 DOI: 10.1111/j.1467-9450.2007.00614.x

Cognition and Neurosciences

Underlying mechanisms of initial feelings of knowing in children

NOA R. BESNER[1] and LISA K. SON[2]

[1]*Teachers College, Columbia University, New York, USA*
[2]*Barnard College, New York, USA*

Besner, N. R. & Son, L. K. (2007). Underlying mechanisms of initial feelings of knowing in children. *Scandinavian Journal of Psychology*, *48*, 449–457.

Past research has pointed to two contrasting mechanisms behind feeling-of-knowing judgments (FOKs). The *trace-based* account proposes a direct internal monitor that detects the presence or non-presence of a target item. The *inferential* account dictates that judgments are actually based on cues external to the memory trace, such as familiarity; FOKs could therefore be based on sometimes misleading information. Thus, while a direct mechanism will lead to good learning strategies (i.e. successful searches), an inferential process could potentially lead to less optimal strategies (i.e. search for unknown target information). The question posed in this research was: What processes underlie the judgments children make prior to fully retrieving a target item? We used a simplified version of Reder and Ritter's (1992) game show paradigm to investigate the accuracy (and mechanism) of children's initial FOKs. The experiment found that initial FOKs were largely driven by familiarity of the cues. These data have important implications for strategies children utilize in educational settings.

Key words: Feeling-of-knowing judgments, cue familiarity, trace-based mechanism, metacognitive judgment, children's metacognition, game-show paradigm.

Lisa K. Son, Department of Psychology, Barnard College, 3009 Broadway, New York, NY 10027. Tel: +1 (212) 854-0114; fax: +1 (212) 854-3601; email: lson@barnard.edu; nb2010@columbia.edu

INTRODUCTION

In the last 40 years, ever since Hart (1965) initially introduced the topic, memory monitoring (Flavell, 1979) and feeling of knowing judgments (FOKs) have received tremendous attention from cognitive psychologists. The ability to understand what a person can and cannot currently recall prior to making a judgment about probable future recognition is clearly an important issue – both in terms of the underlying mechanism and its implications for subsequent learning. Two contrasting types of mechanisms have been highlighted in past research. Early findings favored a *direct access* mechanism, implying the presence of an internal monitor drive by the presence of target information or partial information of items in memory (Hart, 1965, 1967; Nelson, Gerler & Narens, 1984; Yaniv & Meyer, 1987). Recent research, however, has proposed *inferential* mechanisms that function on the basis of external cues like *accessibility* (Koriat, 1993) or *familiarity* (Metcalfe, 1993; Metcalfe, Schwartz & Joaquim, 1993; Reder, 1987, 1988; Reder & Ritter, 1992; Schwartz & Metcalfe, 1992). FOKs driven by inferential mechanisms are at risk of being illusionary, and may subsequently lead to ineffective learning strategies, such as *searching-in-vain*, or perhaps even worse, rehearsing an *incorrect* target. The current experiment investigated the mechanisms – *direct* or *inferential* – that second-grade children use when making initial FOKs, and the consequences of the two types of FOKs on test performance.

Hart (1965) introduced the recall-judgment-recognition (RJR) paradigm, which tests the level of accuracy of the FOK in adults, and which has since become a staple in metacognitive research. In the procedure, participants are presented with general information questions that they are subsequently asked to recall. For answers they cannot recall, they are asked to indicate (yes or no) if they feel that they would be able to recognize the answer on a later multiple-choice test. Thus, his definition of a FOK (which has also prevailed over time) was a feeling that one could identify a target item from memory on a future recognition task despite initially failing to remember it on a free-recall task. Memory monitoring was importantly defined as "the intervening process . . . which leads to feelings of knowing and feelings of not knowing" (Hart, 1965, p. 216). The conclusions of Hart's experiments were that the feelings of knowing (and not knowing) were relatively accurate indicators of what was (and was not) in memory storage – participants were more likely to recognize items to which they initially ascribed a positive FOK.

Hart's (1965) data were the first to support a two-stage model: the first stage consisted of monitoring, the second, retrieval. Nelson (1984) described this model as a *direct monitoring* mechanism. According to this model, FOKs (as well as other types of metacognitive judgments such as the tip-of-the-tongue) simply represent blocked memory states (Koriat, 1993). Moreover, because of the two separate processes, the block does not necessarily inhibit the initial monitoring for correct target items. The opposing literature that followed, however, contrasted the notion of direct access and attributed Hart's initial findings to methodological problems. Koriat (1993) explained that Hart did not differentiate between the participants' and the experimenter's perspectives during the experiment. In addition, he only tested FOKs for items the participants could not overtly recall, mistakenly assuming that all items recalled were correctly derived from the target list. Moreover, Koriat criticized the implementation of "yes" or "no" FOK judgments which take for granted that FOKs are an "all-or-none indicator, rather than a graded signal" (p. 610).

To rectify these problems, Koriat (1993) proposed the *accessibility* model, which posited that FOKs are actually derived from *inferential* mechanisms, as they are "computed" (p. 612) on the basis of external cues during early stages of search and retrieval. Therefore, in contrast to the *direct access* model, the *accessibility* model illustrated a *single* process view of memory monitoring, where the FOKs may be based on how accessible (i.e. the amount) partially retrieved information is, *even if it is incorrect*. Koriat (1993) successfully provided evidence for this account in a series of three experiments using four-letter strings as test items. Using a modified RJR paradigm to address the earlier concerns, his results showed that FOKs increased with the amount of partial information available, regardless of its accuracy. Although the participants produced mostly accurate judgments, the data suggested that the use of accessibility mechanisms could lead to imperfect recall and inaccurate feelings.

Researchers have also shown that FOKs may be based on how quickly or easily information is retrieved from memory (Koriat, 1993; see also Benjamin, Bjork & Schwartz, 1998, for a retrieval-based mechanism for judgments of learning). Koriat (1993) wrote that *intensity* (p. 613), or how quickly information is accessed, was a second inferential cue influencing his participants' FOKs. Like the amount of partial information accessed, quick recall latency is generally indicative of correctly retrieved information. Indeed, the data showed that FOKs increased as response times decreased. In short, Koriat (1993) showed that FOKs were largely based on *inferential* mechanisms – the amount of all accessed information, correct or incorrect, and ease of retrieving that information.

Another inferential cue that FOKs may be based on is how familiar the question or the cue is (see Metcalfe, 1993; Metcalfe *et al.*, 1993; Nhouyvanisvong & Reder, 1998; Reder, 1987, 1988; Reder & Ritter, 1992; Reder & Schunn, 1996; Schunn, Reder, Nhouyvanisvong, Richards & Stroffolino, 1997; Schwartz & Metcalfe, 1992). For instance, Metcalfe *et al.* (1993) used a cue-target paradigm in order to manipulate the repetitions, and thereby familiarity, of the cues, without affecting the memorability of the associated targets. The results showed that FOKs are determined by underlying familiarity processes instead of target-retrievability (direct) mechanisms – FOKs increased with the number of repetitions. In their conclusion, Metcalfe *et al.* stressed that when using the *cue-familiarity* heuristic, people "approximate the uncertain quantity indirectly, rather than measuring the quantity itself" which could therefore "result in biases and errors" (p. 860).

Reder and Ritter (1992) also conducted a set of instrumental experiments that lent support to the cue-familiarity mechanism. The crucial difference in their study and other monitoring studies was that participants were forced to make very quick FOKs, as the researchers wanted to investigate the initial judgments people make *before* any type of search is initiated (in contrast to some studies where FOKs were tested only after subjects failed to recall target items). In a *game-show paradigm* using math problems, adult participants were asked to make initial FOKs of their knowledge by quickly selecting to either "*compute*" (which represents a low FOK) or "*retrieve*" (which represents a high FOK) the answer. By using math, Reder and Ritter were able to manipulate familiarity and avoid potential confounds. They could be fairly certain that these math problems were initially unfamiliar to the participants and therefore it would be possible to observe participants' FOKs change over time with repeated trials. Moreover, by using this paradigm, the researchers were able to manipulate familiarity without difficulty. The procedure was as follows: After several trials, when participants were able to retrieve (without computing them) the answers to repeated equations, novel problems were introduced containing certain question parts – operators or operands – from earlier equations. The findings supported the familiarity account: As the familiarity of problem parts increased, so did participants' attempts to retrieve, without computing. If a direct mechanism were in place, participants would have remained unaffected by the familiarity manipulation, and would not have chosen to retrieve answers that they could not possibly have known. As is reviewed below, few studies have tested FOKs in children, and none have tested their initial FOKs in a manner similar to Reder and Ritter. The current experiment used a simplified version of the game show paradigm to test children's initial FOKs, perhaps prior to conducting a full search and retrieval process.

Researchers have also explicitly addressed the differences between the *accessibility* and *familiarity* account (Benjamin & Bjork, 1996; Koriat, 1993; Koriat & Levy-Sadot, 2001, Schwartz & Metcalfe, 1992). According to the cue-familiarity account (Koriat & Levy-Sadot, 2001), FOKs are thought to emerge from a mechanism operating at a "preretrieval stage" (p. 36), which helps in the selection of the retrieval strategy. Thus, FOKs are based on "properties of the pointer" (p. 36) like familiarity or unfamiliarity rather than on features directly related to the target information. On the other hand, the accessibility account posits that FOKs are derived from the "by-products of the retrieval attempt" (p. 36); only after a person decides to initiate retrieval efforts can the accessibility cues (of amount and intensity) influence one's judgments. In an experiment, Koriat and Levy-Sadot also supported that familiarity increases FOKs in very early stages of the search and retrieval process, but that afterwards, accessibility also comes into play. Thus, accessibility is dependent on familiarity – only when familiarity is high will a person initiate a memory search, where then the amount of accessible information will additionally impact the FOK.

As evident, there is a mass of literature dedicated to FOKs and their underlying mechanisms. However, most of the research, and all of the above-mentioned studies, were tested on and refer only to memory monitoring in adults. Surprisingly, there are only a handful of studies about this topic in relation to children (Brown & Lawton, 1977; Carr & Jessup, 1997; Cultice, Somerville & Wellman, 1983; DeLoache & Brown, 1984; Garofalo & Lester, 1985; Lockl & Schneider, 2002; Swanson, 1990; Wellman, 1977). Do children possess metamemory abilities? Past research (Carr & Jessup, 1997) indicates that even in first grade, both boys and girls can demonstrate metacognitive knowledge, even if they sometimes differ on strategy use (for example, in math, girls use more backup overt strategies, while boys use more retrieval strategies). Even educable mentally retarded children (Mental Age 8 and up) who cannot engage in complicated metacognitive strategies (like predicting recall readiness or efficiently planning study time) have been shown to display FOK abilities (Brown & Lawton, 1977). In fact, while most of the limited developmental studies use children aged 6 and up, DeLoache and Brown (1984) have demonstrated the early appearances of metamemory in infants as young as 18–30 months old. After the experimenters secretly moved a previously known location of a hidden toy, infants showed greater persistence searching for that toy than after simply making an overt mistake, thus showing support for early feelings of certainty. Moreover, the older infants (27–30 months) seemed to consider the possibility of alternative reasons behind the misplaced toy, other than their failed memory; they searched areas related to the initial hiding spot more so than the original location.

Metacognition is a growing area in developmental psychology, especially as it relates to enhancing children's study abilities (Brown & Palincsar, 1989; Palincsar, 1998; Palincsar & Herrenkohl, 1999; Palincsar & Brown, 1984, 1989). After manipulating aptitude and metacognitive ability in a lab setting, Swanson (1990) found that regardless of aptitude, children with higher metacognitive abilities (e.g. hypothetico-deductive and evaluation strategies) perform better than those with lower metacognitive abilities. Garofalo and Lester (1985) stressed the importance of incorporating metacognition into educational practices in order to improve children's problem solving abilities.

While there is a general consensus that metacognition aids children's problem solving abilities, the current research sought to investigate the underlying processes of their individual feelings of knowing. Children's FOKs and the potential accuracy or inaccuracy of these judgments (depending on the underlying processes) can have crucial effects for children's initial search efforts in problem solving situations. When children are beginning to learn how to implement study strategies and approach problems in a sophisticated manner, it is important to consider how their initial FOKs affect these consequent study behaviors. While FOKs do not control study strategies per se, they can affect what a child can and cannot remember. Specifically, if children's FOKs are based on an internal *direct* monitor of the memory store, then educators should encourage children to trust this heuristic when trying to retrieve answers to problem. However, if children's FOKs are grounded in *inferential* mechanisms (as is generally accepted for adults), then these judgments could potentially invoke misleading feelings of certainty about the correctness of information that is "remembered" that might, in reality, be incorrect. Below, we summarize some of the FOK findings regarding children.

Wellman (1977) was the first to investigate the phenomenon in children, reasoning that feelings of knowing are developmentally driven. He tested kindergarten, first-, and third-graders by giving them different items to name. For the items they failed to name, the participants were asked to provide FOKs (yes or no) in relation to future recognition, and if they had seen the picture before. The experimental findings revealed that the third graders' FOKs were the most accurate, as their ability to access information was most developed. Cultice *et al.* (1983) studied the metamemory abilities of preschoolers (aged 4–5) to see if similar results would occur. To ease the experiment for younger participants, they modified the paradigm to incorporate photographs of people (manipulated for 3 levels of familiarity). The methodological adjustment was successful, as the preschoolers seemed to show an understanding of FOKs and produce relatively accurate judgments. Importantly, however, there was a higher correspondence between *seen* judgments and FOKs than between *not seen* judgments and FOKs (showing evidence for a familiarity effect). In recent research, Lockl and Schneider (2002) specifically examined 7-, 8-, 9-, and 10-year-old children's FOKs with the purpose of determining whether the underlying processes follow the traditional *trace-based* view or if they follow Koriat's newer *trace-accessibility* model. Children were given vocabulary tests where they defined certain words. After participants failed to define 30 words, or gave incorrect definitions, they were asked to make FOK judgments and relative FOK judgments before completing recognition tests. The findings showed no accuracy differences among the age groups, and supported the accessibility view. The accessible amount of partial information that came to mind determined their FOKs, rather than the direct access to the knowledge in the memory store.

The above three studies represent the limited span of research conducted on children's metamemory and FOKs. A crucial topic is lacking within these studies; there is no discussion of the *initial* judgments children make before committing to a response. All of the previous children's studies (like the majority of adult studies) have investigated the FOKs children make *after* they cannot recall an answer. What happens in those first few seconds after children hear a question that leads them to either deliver a response (and believe it is correct, regardless of accuracy) or claim uncertainty? Do their answers reflect an actual memory scan (via *direct* mechanisms) or do they quickly respond before they complete a retrieval process (via *inferential* mechanisms of familiarity)?

The current study implemented a simplified version of the Reder and Ritter (1992) paradigm described earlier to investigate how children know when they know. The Reder and Ritter article proposed a unique research question: "Is it necessary to search memory in order to ascertain whether an answer is likely to be known?" (p. 435). By forcing participants to decide to guess or compute the math answer within 850 ms, the researchers were certain that the participants' choices were based on feelings of knowing (which they defined, contrary to others, as a feeling of knowing when you know) rather than on completed retrieval processes. Moreover, by inhibiting the search and retrieval process, Reder and Ritter were able to precisely compare the cue-familiarity (inferential) account with the direct-access account. According to the Koriat and Levy-Sadot (2001) findings, this time restriction inhibited an interaction between cue familiarity and accessibility processes from taking place. As Reder and Ritter found that participant's initial FOKs were influenced largely by familiarity and therefore sometimes led to misleading feelings of certainty, the present experiment investigated the comparable role of familiarity on children's initial FOKs.

METHOD

The purpose of this study was to investigate the role of *cue-familiarity* in children's initial FOKs. A simplified version of Reder and Ritter's (1992) game show paradigm was employed to test initial FOKs. Children were presented with math problems varying in familiarity and given the option to choose between *guessing* the answer (representing a high FOK based on familiarity) and *calculating* the answer (representing a low FOK based on having to calculate the answer). The central question was to see whether children's rate of guessing increased as manipulated familiarity increased.

Participants

Participants were 19 second graders recruited from a local public elementary school. The children participated in the study as part of an after-school program in their own school building – which has been proposed to improve learning in general (see Kahne, Nagaoka, Brown, O'Brien, Quinn & Thiede, 2001). In order to participate, each child's parent read and signed a consent form explaining all methods, procedures, benefits, and risks of the study, adhering to APA guidelines. Their ethnic backgrounds were as follows: 9 Hispanic, 5 African-American, 4 Caucasian-American, and 1 Asian-American child. All were either 6 or 7 years of age.

Materials and design

The experiment used a PowerPoint program for both the study and test phases. The study program displayed 32 different math equations, one at a time, at various repetition rates (0, 1, 4, 10). Eight equations were each repeated zero times, once, four times, or ten times. Four different versions for this study phase were created to vary the sequence of the equations. Each equation was displayed for a total of three seconds. Sixteen (half) of these equations were addition problems and sixteen were subtraction problems. Also, while some problems consisted of single/double digit mixes (i.e. 28 + 7); others consisted of two double-digit numbers (17 + 11).

Each equation slide contained a numerical problem with its answer (28 + 7 = 35), as well as a graphical depiction. For instance, for the example above, there were 28 ladybug images below the digit "28", and seven baseballs under the digit "7". The pictures were presented to maximize the effects of familiarity.

The test phase consisted of a similar PowerPoint program, but now equations (and pictures) were displayed without their corresponding answers. While half of the test equations were the same as those from study, the other half of them were different. Sixteen were manipulated for familiarity by having their operant signs reversed to the opposite operation – addition signs became subtraction signs, and subtraction became addition. For instance, the study equation "21 – 9 = 12" was transformed into "21 + 9 = ?" during the test. Two test versions were used in order to randomize the display order. Therefore, there were two between-subjects factors: Study version (1, 2, 3, 4) and Test Version (1, 2) and two within-subjects factors: Repetition Familiarity (0, 1, 4, 10) and Operator (Original, Reversed).

Procedure

Each participant was randomly assigned to one of four study versions. He or she was seated in front of a computer screen and told that math problems (accompanied by pictures) would appear one at a time. The participant was instructed to pay attention to the numbers and pictures and try to remember as much as possible. Some of the equations were presented only once, others presented four times, and still others presented ten times. A fourth set of equations – the zero condition – was not presented during the study phase, but was presented later during the test phase as a control comparison.

After the study phase was completed, the child was randomly selected into one of the two test versions. As each equation was displayed, the participant had to indicate *as quickly as possible* whether he or she wanted to *guess* or *calculate* the answer. They were instructed to guess if they believed they knew the answer – high FOK – but to calculate if they did not know the answer – low FOK. Although there was no cut-off time for making their guess/calculate decision as there was in the Reder and Ritter study, all children were told that they had to make their choices *immediately*, and understood these instructions clearly. Their understanding of the difference between "guessing" and "calculating" was also good: *It was apparent that when children chose to "calculate" their answer, they needed to work out the answer using paper and pencil step by step. This was not the case for "guess" choices*. Paper and pencil was provided for calculation. After choosing to guess or calculate, participants stated their actual answer as soon as they could. There was no time limit for responding with their answer. The experimenter running the session recorded the following: (i) reaction times for deciding to guess or calculate, (ii) whether the participant chose to guess or calculate, (iii) the correctness of the answer, and (iv) the retrieval times for stating the answer. All reaction times were recorded with a stopwatch.

RESULTS

Analyses were conducted on three components: Probability of guessing, accuracy in test performance, and retrieval response time (for both guessing and calculating) during test. For all analyses, a probability level of $p < 0.05$ was used as the criterion for statistical significance. Bonferroni-Dunn t-tests were used for *post-hoc* comparisons.

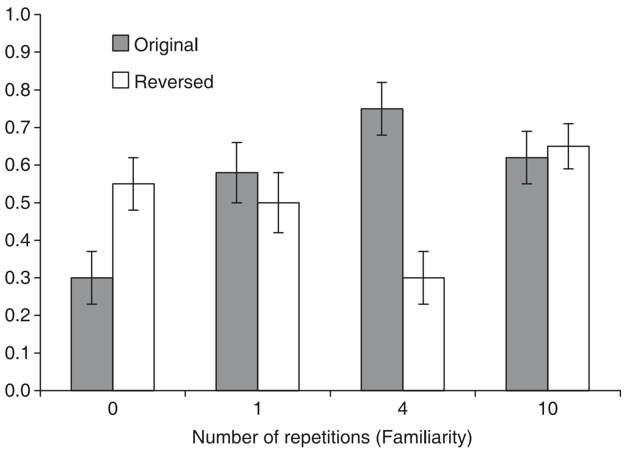

Fig. 1. The mean proportion of guessing during test for Original and Reversed problems following 0, 1, 4, or 10 repetitions.

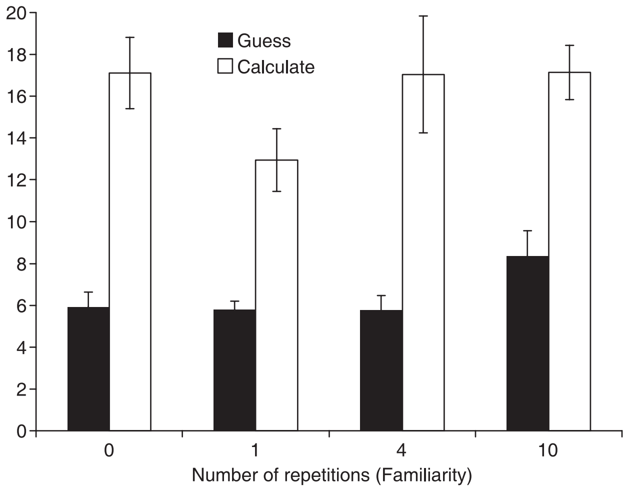

Fig. 2. The mean response times for retrieving the answers for Original problems following 0, 1, 4, or 10 repetitions.

Probability of guessing

The main analysis that spoke to the question of whether children's FOKs were based on familiarity was to see how often they chose to guess the answer as a function of repetition, or familiarity. We conducted an ANOVA with Study Version, Test Version, and Familiarity as factors and guessing rate as the dependent measure. There was no significant effect of Study Version or Test Version. Results showed, however, that there was a significant effect of Familiarity for both the items with the Original operator [$F(3, 33) = 6.50$, $MSE = 0.47$, $p < 0.01$] and the Reversed operator [$F(3, 33) = 5.71$, $MSE = 0.40$, $p < 0.01$]. However, as is shown in Fig. 1, the main effects showed different patterns. For the Original items, paired *t*-tests showed that the probability of guessing increased as familiarity increased. For items that were not studied (repetition = 0), the probability of guessing was 0.30, which was significantly lower than the probability of guessing at repetitions of 1 [$M = 0.58$, $t(18) = 3.63$, $p < 0.01$], repetitions of 4 [$M = 0.75$, $t(18) = 4.54$, $p < 0.01$], and repetitions of 10 [$M = 0.62$, $t(18) = 4.44$, $p < 0.01$]. There were no significant differences among the highest three repetitions. For the Reversed items, the lowest probably of guessing rate occurred for items presented once ($M = 0.50$) and those presented 4 times ($M = 0.30$), and the two were not different from each other. This latter probability was significantly lower than the guessing rate at repetition 0 [$M = 0.55$, $t(18) = 2.85$, $p < 0.01$]. However, at repetition 10, the guessing rate significantly jumped up again [$M = 0.65$, $t(18) = 3.89$, $p < 0.01$].

Of those items that were presented, items that were Original and presented 4 times had a significantly higher guessing probability than items that were Reversed and presented 4 times [$t(18) = 5.93$, $p < 0.01$]. With this one exception, the guessing probability for Reversed and Original items was not different, suggesting that the guessing rate remained high, even for problems that were in fact new (due to the switched operator), but "seemed familiar".

To summarize, the children's guessing rate, or feelings of *knowing* the answer, for both old and new problems was extremely high (averaging approximately 40% for items that had never been studied at all), and increased as the familiarity increased (at rates of 60% for items repeated ten times). In addition, when averaged across the Repetition factor, there were no differences between the groups – children were as likely to guess on the Reversed items as they were to guess on the Original items. These results support the notion that how familiar the cue was had some effect on children's judgments.

Retrieval response times

The response times for choosing to either "guess" or "calculate" the answer to the math problems were 4.14, 3.35, 3.66, 4.64 (for Original problems), and 3.89, 4.32, 4.34, and 4.36 (for Reversed problems), for 0, 1, 4, and 10 repetitions, respectively. (We explore in the Discussion section the challenge in defining what "initial" means when we using the term "initial judgment" given that the mean choice reaction times here were longer than those found for adults in Reder and Ritter's (1992) experiment.) Figures 2 and 3 display (for 0, 1, 4, and 10 repetitions) the retrieval times for "guess" and "calculate" trials. As can be seen, the time it took for children to respond with their guesses (5.94, 5.81, 5.77, 8.36 for Original; 6.30, 8.04, 6.86, 7.30 for Reversed) was much faster than the amount of time it took for them to calculate (17.11, 12.94, 17.04, 17.13 for Original; 17.68, 18.99, 20.89, 20.36 for Reversed) the answers. This difference in time was significant for both the Original items [$F(1, 17) = 19.28$, $MSE = 949.87$, $p < 0.01$] and for the Reversed items [$F(1, 17) = 19.28$, $MSE = 949.87$, $p < 0.01$], suggesting that indeed, the children understood the difference between "guess" and "calculate", and took much longer to answer when they chose to calculate.

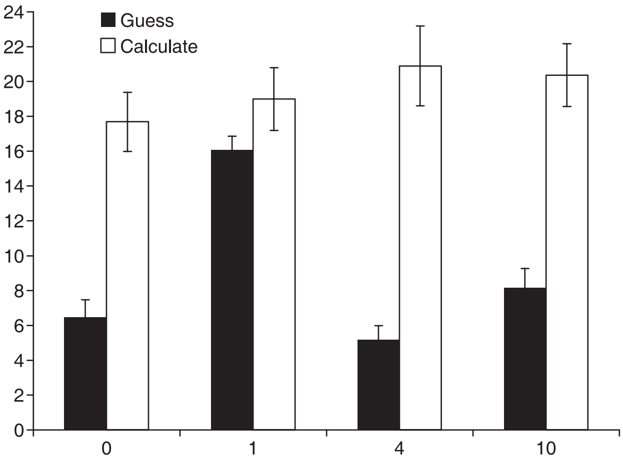

Fig. 3. The mean response times for retrieving the answers for Reversed problems following 0, 1, 4, or 10 repetitions.

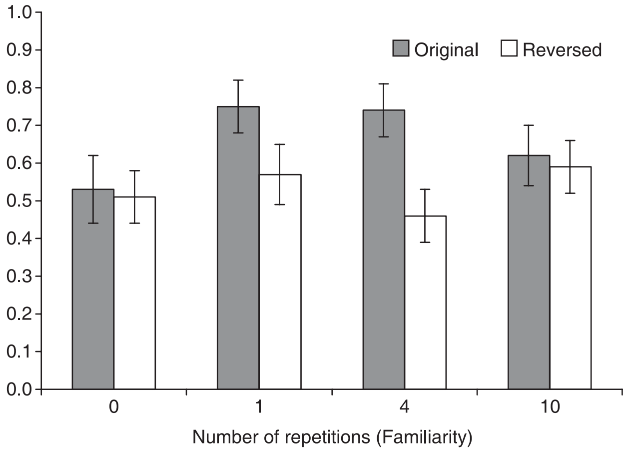

Fig. 4. The mean proportion of correct answers for Original and Reversed problems following 0, 1, 4, or 10 repetitions.

Collapsing across both guessed and calculated problems, the retrieval times for answering Original items *decreased* with increasing repetition rate – ranging from 14.32 s at Repetition 0 to 11.14 s at Repetition 10 [$F(3, 33) = 8.76$, $MSE = 154.57$, $p < 0.01$]. The retrieval rates for the Reversed items, however, revealed an opposite trend – retrieval times *increased* with repetition [$F(3, 33) = 6.86$, $MSE = 89.13$, $p < 0.01$], particularly within the range of Repetition 0 to Repetition 4. These data suggest that there might be a partial interference mechanism taking place for incorrect feelings of knowing (for the Reversed items) that somehow need to be offset, and are during the retrieval process.

Test performance

Given that children's initial FOKs may be driven by feelings of familiarity – accurate *or inaccurate* – it is critical to see whether inaccurate judgments may lead to inaccurate answers. Given that the response times for *guessing* were fast (and were, actually, based on "guessing" and not "calculating"), then it should also follow that the performance for Original problems should be higher than for Reversed problems. This was indeed the case (see Fig. 4). For Original problems, the mean performance levels were 0.53, 0.75, 0.74, 0.61; for Reversed problems, the mean performance levels were 0.51, 0.57, 0.46, 0.59 (for 0, 1, 4, and 10 repetitions, respectively). When collapsing across number of Repetitions and *conditionalizing on only those items that were selected for guessing*, proportion correct on Original items ($M = 0.38$) was marginally higher than the proportion correct on Reversed items ($M = 0.27$) [$t(18) = 2.99$, $p = 0.06$].

As before, we also conducted an ANOVA with Study Version, Test Version, and Repetition as factors, and found that for the Original items, number of Repetitions had a significant effect [$F(3, 33) = 4.55$, $MSE = 0.186$, $p < 0.01$], such that as Repetition increased, so did performance. However, Repetition did not have a significant effect on the Reversed items – children's performance did *not* increase as repetition increased, presumably because the operator was different (and the children had been *misled*). Instead, there were no differences. There were no Study or Test Version effects. Paired *t*-tests showed that for the Original items, lowest performance occurred at Repetition of 0 ($M = 0.53$), and this performance score was significantly lower than that of Repetition 1 [$M = 0.75$, $t(18) = 2.39$, $p < 0.01$], Repetition 4 [$M = 0.74$, $t(18) = 2.58$, $p < 0.01$], and marginally lower than that of Repetition 10 [$t(18) = 1.93$, $p = 0.07$].

In summary, children's performance levels increased as Repetition increased, but only for those items whose operators were *not* reversed (where even a high rate of guessing would not harm performance). When the operators were reversed, however, performance did not increase, presumably because guessing was driven by illusions of knowing the incorrect answer.

DISCUSSION

Before delving into the main discussion of the findings, it is important to elaborate on the assumed definition of "initial" judgments. We have assumed that choosing to guess or to calculate reflected children's initial feelings of knowing, that is judgments made prior to any search having occurred. Although Reder and Ritter (1992) explicitly listed 850 ms as the cut-off for initial feelings of knowing, their participant pool consisted of adults. The methodology of this experiment was designed for children, who took about 3–4 s to make their guesses. While it is impossible to fully confirm that this early decision reflected a "pure" initial judgment before any retrieval has taken place (it is impossible to know this even in the case of adults at 850 ms), we are fairly confident that the children's initial judgments were retrieval-free given the following points: First of all, it is much more difficult to force a child to respond extremely quickly than it is an adult, and thus no cut-off time can be placed. Second, we are not sure that a cut-off period should be placed given that a child's cognitive and metacognitive processes are likely to be significantly slower than those of adults. And most importantly, during the procedure, we saw that each child was entirely dependent on paper and pencil in order to *begin* solving a given math problem during the calculation period. When they did use paper and pencil, each child always wrote out the *entire* math problem before beginning the steps of solving the problem. On the other hand, when a child chose to guess (presumably having high initial FOKs), external materials were never required. The fact that the children always used paper and pencil after deciding to calculate, and never when choosing to guess, increased the chances that the experimental design accurately tapped children's *initial* FOKs – those quick FOKs defined as judgments prior to initiating a search process.

Given our confidence in our measure, the main finding in the current experiment is that children's initial FOKs are driven more so by *cue familiarity* than by *direct access* mechanisms. Using Reder and Ritter's (1992) measure of guessing as a strong FOK, we found that children were more likely to *guess* the answers (as opposed to *calculate* the answers) to items that were more familiar due to a higher number of repetitions. Ideally, if FOKs were driven by only direct-access mechanisms, then children should have chosen to calculate all items that were Reversed, and therefore "novel" in the sense that they were not presented during the experimental paradigm.[1] Given that guessing rates were just as high for Reversed problems as they were for Original problems, illusions of knowing occurred, unfortunately resulting in lowered performance.

The data also showed that children responded quickly to the zero- and ten-repeated Reversed problems, and more slowly for the one-time and four-time repeated questions. This contrasts the initial hypothesis whereby false familiarity would increase response time. It was predicted that children would experience the greatest false feelings of certainty for the highly repeated "familiar" (manipulated) problems, and therefore answer quickly. Nevertheless, the actual results can be explained through the two-process account of FOKs (Reder & Ritter, 1992). In the initial stage of metacognition, FOKs are based on familiarity and can contribute to false feelings of certainty. However, the second stage consists of retrieval, during which participants may recognize their mistake. The additional time needed to correct these errors can explain the slower retrieval times found for the Reversed problems.

These initial-FOK results contrasted a direct mechanism explanation. Children did not realize that they were missing the answers to the new problems and that calculation was necessary. Rather, they showed evidence for an inferential account as judgments were based on external cues – in this case, familiarity. Participants were misled into thinking that more familiar problems (old and new) were better learned, therefore displaying misleading feelings of certainty and choosing to guess more often. Indeed, the highest proportion of guessing for new problems was found in the ten-time repeated subset.

Metamemory and feelings of knowing are fascinating abilities that come into play in our daily lives. Knowing what we know is crucial to our self-improvement, as understanding and remembering our past experiences allows us to expand ourselves through future learning. Nevertheless, as the majority of the past research implies, our feelings of knowing are based on inferential mechanisms, which cannot always be trusted to produce accurate results. Yet, as Koriat has noted, just because feelings of knowing are influenced by familiarity and accessibility, it does not mean that all FOKs that come to mind should be disregarded. Quite on the contrary, in fact, most of our FOKs are accurate (as explained earlier in the research) because of the way our memories are organized and the time it takes us to recall information. Nevertheless, while adults *should* pay attention to their initial FOKs prior to initiating a search and retrieval process (in terms of the familiarity account; or pay attention to partial information, as per the accessibility account), they should always remember, in the back of their minds, that this information *could* potentially be misleading.

As seen with the present study, this line of reasoning in relation to children is far different. It is very difficult for children to grasp the notion that their initial FOKs are not based on direct truths, but rather on external cues. It is hard for children to view their initial feelings of knowing as a helpful heuristic that should be consulted, but not taken for granted. Teachers must try to explain this concept to their students in order to prevent children from trusting potentially misleading feelings of certainty. Perhaps it would be helpful to teach students an array of metacognitive strategies to supplement their problems solving attempts. For instance, teaching them retrospective strategies to balance their prospective judgments (like FOKs) could be quite beneficial. All of the above could be possible with grade school students.

Yet, with very young students like kindergarteners, who are capable of FOKs (Wellman, 1977), it might be difficult for their educators to convince them to be wary of these judgments. Moreover, these young children would probably be quickest to attribute a FOK to a test item that is highly accessible or familiar. Kindergarten teachers should therefore initially force students to take time to think about question answers before they actually respond. For instance, young students should listen to a question and be given a few seconds to mull over it before physically writing down the answer. This will encourage young children to get into the habit of consulting, but not always trusting, their initial FOKs, which could be misleading.

More research, like the present experiment, is needed to understand the role of initial FOKs for children's metacognition.

Future work gathering more support for the underlying mechanisms of children's initial FOKs is essential. Another area requiring exploration is the testing of possible solutions to the problems mentioned above. How can children be taught to best implement their feelings of knowing without relying on them too heavily to lead to accurate retrieval? Proper use of metacognitive strategies has been repeatedly proven to be associated with higher aptitude students. Thus, the answer to this previous question can benefit classroom settings. As evident from the literature and the present study, what started with a simple set of experiments by Hart in 1965 has developed into nearly 40 years of extensive research, with endless possibilities for future study – particularly in classroom settings – in sight.

We thank the teachers, administration, students, and principal Robert O'Brien, of P.S. 75, the Emily Dickinson School, for their participation. This research was funded by the Spencer Foundation.

NOTE

[1] Although children are exposed to math problems in school, it is unlikely that they would "memorize" the answers to math problems, particularly those that consist of two-digit numbers.

REFERENCES

Benjamin, A. S. & Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. M. 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.

Brown, A. L. & Lawton, S. C. (1977). The feeling of knowing experience in educable retarded children. *Developmental Psychology*, *13*, 364–370.

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

Brown, A. L. & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 393–451). Hillsdale, NJ: Lawrence Erlbaum.

Butterfield, E. C., Nelson, T. O. & Peck, V. (1988). Developmental aspects of the feeling of knowing. *Developmental Psychology*, *24*, 654–663.

Carr, M. & Jessup, D. L. (1997). Gender differences in first-grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology*, *89*, 318–328.

Cultice, J. C., Somerville, S. C. & Wellman, H. M. (1983). Preschoolers' memory monitoring: Feeling-of-knowing judgments. *Child Development*, *54*, 1480–1486.

DeLoache, J. S. & Brown, A. L. (1984). Where do I go next? Intelligent Searching by Very Young Children. *Developmental Psychology*, *20*, 37–44.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist*, *34*, 906–911.

Garofalo, J. & Lester, F. K. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education*, *16*, 163–176.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*, 208–216.

Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior*, *6*, 685–691.

Kahne, J., Nagaoka, J., Brown, A., O'Brien, J., Quinn, T. & Thiede, K. (2001). Assessing after-school programs as contexts for youth development. *Youth and Society*, *32*(*4*), 421–445.

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. & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feeling of knowing. *Journal of Experimental Psychology*, *27*, 34–53.

Krinsky, R. & Nelson, T. O. (1985). The feeling of knowing for different types of retrieval failure. *Acta Psychologica*, *58*, 141–158.

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

Metcalfe, J. (1993). Novelty monitoring, metacognition, and control in a composite holographic associative recall model: Implications for Korsakoff amnesia. *Psychological Review*, *100*, 3–22.

Metcalfe, J., Schwartz, B. L. & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology*, *19*, 851–861.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing Predictions. *Psychological Science*, *2*, 267–270.

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

Nelson, T. O., Dunlosky, J., Graf, A. & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, *5*, 207–213.

Nelson, T. O., Gerler, D. & Narens, L. (1984). Accuracy of feeling-of-knowing judgments for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General*, *113*, 282–300.

Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125–173). San Diego, CA: Academic Press.

Nhouyvanisvong, A. & Reder, L. M. (1998). Rapid feeling-of-knowing: A strategy selection mechanism. In V. Y. Yzerbyt, G. Lories & B. Dardenne (Eds.), *Metacognition: Cognitive and social dimensions* (pp. 35–52). London: Sage.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, *49*, 345–375.

Palincsar, A. S. & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*(*2*), 117–175.

Palincsar, A. S. & Brown, A. L. (1989). Classroom dialogues to promote self-regulated comprehension. In J. Brophy (Ed.), *Advances in research on teaching* (pp. 35–72). Greenwich: JAI.

Palincsar, A. S. & Herrenkohl, L. R. (1999). Designing collaborative contexts: Lessons from three research programs. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning: The Rutgers Invitational Symposium on Education Series* (pp. 151–177). Mahwah, NJ: Lawrence Erlbaum.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, *19*, 90–138.

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

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

Reder, L. M. & Schunn, C. D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 45–78). Hillsdale, NJ: Erlbaum.

Ryan, M. P., Petty, C. R. & Wenzlaff, R. M. (1982). Motivated remembering efforts during tip-of the-tongue states. *Acta Psychologia*, *51*, 137–147.

Schunn, C. D., Reder, L. M., Nhouyvanisvong, A., Richards, D. R. & Stroffolino, P. J. (1997). To calculate or not to calculate: A source activation confusion model of problem familiarity's role in strategy selection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 3–29.

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

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

Schwartz, B. L. & Smith, S. M. (1997). The retrieval of related information influences tip-of the-tongue states. *Journal of Memory and Language*, *36*, 68–86.

Swanson, L. H. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology*, *82*, 306–314.

Wellman, H. M. (1977). Tip of tongue and feeling of knowing experiences: A developmental study of memory monitoring. *Child Development*, *48*, 13–21.

Yaniv, I. & Meyer, D. E. (1987). Activation and metacognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 187–205.

Received 26 August 2006, accepted 12 April 2007