that young children are capable of monitoring their knowledge under favorable circumstances, but also suggest that their JOLs are affected by internal, mnemonic cues.

The second aim was to examine the hypothesis that the process underlying the delayed-JOL effect is the same as that underlying the effects of practice on JOLs. This hypothesis has not been tested so far on either adults or children. We have previously proposed that the improvement in JOL accuracy that occurs with practice derives from both the increased diagnosticity of the mnemonic cues underlying JOLs, and increased reliance on these cues. Similarly, the delayed-JOL effect has been explained in terms of a better diagnosticity of the cues underlying delayed JOLs compared to those underlying immediate JOLs. If so, we should expect an interaction between the effects of practice and the effects of delay so that both of these manipulations can be considered to constitute roughly alternative means to achieve the same goal. Therefore practice should have little effects beyond those that are due to delaying JOLs.

2.2.1 Method

Participants. As in Experiment 1, participants were 32 2nd graders and 32 4th graders. In each group participants were assigned randomly to the stimulus-alone and stimulus-response conditions.

Materials and Procedure. The same list of 24 Hebrew pairs as in Experiment 1 was used. The procedure was similar to that of Experiment 1 except for the following. First, in the stimulus-alone condition JOLs were cued by the stimulus word, whereas in the stimulus-response condition the intact stimulus-response pair was presented as a stimulus for JOLs.

Second, for each participant, the elicitation of JOLs was immediate for 12 items and delayed for the remaining 12 items. The assignment of items to the immediate and delayed JOL conditions was random except that in each condition there were exactly 6 easy and 6 hard items. For the immediate-JOL items the stimulus for JOL appeared immediately at the end of the study trial. For the delayed-JOL items, in contrast, the stimulus for JOLs appeared after all 24 items had been studied. The order of JOL elicitation for these items was such that the first 4 items studied that were assigned to the delayed-JOL condition, appeared first, in random order, then the next four items, and finally the last set of four items studied. Finally, unlike in Experiment 1, no feedback was given.

There were 4 study-test blocks. Participants were instructed about the difference between immediate- and delayed-JOL items, and were given practice with a 6-item list

2.2.2 Results.

Let us consider first the results for the first presentation. Recall was overall better for 4th graders than for 2nd graders in the first study-test cycle (47% and 42%, respectively). Recall was also better for the stimulus-response condition (47%) than for the stimulus-alone condition (41%), and for delayed-JOL (53%) than for immediate-JOL items (35%). However, there was an interaction such that the advantage of the stimulus-response condition over the stimulus-alone condition was found only for delayed-JOL items but not for immediate-JOL items.

With regard to JOLs, an interactive pattern was observed: There was little difference between immediate (3.65) and delayed JOLs (3.70) when JOLs were made in response to the stimulus-response pair. When JOLs were cued by the stimulus alone, in contrast, delayed JOLs were significantly lower (3.33) than immediate JOLs (3.98). These results suggest that delaying JOLs can mend the overconfidence experienced by children during study.

Note the interesting dissociation between the effects of JOL interval (immediate-delayed) on recall and JOLs in the stimulus-alone condition: Delaying JOLs improved recall significantly ($p<.0001$) but reduced JOLs significantly ($p<.0001$). The former effect is consistent with the finding that retrieval experience is more beneficial to recall when retrieval is difficult than when it is easy (Whitten & Bjork, 1977). The latter effect, on the other hand, presumably derives from the greater retrieval fluency that is experienced during immediate JOL compared to that characteristic of delayed JOL.

With regard to the effects of practice, the results were similar to those of Experiment 1, exhibiting increased JOL and recall with practice.

We turn next to resolution, that is, the accuracy of JOLs in predicting inter-item differences in recall. In the analyses of resolution we focused only on the results from the first three presentations because 12 participants exhibited little variance in JOLs on the 4th presentation. In addition, the results for the first 3 blocks were based only on 27 4th graders and 24 2nd graders because the remaining participants also yielded no variance in JOLs on one of these presentations.

First, consider the delayed-JOL effect. The results were consistent with those obtained for adults and with those reported by Schneider et al. (in press) for children: For the stimulus-alone condition the JOL-recall gamma correlation averaged .60 and .92 for the immediate and delayed JOLs in the first block. The respective values across all 3 blocks were .77 and .91. There was no similar difference for JOLs cued by stimulus-response pairs. We should also note that there were no significant age differences in JOL-recall accuracy, unlike what was found in Experiment 1 (see also Schneider et al., in press).

Second, the effects of practice on resolution were generally similar to those obtained in Experiment 1: JOL accuracy improved with practice. However, the results, presented in Figure 3 also disclose the expected interactive pattern between practice and delay. For three of the four conditions the JOL-recall correlation generally increased with practice, averaging .59, .80 and .85 for blocks 1-3, respectively. In contrast, delayed JOLs cued by the stimulus alone yielded a high resolution on the first block, consistent with the delayed-JOL effect, and this high resolution remained stable across blocks. In fact, practice seemed to achieve practically the same level of JOL accuracy as that achieved by delaying JOLs (cued by the stimulus alone). This pattern of results is consistent with the idea that practice and delay constitute alternative means for enhancing JOL accuracy, and that the effects of both are mediated by similar processes, presumably the increased diagnosticity of the mnemonic cues on which JOLs are based. We should stress, however, that because JOL accuracy was very high even on the first block for stimulus-alone delayed JOLs, the pattern of results depicted in Figure 3 could simply stem from a ceiling effect.

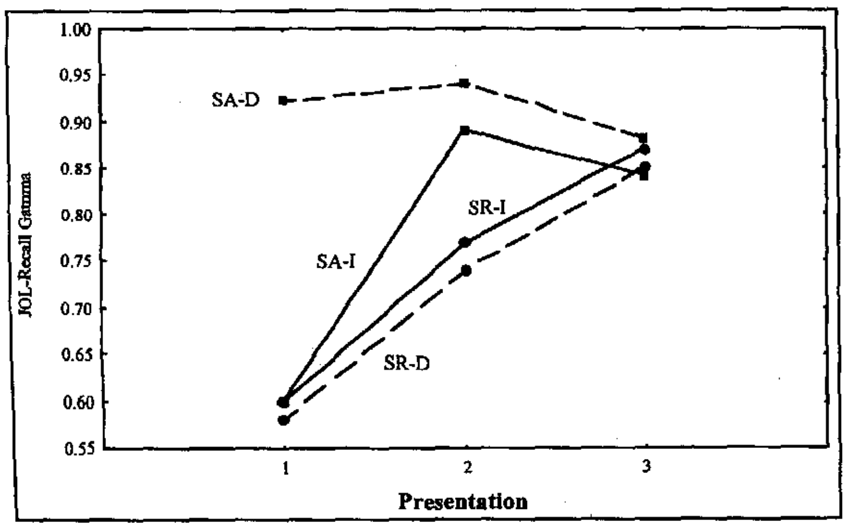

Figure 3. Mean within-subject gamma correlations between JOLs and recall plotted as a function of presentation for immediate JOLs cued by the stimulus alone (SA-I) and by the stimulus-response pair (SR-I), and for delayed JOLs cued by the stimulus alone (SA-D) and by the stimulus-response pair (SR-D) (Experiment 2).

In sum, the present study concerned two factors that have been found to have marked effects on the accuracy of item-by-item JOLs among adults –repeated practice studying the same materials, and the elicitation of JOLs immediately after study or at some delay. The results obtained with adults suggest that the improved monitoring resulting from both practice and delay is mediated by reliance on internal, mnemonic cues that are diagnostic of the extent to which the studied items have been mastered. In this study we obtained results suggesting that children's JOLs are similarly affected by mnemonic cues, and furthermore, that the effects of practice and delay on JOL accuracy are mediated by similar processes. This study, then, provides some insight into the mechanisms underlying children's monitoring of their own knowledge during study, and the processes that contribute to the accuracy of that monitoring.

3. CONCLUDING REMARKS

While the present study concerned metamemory in children, our main interest was not simply to assess children's monitoring proficiency or to examine age-related effects in that proficiency. Rather, the focus of the study was on the processes mediating metacognitive judgments and their accuracy in young children. As noted in the introduction, developmental studies in metamemory have been primarily descriptive and correlational, attempting to identify age differences in metamemory and their possible effects on memory performance. In contrast, the study of metacognition by experimental cognitive psychologists has concentrated more narrowly on testing specific hypotheses about the dynamics of metacognitive monitoring and control processes (e.g., Barnes et al., 1999; Koriat & Goldsmith, 1996; Nelson & Narens, 1990). Several recent studies, however, have attempted to bring insights from cognitive psychology to the investigation of developmental aspects of metacognition, concentrating on the on-line monitoring and control processes that occur during learning and remembering. For example, Butterfield, Nelson, and Peck (1988) applied experimental paradigms that have been in use in the study of adult metacognition to investigate developmental trends in the accuracy of the feeling of knowing. A subsequent study by Lockl and Schneider (submitted), while extending this study, also addressed the question of the basis of feeling-of-knowing judgments in children. With regard to monitoring and control processes during learning, both the Schneider et al. (in press) and the present study extended investigation of the accuracy of JOLs to children, focusing on the on-line monitoring of degree of learning that occurs in item-by-item learning (see Thiede & Dunlosky, 1999). A study by Dufresne and Kobasigawa (1989), however, provides important insight into developmental aspects of the control function. In general, the on-line monitoring of learning is important because it guides the allocation of study time and study effort among different items (Nelson & Leonesio, 1988). What Dufresne and Kobasigawa showed is that the main difference between younger and older children lies in the ability to put the output of monitoring to use in the self regulation of study time. A more recent study by Koriat, Goldsmith, Schneider, and Nakash-

Dura (in press; see also Roebers, Moga, & Schneider, in press) focused on the strategic regulation of memory performance during the report of information from memory. Their results indicate that young children, like adults, are capable of enhancing the accuracy of their memory testimony by screening out wrong answers under free-report conditions. The results also suggest a developmental trend in the level of memory accuracy that can be achieved through the strategic control of memory reporting.

These recent studies, among others, illustrate some of the benefits that ensue from the attempt to merge contributions from cognitive and developmental psychology in the investigation of metacognition. While the cognitive approach to metacognition stresses depth of understanding, the developmental approach offers conceptual breadth and richness. A combination of both approaches is likely to offer interesting and important new venues for investigation.

Authors' note

The experiments reported were conducted at the Institute of Information Processing and Decision Making, University of Haifa. We are grateful to Yaffa Lev for programming the experiments, and to Limor Sheffer for the analyses of the data. Correspondence concerning this chapter should be addressed to Asher Koriat, Department of Psychology, University of Haifa, Haifa 31905, Israel.
akoriat@psy.haifa.ac.il