Journal of Experimental Psychology,
Learning, Memory and Cognition, (in press).

Accepted November 2002

The dynamics of metacognitive judgements: Pre and post retrieval mechanisms

David Vernon[1] and Marius Usher[2]

[1]Dept. of Cognitive Neuroscience & Behaviour
Imperial College,
St Dunstan's Road
London
W6 8RP
United Kingdom
Email: D.Vernon@ic.ac.uk
Tel: 44 (0) 208 383 0584
Fax: 44 (0) 208 846 1670

[2]School of Psychology
Birkbeck College,
University of London
Malet Street,
London
WC1 E7HX, UK
M.Usher@bbk.ac.uk
44 (0) 207 631 6201.
44 (0) 207 631 6312

Abstract

Two experiments were conducted to investigate the temporal course of pre- and post-retrieval mechanisms using a new kind of dynamic metacognitive judgment that has not been reported previously. In experiment 1, participants were presented with primed and unprimed triples of remote associates to a target word (RAT test) and required to provide repeated metacognitive judgements, at four times during a 12 sec interval, about the likelihood they will recognise the target at the end of the interval. Both, familiarity with the words (priming) and the processing time, were associated with changes in metacognitive evaluations. In experiment 2, pre- and post- retrieval mechanisms were placed in opposition, by transforming a critical element of a previously primed question. This led to high initial ratings, which subsequently decreased over time. In contrast the metacognitive ratings of unprimed stimuli were initially lower and increased with time. We discuss these results in terms of pre- and post-retrieval mechanisms interacting over time.

When tested on a memory target that they cannot recall, participants in experimental studies are often able to provide a judgement about their ability to recall or recognize the target at a later stage. This feeling of knowing (FOK) is proposed to act as a metacognitive process providing information on stored knowledge, guiding retrieval (Barnes, Nelson, Dunlosky, Mazzoni & Narens, 1999; Koriat & Levy-Sadot, 2001; Nelson, 1996; Nelson, Gerler & Narens, 1984; Nelson & Narens, 1990; Reder, 1987; Schunn, Reder, Nhouyvanisvong, Richards & Stroffolino, 1997). While FOK judgements are typically obtained after memory failure, it is remarkable that preliminary FOK judgements (e.g., Nelson & Narens, 1990) can be made very fast within about 850 ms of a query (Reder and Ritter, 1992, Reder, 1987). The FOK and the *prediction of knowing* (POK) (Schreiber, 1998) are among a variety of metacognitive judgements, some of which unfold over longer time intervals (e.g., feelings of warmth [Metcalfe, 1986b] and confidence ratings) and are instrumental for the control of cognitive processing (Nelson, 1996; Barnes, Nelson et al., 1999). The purpose of the present research was to study the time course of processes mediating such metacognitive judgements. This should enable us to evaluate the contributions of pre-retrieval and post-retrieval mechanisms that are often contrasted in theories of metacognition (Koriat, 1993, 1995, 1996; Koriat & Levy-Sadot, 2001; Maki, 1999; Metcalfe, 1986a; Reder, 1987; Reder & Ritter, 1992).

In order to examine how metacognitive judgements change over time we utilised a method based on the principles of psychophysics, in which each participant is required to make a sequence of judgements about the likelihood to recognise/recall the target at the end of the trial, at four different times during the unfolding of the task in each trial (and enough trials are presented to obtain consistent trends for each participant). This method enabled us to track the metacognitive judgements online utilising a new judgement we termed the *dynamic prediction of knowing* (dPOK) (after Schreiber, 1998; Schreiber & D. Nelson, 1998). Consistent with the prediction of knowing (POK) outlined by Schreiber (1998), the dPOK requires participants to provide a fast initial judgement about their likelihood to recall the target in a subsequent test, immediately after the presentation of a cue on all items, rather than after unrecalled items only. However, as in the feeling of warmth (FOW) method (Metcalfe, 1986b), where participants are required to track their perceived closeness to a solution over time, the dPOK also requires participants to make a sequence of successive judgments (these judgements however, involve the prediction of successful recall).

Tracking the temporal course of metacognitive judgements should enable us to evaluate the contributions of pre-retrieval and post-retrieval mechanisms thought to mediate such processes. For instance, the fast initial dPOK made at time 1 resembles the fast metacognitive judgement used by Reder (1987), which provide support for a pre-retrieval mechanism based on *cue-familiarity*. Strong support for the cue-familiarity account comes from research showing a direct relationship between metacognitive judgements and the number of times the cues (but not the target) are primed prior to testing (Metcalfe, Schwartz & Joaquim, 1993; Reder, 1987; Reder & Schunn, 1996; Reder & Ritter, 1992; Schunn, et al., 1997; Schwartz & Metcalfe, 1992). The latter dPOK judgements resemble both the FOW (Metcalfe, 1986b) and later FOK ratings (Koriat, 1993, 1995, 1996), which are likely to be influenced by post retrieval mechanisms, such as the accessibility[1] of (correct or incorrect) information generated during the retrieval (Koriat, 1993, 1996) or the competition of retrieved information (Maki, 1999; Schreiber, 1998).

If metacognitive judgements are dynamic and change over time, we should expect that while the initial dPOK judgement (immediately after cue presentation) is influenced mainly by the familiarity with the cue, as more information is activated during the memory search, the dPOK evaluation changes. The proposal of a combined dynamic mechanism driving the metacognitive process is supported by research showing that *both* the number of cue presentations and the amount of information retrieved can influence metacognitive performance (Metcalfe, 1993). This is consistent with recent research suggesting that the familiarity of the cue and the accessibility of the target operate together to produce higher metacognitive judgements (Koriat & Levy-Sadot, 2001). By examining dPOK judgements made over time, it should thus be possible to clarify the interplay between cue familiarity and accessibility of target information.

[1] In the most recent version of this account Koriat and Levy-Sadot (2001) make a distinction between fast automatic and slow effortful (or analytic) retrieval stages. It is beyond the scope of this study to distinguish between the automatic accessibility and the cue-familiarity contributions to FOK, which are both assumed to be fast and automatic. In the following, we refer only to the slow retrieval component of the accessibility model, which we contrast to the fast process triggered by cue-familiarity.

Experiment 1

The experiment utilised Mednick's (1962) remote associate task (RAT), in which participants are given a triple of remote associates and required to produce a fourth word related to *all* of the three words in the triple. The RAT was chosen as it generates a slow search process that is convenient for temporal tracking. Familiarity was operationalised by having participants complete a pre-dPOK recognition task using a set of triples that were used again in the subsequent dPOK judgement task. After the presentation of the RAT problem, the participants were required to make four temporally successive dPOK judgements, concerning their ability to identify the correct associate in a subsequent recognition task. Each of the four judgements was prompted by a temporal cue (at 2, 4, 8, and 12 seconds[2]). In order to minimize the involvement of post-retrieval processing in the first dPOK response, we followed Reder and Ritter (1992) by imposing a deadline of 850 ms for participants to enter their response following the first cue. After entering the four dPOK judgements participants were required to complete a three alternative forced choice (3-AFC) recognition task. We predicted that participants' dPOK levels would be higher at time 1 for the primed-familiar stimuli in comparison to the unfamiliar items. In addition to this, as time to search memory increased, we expected participants' dPOK levels to increase for *both* familiar and unfamiliar stimuli. Finally, over time, as more target information is activated and retrieved, we also expected the association between dPOK judgements and accuracy in the recognition task to increase.

METHOD

Participants

Six participants were recruited from the undergraduate psychology course at the University of Kent. All participants had normal/corrected to normal vision and were native English speakers.

[2] A pilot study revealed that the initial 2 sec interval was sufficient to allow participants to read the three words under speeded conditions.

Materials

An original set of 34 triples (e.g., SALT-DEEP-FOAM) and their answers (e.g., SEA) were selected from Bowers, Regehr, Balthazard and Parker (1990) and a pilot study was conducted to reduce the set to the 20 most completed triples, which were subsequently used to construct two lists of matching difficulty. In addition, two further lists, each containing 30 words, were selected from the Kucera and Frances (1967) database. The first list was used alongside the triples in the study phase of the pre-dPOK memory task and the second list used alongside the test phase. All word lists were matched as closely as possible for familiarity and frequency. The post-dPOK recognition task offered three alternatives, consisting of 40 new words used alongside the 20 answers from the triples. The new words were selected to resemble (semantically or associatively) the correct answers in that they were clearly associated with at least one of the triple-associates (e.g., SALT-DEEP-FOAM; correct answer = SEA; possible answers = PEPPER or BATH). All stimuli were presented on an Apple Mac PC.

Procedure

Each participant was tested individually. During the study phase of the pre-dPOK memory task participants were presented with 60 words (words belonging to each triple associate were presented successively, however, the order of the triples was randomised by the program) with each word remaining on screen for 5 seconds. They were informed that after viewing this list a memory test would be administered and that they should attempt to memorise the words for future test. During the test phase the participants saw a further 60 words (30 ‘old’ and 30 'new') and were required to recognise as either ‘old’ (e.g., seen in the earlier list) or ‘new’ (e.g., not previously seen) each of the words presented. The order of presentation for the test words was also randomised and the computer scored on-line whether the response was correct or incorrect. On completion of this task participants read aloud from a passage of prose for three minutes before completing the RAT (and dPOK) task.

Prior to completing the dPOK judgements participants completed 4 practice trials to familiarise themselves with the procedure and required responses. They were then presented with 20 triple-associates trials (e.g., SALT-DEEP-FOAM), containing 10 *familiar* (e.g., used in the pre-dPOK memory task) and 10 *unfamiliar* (e.g., not seen before) triple-associates. Each set of triple-associates was presented centre screen and remained on screen until the fourth response was entered. Four tonal cues prompted participants to respond via the keyboard, according to how well they thought they knew the common associate and will be able to identify it in a subsequent recognition task. Participants used one of four designated keys to indicate dPOK levels of 25%; 50%; 75% and 100%. The deadline of 850 ms for the first response was imposed by providing on-line feedback with the message "too slow" appearing on screen for responses that did not meet the deadline; only responses that met the deadline were recorded. After the fourth dPOK response the triple-associate was removed from the screen and replaced by three words consisting of the correct answer and two lures. Participants were required to select a word that related to *all three* of the words presented in the earlier triple using one of three designated keys. No feedback regarding the accuracy of their response was given. The instructions emphasised the need to respond as quickly and as accurately as possible. After participants keyed in their response indicating their selection the next trial began. This process was then repeated for all remaining triple-associates with trial order randomised by the software.

Results and Discussion

Pre-dPOK recognition performance was high [Total correct-rate (hits + correct rejections) = 94.6%, which was significantly higher than chance, $t(1,5) = 655.2$, $p<.001$, Mse = 9.1)], indicating that participants were able to create and maintain a memory representation for the triples, which should influence their familiarity.

The dependency of the mean dPOK judgements, across all the participants, as a function of time is shown in Figure 1, together with the individual data for each of the six participants.

Figure 1 About Here

A 2 x 4 repeated measures analysis of variance (ANOVA) was carried out on participants mean dPOK for each list (*familiar* vs. *unfamiliar*) across the four times (*time1; time 2; time 3* and *time 4*). This revealed a main effect of Time, F (3,15) = 54.2, p< .001, Mse = 32.3, a marginal effect of List F (1,5) = 4.5, p= .09, Mse = 13.4 and a significant List x Time interaction F (3, 15) = 5.3, p< .01, Mse = 51. Planned comparisons were conducted between List type at Time 1 and Time 4 and within each List type between Time 1 and Time 4. These revealed that participants gave higher dPOK ratings to *familiar* stimuli relative to *unfamiliar* stimuli at *time 1* (59.4% and 44.1% respectively) F (1,5) = 8.8, p< .05, Mse = 79[3]. In contrast there was no difference in the dPOK ratings to *familiar* and *unfamiliar* stimuli at *time 4* (76.2% and 81.2% respectively) F (1,5) = 3.3, p= .13 Mse = 22. Within List A (*familiar*) participants dPOK ratings increased from *time 1* to *time 4* (59.4% and 76.2% respectively) F (1,5) = 13.4, p< .05, Mse = 63. Ratings for stimuli in List B (*unfamiliar*) also increased from *time 1* to *time 4* (44.2% and 81.2% respectively) F (1,5) = 68.6, p< .001, Mse = 60. Visual inspection of the individual participants' data (Fig. 1) indicates that the effect of Time is evident for each participant. However, the Time x List interaction is clearly observed in only four of the participants.

This pattern of results is consistent with our earlier predictions that the familiarity of the cue at time 1 would influence dPOK ratings and that over time, as more information is accessed from memory, dPOK levels for both familiar and unfamiliar stimuli would increase. Cue familiarity, which resulted from encoding items in a previous memory task led to higher initial metacognitive judgements. This is consistent with research showing that cue familiarity is a key mechanism used to drive early metacognitive processes (Metcalfe, 1993; Reder, 1987; Reder & Ritter, 1992; Reder & Schunn, 1996; Schunn, et al., 1997). Nevertheless, the individual data indicates that the familiarity manipulation did not have its effect on two of the participants (S4 and S6). This could be the result of the fact that the three words were

[3] It could be argued that the effect of List type at time 1 is an artefact of the stimuli used. To address this possible criticism we administered the same lists to a separate group of six subjects without requiring them to complete a pre-dPOK recognition task. Hence both lists were suggested to be of equivalent familiarity. Analysis revealed no difference in dPOK ratings at *time 1* between List A and List B (39.7% and 40.7% respectively) F, (1,5) = 0.5, p= .49, Mse = 6.1. However, there was an increase in the level of dPOK from *time 1* to *time 4* for both List A, F (1,5) =11.2, p< .05, Mse = 101 and List B, F (1,5) = 26.6, p< .01, Mse = 96.

primed in a successive, rather than simultaneous presentation (e.g., Reder, 1987). The rationale for doing so was to prevent any activation of associated targets during the priming session. For the four remaining participants, however, a sufficient level of familiarity was achieved even with this weaker priming manipulation. In our second experiment we tried to generate a more effective familiarity manipulation, which is expected to influence all participants.

The increase in dPOK levels over time for both types of stimuli suggests that as target information was accessed from memory, the information was fed back into the meta-memory process influencing subsequent responses. This later effect on dPOK levels is consistent with the view that information activated during the memory search is also utilised in driving metacognitive processes (Koriat, 1993, 1996; Koriat & Goldsmith, 1994; Maki, 1999; Schreiber, 1998).

Analysis of post-dPOK recognition scores revealed no significant difference in performance between *familiar* and *unfamiliar* stimuli (78.3% and 85% respectively) ($F < 1$). Finally, we examined the relationship between participants' dPOK ratings and their subsequent recognition performance using the Goodman-Kruskal gamma as an index of the association between items at *time 2* and *time 4* with accuracy of recognition and latency of responses to correctly recognised targets. As can be seen from Table 1, the mean dPOK judgements are positively associated with the accuracy in the recognition test and negatively with the response time. Moreover, the magnitudes of these associations increase with time. For accuracy, the magnitude increased from *time 2* to *time 4* (0.24 and 0.47, respectively) $t(5) = 3.53$, $p < .05$. For response latencies there was a trend showing an increase in the correlation magnitude from *time 2* to *time 4* (–0.25 and –0.46 respectively)[4].

Table 1 About Here

[4] Due to variability this trend failed to reach statistical significance [t (5) = 1.84, p= .12; two tail] but is marginally significant at one-tail [p=.060].

The fact that participants took less time and more accurately recognised the target when dPOK levels at time 4 were higher suggests that the information mediating such judgements was based on specific target knowledge. In addition to this, the increase shown in the association between dPOK judgements and recognition accuracy from time 2 to time 4 suggests that across time more target information was retrieved from memory. This pattern of results is consistent with the suggestion that the accuracy of later metacognitive judgements may be due predominantly to the high accuracy of memory itself (Koriat, 1993; 1995; 1996; Koriat & Goldsmith, 1996; Koriat & Levy-Sadot, 2001).

EXPERIMENT 2

Experiment 2 was conducted to examine the familiarity and retrieval mechanisms driving dPOK judgements when placed in opposition. A paradigm that places pre and post retrieval processes in opposition may be particularly useful in clarifying the nature of these processes and their interactions. A second aim of the experiment was to strengthen the familiarity manipulation so that its effect will be seen on all participants and not only on the group average. This was achieved by having participants complete dPOK ratings on riddles seen previously (i.e., *familiar*), new riddles (i.e., *unfamiliar*), and riddles seen previously that had undergone a cue manipulation (i.e., *transformed*) (see also Reder & Ritter, 1992). The manipulation involved a change to the letter cue of the riddle (e.g., from 'what is small can sting and ends in E' = Bee: to 'what is small can sting and ends in N' = Scorpion). We predict that stimuli within the transformed condition will appear more familiar, resulting in high initial dPOK ratings at time 1. Such a prediction is consistent with research showing that a partial matching process is sufficient to enable the participant to provide a response concerning information stored in memory (Kamas, Reder & Ayers, 1996; Reder & Kusbit, 1991). However, over time, as information is accessed from memory and the participants realise that in fact they do not know the answer, we expect to find a reduction in the level of dPOK ratings. As in experiment 1, dPOK levels for both familiar and unfamiliar stimuli are expected to increase from time 1 to time 4. Thus experiment 2 aims to show that the post-retrieval dPOK judgement can *both* increase and decrease over time, after the initial influence of cue familiarity at time 1. In addition, in order to facilitate participants ability to produce a fast dPOK response to the first cue (within an 850ms deadline), we replaced the four alternative responses used in experiment 1 with binary ones (high/low).

METHOD

Participants

A separate group of seven participants were recruited from the University of Kent to complete this experiment. All had normal or corrected to normal vision and were native English speakers.

Materials

A list of 42 riddles (e.g., what has wheels, can move, and ends in 'R') and their answers (e.g., car) were specifically constructed by the authors[i]. All stimuli were presented on an Apple Mac PC.

Procedure

Each participant was tested individually for both stages of the task. In stage one, participants were presented with a list of 26 riddles using a similar paradigm to experiment 1. They were required to respond via the keyboard with either a 'high' or a 'low' dPOK rating, immediately after each tone (as in Experiment 1), reflecting their ability to produce the answer at the end of the trial. In stage two participants were presented with 42 riddles, 16 unchanged from the previous stage (e.g., *familiar*), 16 new previously unseen riddles (e.g., *unfamiliar*), and 10 from the previous stage that underwent a cue manipulation (e.g., *transformed*). Similar to stage one, participants gave either a 'high' or a 'low' dPOK rating across the 4 times. Following the fourth rating, participants were presented with a prompt requiring them to recall the correct answer. Stimulus presentation for both stages was randomised and the computer recorded participants' dPOK ratings.

Results and Discussion

The dPOK judgements as a function of processing time both for the group average and the individual participants are shown in Figure 2.

Figure 2. About Here

Unlike in Experiment 1, the temporal dynamics of the dPOK judgements, for each of the participants, reflected exactly the same pattern as the group average (Fig. 2). A repeated measures analysis of variance (ANOVA) of participants mean dPOK ratings revealed a main effect of List, F (2,12) = 61.8, p< .001, Mse = .01, a main effect of Time, F (3,18) = 7.2, p< .01, Mse = .01 and a highly significant List x Time interaction, F (6,36) = 25, p< .001, Mse = .01.

The List x Time interaction was analysed by looking at the effects of Time within each List. For *familiar* stimuli there was no main effect of Time (F < 1). For *unfamiliar* stimuli there was a main effect of Time, F (3,18) = 37.1, p< .001, Mse = .01. Planned comparisons revealed that participants dPOK ratings increased from *time 1* to *time 2* (55% and 80% respectively) F (1,6) = 36, p< .001, Mse = .03, from *time 2* to *time 3* (80% and 85% respectively) F (1,6) = 33.3, p< .001, Mse = .02 and from *time 3* to *time 4* (85% and 90% respectively) F (1,6) = 48.5, p< .001, Mse = .09. For *transformed* stimuli there was a main effect of Time, F (3,18) = 12.9, p< .001, Mse = .01. Planned comparisons revealed that participants dPOK ratings decreased from *time 1* to *time 2* (85% and 60% respectively) F (1,6) = 47.6, p< .001, Mse = .03. In addition to this their dPOK ratings increased from *time 2* to *time 3* (60% and 65% respectively) F (1,6) = 6, p< .05, Mse = .03.

Analysis of the List type at *time 1* and *time 4* revealed a main effect of List, [F (2,12) = 35.3, p< .001, Mse = .02 and F (2,12) = 12.69, p< .001, Mse = .02 respectively]. Comparisons, at *time 1* using a corrected alpha level of p< .016 (Bonferroni correction of .05/3 = .016) revealed that dPOK ratings were higher for *familiar* relative to *unfamiliar* (90% and 55% respectively) t (6) = 8.9, p< .001, and *transformed* relative to *unfamiliar* lists (85% and 55% respectively) t (6) = 6.2, p< .001. There was no difference between *familiar* and *transformed* lists (t < 1). In contrast, similar comparisons, at *time 4* using a corrected alpha level of p< .016 (Bonferroni correction of .05/3 = .016) revealed that dPOK ratings were higher for *familiar* relative to *transformed* (91.7% and 72.8% respectively) t (6) = 4.76, p< .016, and marginally higher for *unfamiliar* relative to *transformed* lists (88.2% and 72.8% respectively) t (6) = 2.96, p= .025. There was no difference between *familiar* and *unfamiliar* lists (t <

2). Analysis of post-dPOK recall scores revealed a main effect of List, $F(2,12) = 20.2$, $p < .001$, $Mse = 225$. Comparisons using a corrected alpha level of $p < .016$ (Bonferroni correction of $.05/3 = .016$) revealed that participants were more accurate at recalling the correct answer for riddles from *familiar*, relative to *transformed* lists (89.2% and 40% respectively; $t(6) = 5.9$, $p < .001$), and *unfamiliar* relative to *transformed* (76.1% and 40% respectively; $t(6) = 3.6$, $p < .016$). These results show that the differences in dPOK ratings made at time 4 are consistent with the actual recall performance, indicating that these later metacognitive judgements are a good indicator of memory performance.

The fact that participants gave higher dPOK ratings to familiar stimuli at time 1 replicates the pattern found for experiment 1 and is consistent with research showing effects of cue familiarity on early metacognitive judgements (Metcalfe, 1993; Reder, 1987; Reder & Ritter, 1992; Reder & Schunn, 1996; Schunn, et al., 1997). The pattern of dPOK judgements over time for the unfamiliar stimuli replicated that reported from experiment 1, showing an increase from time 1 to time 4. However, unlike in experiment 1, overall dPOK levels for familiar stimuli did not increase over time. One explanation for this is that the participants' dPOK ratings for familiar stimuli were already at ceiling at time 1. Another possibility is that the familiarization stage, led to a priming of the targets and not only of the questions, which may have influenced early dPOK levels at time 1. Consequently, if as soon as the riddle was presented the participants already had the solution in mind then one would expect to see little or no change in the individual ratings from time 1 to time 4. Although it is not possible to fully refute this explanation (it could hold for one participant, S1), we believe that this is unlikely to hold for most of the participants (six out of the seven participants were not at ceiling in this condition; four participants showed an increase and one showed a decrease in dPOK with time). Moreover as the first dPOK is required very soon (2.85 sec) after the presentation of the riddles, it is more likely that participants generated the initial response on the basis of the familiarity with the question cues, and then updated it on the basis of the retrieval process. In most cases (four subjects) this led to an increase in the dPOK levels. The decreasing pattern (e.g., S4) can be explained as resulting from an initially overconfident dPOK judgement (due to the familiarity of the question) in conjunction with a much slower retrieval process for the target (at time2 the participant thus realised that the target is not yet available).

The key result of this study stems from the pattern of dPOK ratings for the transformed stimuli. Initially participants gave high dPOK ratings to the transformed stimuli based on the familiarity of the cues (see also Metcalfe, 1993; Reder, 1987; Reder & Ritter, 1992; Reder & Schunn, 1996; Schunn, et al., 1997). This is consistent with research showing that a partial matching process is sufficient to elicit a metamemory response (Kamas, Reder & Ayers, 1996; Reder & Kusbit, 1991). Nevertheless, as participants searched memory for the target and realised that in fact they did not know the answer, they altered their dPOK level accordingly, as indicated by the clear reduction in ratings from time 1 to time 2. Moreover, the increase in the level of dPOK judgements from time 2 to time 3 suggests that once the participants' realise the cue does not relate to a familiar riddle, the continual search of memory led to activation of some target information resulting in an increase in dPOK.

The dPOK levels for transformed stimuli showed only a limited increase from time 2 to time 4 (relative to familiar and unfamiliar stimuli, see Fig 2). One possible reason for this is that the original cue encoded during stage one may have interfered with the search for the correct solution resulting in a less robust rating across time. Such a possibility is consistent with research showing that learning one stimulus-response association may impair the ability to learn a new response to the same original stimulus (e.g., Shapiro & Olton, 1994). Alternatively, this pattern of dPOK responses may be the result of inter-trial variability. For instance, the time at which participants realised that the transformed cue was different may have occurred at different times within each trial. When averaged over trials, this inter-trial variability results in a more gradual increase in dPOK levels over time.

General Discussion

The reported results indicate that the dPOK judgement reflects a dynamic process that can be influenced *both* by the familiarity of the question-terms and by the information activated during the search for the target. Accordingly, such metacognitive judgements are driven initially by the familiarity of the cue (Metcalfe et al, 1993;

Reder, 1987; Reder & Schunn, 1996; Schunn, et al., 1997), however subsequent information accessed during the search for the target provides an additional contribution (Koriat, 1993; Koriat & Levy-Sadot, 2001; Schreiber, 1998; Maki, 1999). As first suggested by Reder (1987; Reder & Ritter, 1992; Reder & Shun, 1996), fast evaluations (made in less than 850 ms) based on the familiarity of the question-terms can, in certain situations, be instrumental for the selection of retrieval strategies that are well suited for the task. For example, in an exam situation, when multiple questions need to be answered under a shared time deadline, fast metacognitive evaluations based on familiarity with the question-terms and topic, can provide a person with a useful guide in deciding which questions to work on and which ones to abort. When a person decides (or is required) to work on a specific problem, however, the metacognitive process is not frozen at the initial evaluation, but is updated dynamically with the results obtained from the retrieval process. Such late meta-judgements can also be instrumental in controlling cognitive processing, as they may lead to the decision to give up, or alternatively, to invest more effort in solving the problem. Finally, after responding, a metacognitive evaluation of *confidence* is generated, guiding future attitudes and behaviour. A complete scheme of the unfolding metacognitive process and its role in cognitive control, consistent with this verbal description, has been proposed by Nelson (1996; Barnes, Nelson et al., 1999).

The decrease in metacognitive judgements made over time for the transformed stimuli provides a challenge to the accessibility account, which suggests that such judgements should increase as the search process unfolds activating and accumulating both correct and incorrect information (Koriat, 1993). Notice, however, that we refer here to the effortful slow and not the automatic fast component of the accessibility theory (Koriat & Levy-Sadot, 2001). Moreover, the interaction between familiarity and retrieval time shows a different pattern from that reported by Koriat and Levy-Sadot (2001), who by manipulating orthogonally the familiarity and accessibility components found a super-additive effect on metacognitive judgements. In contrast, our results exhibit a sub-additive effect (i.e., less increase with retrieval time in the familiar condition). This may indicate that the way in which the slow retrieval component contributes to the metacognitive process is different from the way in which the fast automatic component does. It is indeed expected that on a time scale that allows conscious processes to operate and monitor the quality of performance, meta-judgements such as dPOK become as time progresses predominately determined by the quality of the retrieval and closeness to the solution, and less by the familiarity with the question terms. One explanation of this data may be provided by a competitive theory of metacognition (Maki, 1999; Schreiber, 1998; Schreiber & D. Nelson, 1998). According to this view, conflicting information activated by the familiar primed riddles and the unfamiliar transformed cues competes to reduce the initially high dPOK judgement. Alternatively, a combined accessibility-competition theory could explain these findings by assuming that such judgements depend on the *total* activation accumulated in the retrieval process, which is however, subject to competitive interactions. Accordingly, early-on activation is dominated by the cue items (determining fast judgements), while later on in the retrieval process, the activation reaches knowledge schemas in long term memory that generate competitive constraints (Rumelhart, Smolensky, McClelland & Hinton, 1986).

Although the dPOK judgement used here is formally different from each of the more traditional metacognitive judgements (e.g., POK, FOW and FOK), we suggest all these judgements tap similar metacognitive processes that unfold at different stages. However, in order to demonstrate the generality of our results, a number of further investigations are needed. First, online metacognitive judgments at various temporal delays from stimulus presentation should be obtained in the more traditional FOK methodology, where responses are solicited following an unsuccessful recall (utilising both within and between participants designs). Second, it is important to perform similar experiments in other domains of memory, such as general-knowledge. Third, psychophysiological measures (e.g., fMRI and event related brain potentials) may help to provide a more direct measure of the neurophysiological correlates of metacognitive judgements (Botvinick et al., 2001). Finally, more formal computational models addressing various metacognitive judgements need to be developed and tested against the data.

In conclusion, the present experiments illustrate the temporal aspects of the different mechanisms driving the metacognitive process. Using methods of temporal tracking for metacognitive judgements, first introduced by Metcalfe (1986b; Metcalfe &

Weibe, 1987), we demonstrated that such metacognitive judgements are based on *both* cue and target information at different times in the process. One implication of this is that the metacognitive process is dynamic, changing over time according to the quality and quantity of information accessed from memory. Consistent with other recent studies (Barnes et al., 1999; Koriat & Levy-Sadot, 2001; Metcalfe et al., 1993; Nelson, 1996), we view the metacognitive process as operating in parallel with the memory search and retrieval processes. Nevertheless, the initial output can be issued very quickly, determined predominantly by the familiarity of the question cue (Reder, 1987; Reder & Ritter, 1992). This suggests that such searches are not made 'blindly' but are controlled by information gained both from the familiarity of the cue as well as the information activated from the search itself.

Acknowledgements

We would like to thank Asher Koriat, Janet Metcalfe and Lynne Reder for valuable advice and helpful comments during the review.

Figure Captions

Figure 1.

Figure 1A represents participants mean dPOK ratings (with standard error bars) for familiar and unfamiliar Lists, across each of the four Times. Figures 1B to 1G represent each individual participants mean dPOK ratings across time.

Figure 2.

Figure 2A represents participants mean dPOK ratings (with standard error bars) for familiar, unfamiliar and transformed Lists, across each of the four times. Figures 2B to 2H represent individual participants dPOK responses across time.

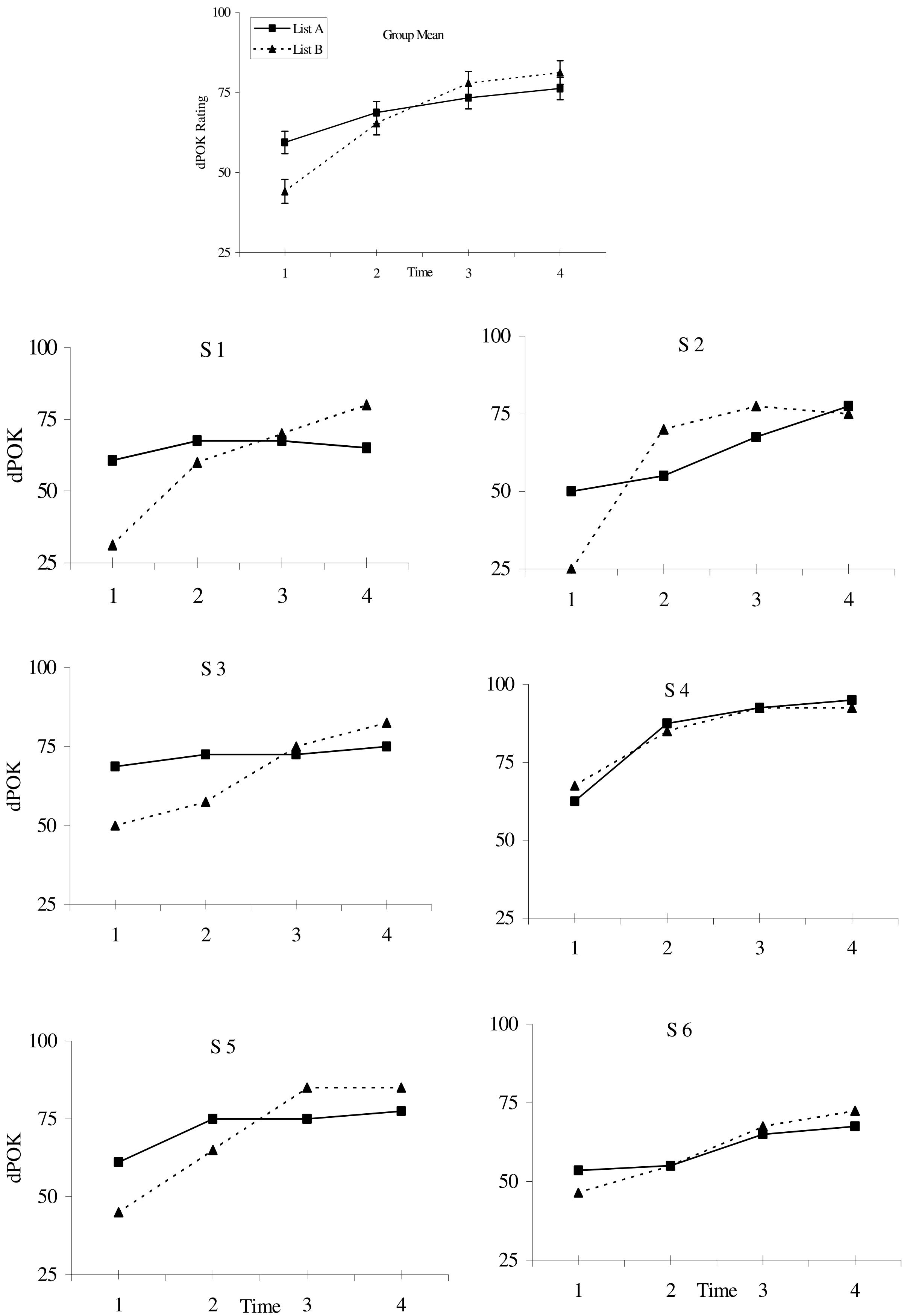
Table 1.

Showing mean and standard errors of the correlations between dPOK ratings and recognition accuracy and latency at time 2 and time 4.

	dPOK at time 2	dPOK at time 4
Reaction Time	-.25* (0.09)	-.46* (0.14)
Recognition accuracy	.24 (0.14)	0.47* (0.14)

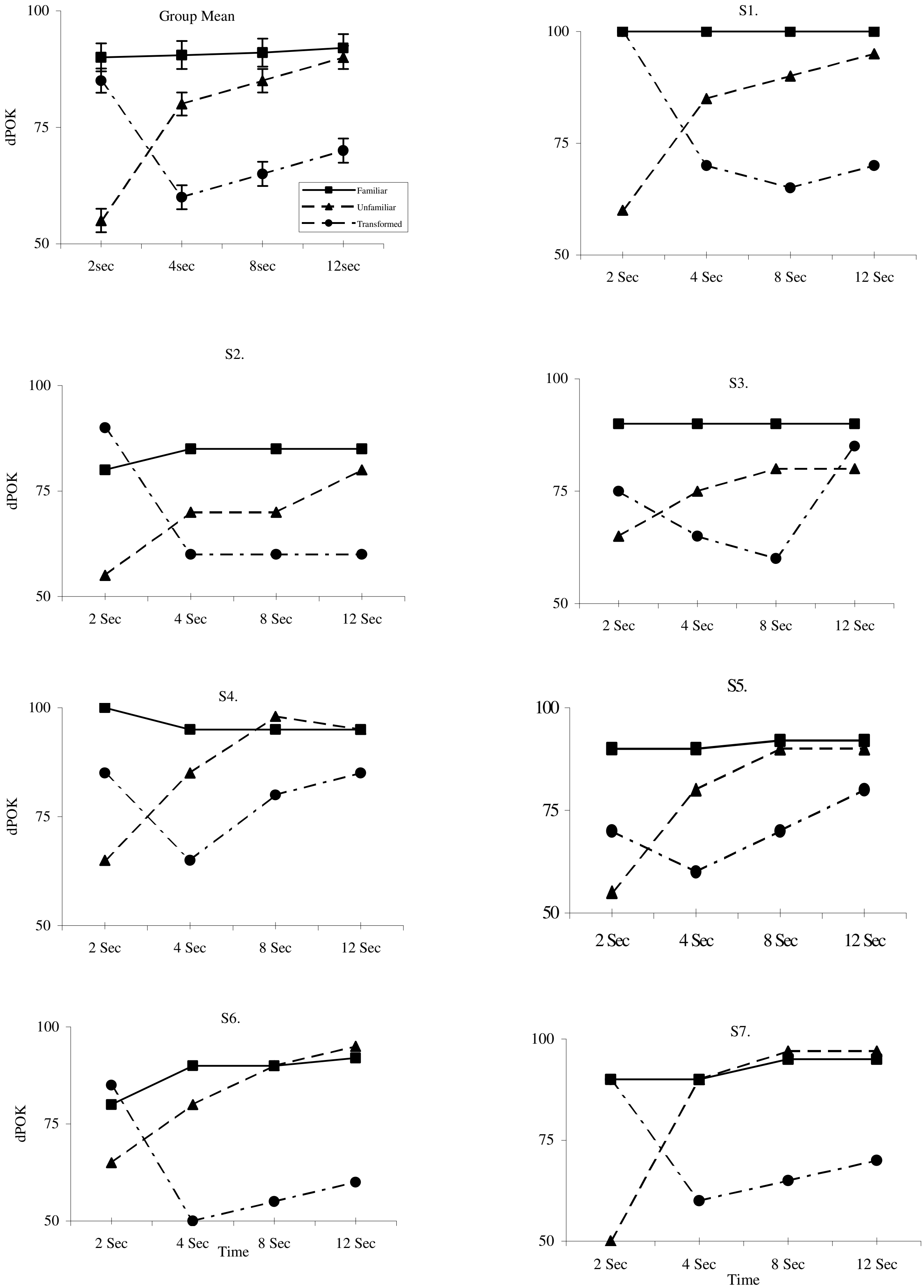
* Significantly different from zero ($p < .05$)

FIG 1.



Group Mean
List A
List B
dPOK Rating
Time
S 1
S 2
S 3
S 4
S 5
S 6
dPOK
Time

FIG 2.


Group Mean
S1.
S2.
S3.
S4.
S5.
S6.
S7.
dPOK
Time
100
75
50
2sec
4sec
8sec
12sec
2 Sec
4 Sec
8 Sec
12 Sec
Familiar
Unfamiliar
Transformed

References

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

Botvinick, M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review,* 108, 624-652.

Bowers, K. S., Regehr, G., Balthazard, C., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology, 22,* 72-110.

Kamas, E. N., Reder, L. M., & Ayres, M. S. (1996). Partial Matching in the Moses illusion: Response bias not sensitivity. *Memory and Cognition, 24,* 687-699.

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

Koriat, A. (1995). Dissociation knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology, General,* 124, 311-333.

Koriat, A. (1996). Memory's knowledge of its own knowledge: The accessibility account of the feeling of knowing: In Metcalfe, J., & Shimamura, A. P. (Eds.). Metacognition: knowing about knowing. A Bradford Book. MIT.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123,* 297-315.

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

Kucera, H., & Francis, W.N. (1967). Computational analysis of present day American English. Brown University Press.

Maki, R. H. (1999). The roles of competition, target accessibility, and cue familiarity in metamemory for word pairs. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25,* 1011-1023.

Mednik, S. (1962). The associative basis of the creative process. *Psychological-Review, 69,* 220-232.

Metcalfe, J. (1986a). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 12, 288-294.

Metcalfe, J. (1986b). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 12, 623-634.

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: Learning, Memory and Cognition,* 19, 851-861.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory and Cognition, 15*, 238-246.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist,* 51, 102-116.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation,* 26, 125-141.

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

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

Reder, L. M., & Kusbit, G. W. (1991). Locus of the Moses illusion: Imperfect encoding, retrieval, or match? *Journal of Memory and Language,* 30, 385-406.

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

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

Rumelhart, D. E., Smolensky, P, McClelland, J. L. & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland and D. E Rumelhart and the PDP research group, *Parallel distributed processing* (vol. 2). A Bradford book, Cambridge MA: MIT press.

Schreiber, T. A. (1998). Effects of target set size on feelings of knowing and cued recall: Implications for the cue effectiveness and partial retrieval hypothesis. *Memory & Cognition, 26,* 553-571.

Schreiber, T. A., & Nelson, D. L. (1998). The relation between feelings of knowing and the number of neighbouring concepts linked to the test cue. *Memory & Cognition, 26,* 869-883.

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 & Cognition,* 23, 1-27.

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

Shapiro, M. L., & Olton, D. S. (1994). Hippocampal function and interference. In D. L. Schacter, and E. Tulving (Eds). *Memory systems 1994*. (pp. 87-117). Cambridge. MIT Press.

---

[i] A copy of the riddles used in experiment 2 is available at the following web address: http://www.med.ic.ac.uk/divisions/49/Davidpub.htm