Changes in Memory Awareness During Learning: The Acquisition of Knowledge by Psychology Undergraduates

Martin A. Conway
University of Bristol

John M. Gardiner
City University, London

Timothy J. Perfect and Stephen J. Anderson
University of Bristol

Gillian M. Cohen
The Open University

First-year psychology students took multiple-choice examinations following each of 4 lecture courses and 3 laboratory research methods courses. One lecture course was later retested. Students indicated state of memory awareness accompanying each answer: recollective experience (remember), "just know" (know), feeling of familiarity (familiarity), or guess. On the lecture courses, higher performing students differed from other students because they had more remember responses. On research methods, higher performing students differed because they knew more, and in the delayed retest, higher performing students differed because they now knew rather than remembered more. These findings demonstrate a shift from remembering to knowing, dependent upon level attained, type of course, and retention interval, and suggest an underlying shift in knowledge representation from episodic to semantic memory. The authors discuss theoretical and educational implications of the findings.

In the process of acquiring a new knowledge domain, knowledge retained by the learner is open to a number of different conscious states. Consider an ungraduate student taking an introductory course that includes an account of classical conditioning (Pavlov, 1927) in a lecture on learning theory. How does the student respond to questions about classical conditioning shortly after the lecture? Presumably the student remembers some of the comments of the lecturer. Perhaps images from the overhead projections also come to mind and, with luck, at least some of the logic of classical conditioning is retained. In the weeks and months following the initial lecture and before being examined on it, the student reads about conditioning in text books, discusses it with others, and completes relevant course work assignments. Again, some knowledge may be remembered from these various sources, but now, in addition to remembering, the student may know what the various unconditioned and conditioned responses and stimuli are and how these combine in the processes of acquisition and extinction. That is, this knowledge may come to mind without any recollection of any particular learning episode. Thus, remembering particular learning episodes may no longer be the state of consciousness that usually accompanies knowledge access, and instead, the knowledge may more often come to mind only with the consciousness that it is known.

In this article, we describe what we believe to be the first empirical study of how these states of consciousness change during a protracted period of learning: appropriately enough, with large groups of 1st-year undergraduate students taking various introductory psychology courses. The study addresses four main questions: What role, if any, does episodic remembering play in this type of learning? How do these remembering and knowing states of awareness relate to differences in the overall level of performance that the students achieve? How do these states of awareness change during acquisition? And how are they related to differences in the nature of the courses? Theoretically, the findings have implications for understanding how states of awareness during memory access are affected by changes in the underlying representation of knowledge. Practically, the findings suggest new ways in which to assess the type of knowledge a student has acquired and to detect changes in knowledge structures as learning proceeds. The distinction between remembering and knowing as subjective states of consciousness that can be measured was introduced by Tulving (1985a), and before turning to further details of our study, we briefly review what is currently known about these
states of awareness and how they might relate to underlying knowledge structures.

Remembering and Knowing

Tulving (1985a) introduced a distinction between "remember" and "know" responses and provided the first demonstration that people can make these responses in a memory test, item by item, to report their conscious states. Remembering was defined as having recollective experience, a state in which images, feelings, and other context-specific details relating to a past event come to mind, such that the experience is that of mentally reliving a particular episode. Recollective experience is accompanied by a "sense of pastness" and of the role of the self in the past (see Conway & Dewhurst, 1995), and Tulving (1985a) referred to this kind of awareness as autonoetic consciousness or self-knowing. In contrast, awareness that something has been encountered or experienced before need not entail any of the sequelae of recollective experience and may, instead, only be characterized by a sense of knowing or familiarity. Tulving (1985a) referred to this kind of awareness as noetic consciousness or knowing. In Tulving's theory, remember responses, which are reports of recollective experiences indicating autonoetic consciousness, are the hallmark of an episodic memory system. In contrast, know responses, which are reports of knowledge of facts or events in the absence of recollective experiences indicating noetic consciousness, are the hallmark of a semantic memory system (Tulving, 1983, 1993, in press). These measures have been used to elucidate the contributions of the two kinds of consciousness to many different memory phenomena (e.g., Conway, Collins, Gathercole, & Anderson, 1996; Dalla Barba, 1993; Gardiner & Java, 1991; Huron et al., 1995; Parkin & Walters, 1992). The evidence currently documents an impressive list of dissociations and associations between remember and know responses, especially in recognition memory tests. These findings are systematic, replicable, and interpretable theoretically, and so confirm the validity of the judgments (for reviews, see Gardiner & Java, 1993; Rajaram & Roediger, in press). Hence the state of awareness measured by know responses has sometimes been interpreted with respect to differences in processing rather than differences in the way knowledge can be represented. For example, it has been suggested that know responses might reflect data-driven processing, enhanced perceptual fluency, or automatic unconscious influences of memory (e.g., Gardiner, 1988; Jacoby, Yonelinas, & Jennings, in press; Rajaram, 1993). But despite the growing body of research showing different effects on recollective experience and on familiarity as measured by remember and know responses, there is currently no generally accepted theoretical account that encompasses the full range of the findings (though see Dewhurst & Conway, 1994; Gardiner & Java, 1993; Knowlton & Squire, 1995; Rajaram, 1993, 1996; Rajaram & Roediger, in press; and Richardson-Klavehn, Gardiner, & Java, 1996, for a range of current suggestions), and the interpretation of know responses in particular continues to be controversial (see, e.g., Gardiner, Java, & Richardson-Klavehn, in press; Jacoby et al., in press; Richardson-Klavehn et al., 1996).

Fractionating Knowing and Familiarity: The Just Know State

Students who have acquired sufficient knowledge of classical conditioning or any other psychological topic, such that they can answer examination questions correctly, produce coherent and knowledgeable course work and use their knowledge in laboratory classes, for example, would not necessarily feel that their knowledge was familiar in the way that know responses have previously been defined. For instance, when asked which was the conditioned stimulus in Pavlov's experiments, the salivation or the light, the student may simply know the answer. In pilot studies preceding the present study, we found that students presented with multiple-choice questions often indicated that neither recollective experience nor familiarity accurately captured their conscious state when selecting an answer. Instead, the students, who were often extremely confident (and accurate) in their choice of answer, commented that they just knew that answer.

The problem is that familiarity has at least two quite distinct meanings. Familiarity can be defined as the feeling that something has been encountered or experienced recently, although nothing about this recent occurrence can be remembered. A student might be aware of having recently come across Pavlov's name without being able to recollect when or where. It is this sense of familiarity that know responses have measured in previous studies. When students feel a word from a study list has been encountered recently
but do not recollect anything about that encounter, they infer that the word was on the study list. They believe the word occurred in the study list as a fact, because it feels familiar. However, an item may also be familiar because it has been frequently, rather than recently and uniquely, encountered. Such highly familiar items might come to mind without recollecting any particular encounter or any feeling of a recent encounter that cannot be placed: These items are known and give rise to just know responses. It is in this sense that a student might be aware, might just know, that Pavlov developed classical conditioning techniques, not Freud.

Thus, in this study, we further distinguished between these two kinds of familiarity. We did so by allowing students to use know responses for the semantic sense in which knowledge was highly familiar, and its use was unaccompanied either by any recollection of when or where it was acquired, or last encountered, or any feelings of some recent encounter that could not be recollected. When students were aware of some recent encounter they could not place, they used a familiar response. Note that not only does this definition of a know response differ sharply from that in previous studies, which were concerned with familiarity in the sense of an awareness of some recent encounter, here we were not particularly concerned with this kind of awareness. Instead, we were concerned with knowledge that students simply knew compared with knowledge students remembered because they could recollect the learning episode. It is after all the aim of higher education that students come to know, in the semantic sense, a body of knowledge. If a graduate could only recollect knowledge retained in memories of learning episodes or if knowledge just seemed familiar, in the sense that it had been recently studied, then it could not be concluded that the student had acquired conceptual knowledge of the area.

The Transition From Episodic to Conceptual Knowledge

For present purposes, the essential aspect of Tulving's (1985a, 1985b) proposals that we wish to retain is the suggestion that the nature of underlying knowledge can influence and even determine the type of awareness that is experienced during knowledge access (cf. Dewardt & Conway, 1994). We suggest that when a new knowledge domain is to be acquired, memory is represented initially in a way that supports or even compels recollection of the learning episode. As learning proceeds, the underlying representations may change such that they no longer primarily lead to recollective experiences and instead become so highly familiar that they are simply known. Thus, we postulate a shift in the basis of learning from that is episodic and literal to learning that is semantic and conceptual.

Surprisingly little is known about how knowledge acquired in a specific episode (e.g., in a lecture) can, over the course of extended learning, become conceptual knowledge free of reference to the details of the initial and subsequent learning episodes (but see Ross, Perkins, & Tenpenny, 1990). One view is that knowledge becomes schematized (Bartlett, 1932; Neisser, 1976; Schank, 1982; Schank & Abelson, 1977). For instance, Linton (1986) describes how her own autobiographical memory for attending (academic) meetings changed as she attended more meetings. At the time she tried to recall the many meetings she had attended, she was only able to remember one or two recent meetings and one or two of the very first meetings at which she was present. Linton argued that her memory for meetings had become schematized into an abstract representation of what usually took place at the types of meetings she attended, the people typically present, locations, and topics. This account is much in the tradition of Schank & Abelson (1977) who proposed that abstract scripts are formed for often repeated stereotypical action sequences (e.g., going to a lecture; Bower, Black, & Turner, 1979). This view of the schematization of event knowledge may suggest how some aspects of semantic knowledge arise in the acquisition of complex and large-scale knowledge domains such as psychology, physics, mathematics, geography, sociology, and so forth. For example, there could be schema for certain subsets of knowledge. In psychology, there might be schematic knowledge of experimental design (Conway, Cohen, & Stanhope, 1991), learning theory, or other well-demarcated subdomains. However, knowledge domains must also represent rules, algorithms, concepts, and specific aspects of knowledge, and it is difficult to see how the schematization of event knowledge might account for acquisition of these aspects of semantic knowledge. Yet, rules and algorithms are acquired. For example, research by Ross (1984) has shown how people learning to use various word-processing programs move from relying on very specific episodic memories (i.e., "I deleted this word using this chain of commands in a previous learning episode") to relying on rules ("This is the command sequence to delete a word."). Although the processes underlying the transition from episodic to semantic knowledge are unknown, it is nevertheless useful to use the term schematization to denote this change, bearing in mind that we mean more by this than the schematization of event knowledge: Our use of the term encompasses the acquisition of semantic representations of rules, concepts, algorithms, and so forth, as well some general abstract representation of frequently encountered or stereotypical aspects of a knowledge domain. Moreover, if the transition is from knowledge embedded in episodic or autobiographical memories of specific learning experiences to knowledge represented semantically, then as argued earlier, there should be a corresponding shift in memory awareness from remembering to knowing. Thus, memory awareness should vary systematically with the degree of schematization of knowledge, as Tulving (1985a) implied.

Research into the very long-term retention of knowledge acquired through formal education (i.e., at high school and university) has tended to oscillate around the issue of schematization of knowledge (see Conway, Cohen, & Stanhope, 1993, and Semb & Ellis, 1994, for reviews). For instance, Bahrick (1984) suggested that very long-term knowledge was preserved in a state he called permastore in which residual specific details and more abstract conceptual knowledge were retained in a form resistant to forgetting.
Alternatively, Neisser (1984) proposed that what was retained were schema that supported the reconstruction of knowledge. Thus, when presented with a cue such as a multiple-choice question from a previously acquired knowledge domain, the schema and cue could be combined to reconstruct a response and produce above-chance memory performance after a retention interval of many years. According to the schema view, no specific knowledge is retained. In our own research, however, into the very long-term retention of cognitive psychology (Conway et al., 1991, 1992, 1993), we found that specific details, such as the names of researchers, were retained for long periods of time but that an initial period of forgetting was more marked in the first 3 years after acquisition for specific compared to conceptual knowledge. These latter findings suggest that very long-term knowledge structures, at least in this domain, are composed of a mixture of detailed and schema knowledge bound together in representations resistant to forgetting. Indeed, conceptual knowledge more generally is probably closely associated in long-term memory with some specific knowledge and even with autobiographical memories (Conway, 1990; Schank, 1982), as the embedded view of episodic and semantic memory suggests (Tulving, 1985b).

The previous considerations suggest various predictions for the present study, and here we consider conjectures derived from three different views: the pure schema view, the schema-plus-episodic view, and the pure episodic view. According to the pure schema view, when knowledge has been completely schematized such that all that is represented in long-term memory are abstract schema, which in conjunction with externally presented cues can be used to generate appropriate responses (i.e., select the correct options in a multiple-choice test), then recollective experience and feelings of familiarity do not occur during knowledge access. Instead, responding would be characterized by states of conscious awareness associated with access to these more specific schema. Instead, and rather as in the case of Linton’s (1986) memory for meetings, access of the schema would give rise to the “just know” state and to occasional incidents of recollection. Occurrence of these latter states might be dependent upon the nature of the actual knowledge made available by the schema in particular processing tasks (Dewhurst & Conway, 1994). Counter to both the pure schema view and the schema-plus-episodic view, the pure episodic view proposes that knowledge that has not been schematized or that is only partly schematized (knowledge may pass through this state during the course of schematization as this takes place in long-term learning) is represented in memories of specific learning episodes. When this knowledge is accessed, the experience of recollection is frequent as are feelings of familiarity, whereas at the same time the “just know” state is only rarely experienced.

During the course of acquisition of a new knowledge domain a learner may pass through all three forms of long-term knowledge. Initial exposures to the to-be-acquired knowledge may result in the formation of episodic memories of the specific learning episodes and therefore give rise to a dominance of recollective experience during knowledge access. As learning proceeds, schema develop, but these still maintain indices to some episodic information and so knowledge access may be distinguished by a greater frequency of just know states along with some recollective experience and feelings of familiarity. When the learner has reached a level of expertise, then knowledge representations may be largely abstract and conceptual and maintain few if any direct links with memories of specific learning episodes. In this case knowledge access would be very largely dominated by the “just know” state. In the case of the 1st-year students sampled in this study it seems implausible that they would reach this expert state of acquisition of knowledge of psychology. Instead we predict that initially they will have the pattern of memory awareness states suggested by the pure episodic view. Possibly the better students will have a pattern more indicative of the schema-plus-episodic view. Over the course of the year we predict a shift in characteristic patterns of memory awareness during knowledge access from one typical of the pure episodic view to one more representative of the schema-plus-episodic view showing a dominance of just know states along with some recollective experience and feelings of familiarity. It also seems possible that certain courses may facilitate the shift more than others. Following Neisser’s (1984) suggestion that some knowledge domains may be more easily schematized than others, it may be that those courses that teach prescribed and integrated areas of knowledge facilitate the rapid creation of schematized knowledge. In the case of the present study, candidates for accelerated development of schema are the courses in research methods that cover a more proscribed range of knowledge than the lecture courses, which tend as with most introductory courses to be wide ranging. Thus, we expect to observe some differences in the patterns of memory awareness in the tests of lecture courses compared to courses in research methods. These differences may take the form of more just know responses to materials from the courses in research methods compared to materials from the lecture courses.

In summary, it is not known how knowledge originally presented in single learning episodes makes the transition from an episodic type of representation to a semantic or conceptual representation. It is clear, however, that this is what must occur in the process of the acquisition of complex knowledge domains acquired during the course of education. The purpose of the present research is to examine how states of memory awareness that accompany knowledge access change over the course of learning. We hypothesize that shortly after the first few learning experiences knowledge is retained in a predominately episodic form and this gives rise to a preponderance of recollective experiences (remember responses) when this knowledge is accessed. As learning progresses through the academic year there is a transition in the underlying memory representations from
the episodic type of representation to a more conceptual and schematized type of representation, which nonetheless contains specific details and possibly some memories of perhaps critical learning experiences (cf. Pillemer, Picariello, Law, & Reichman, 1996). When this occurs, knowledge access is associated with knowing (in the semantic sense or a just know state) and feelings of familiarity (a sense of previous occurrence), whereas recollective experience becomes less frequent. An implication of this view is that students who attain high grades in initial tests do so because they remember more. In later tests, however, such as end-of-year examinations, the better students succeed because they know more. This shift from remembering to knowing may be more rapid on certain types of courses that present materials that can be readily represented in long-term memory in schematic form.

### General Method

#### The Lecture Courses

The first year course consisted of four courses, two research methods courses, and a course of laboratory classes. The latter course did not feature in the study. The four lecture courses were Introduction to Psychology, Physiological Psychology, Cognitive Psychology, and Social and Developmental Psychology. Each course used several text books, chapters, and some journal articles. The courses ran consecutively in the order stated in 6-week blocks with three lectures per week. Thus, each lecture course comprised 18 one-hr lectures with associated tutorials and a course essay. For the course Social and Developmental Psychology, there were 8 lectures in social psychology and 9 lectures in developmental psychology, and students were free to choose either area for their course essay. In the 18th lecture slot, for all four courses, a class test was administered under examination conditions.

The courses in research methods had been specifically designed for single- and joint-honor psychology students. These courses covered issues in experimental design and statistics and were yoked to the laboratory practicals. In Research Methods Part I, the students learned basic statistical theory and covered simple non-parametric tests concluding with the t test. At the same time, they learned elementary design and hypothesis testing. Research Methods Part II, taken by single-honors students only, encompassed advanced issues in experimental design and the ethics of research and introduced the analysis of variance (ANOVA), culminating in a simple $2 \times 2$ mixed design. The research methods courses were taught in two blocks each of 12 three-hour laboratory classes at the rate of one per week. The class test for Research Methods Part I occurred in Week 2 of Term 2 (the same week as the class test for Physiological Psychology), and for Research Methods Part II, the test was in Week 3 of Term 3 (24 weeks into the 1st-year program, only a few days before the first of the summer examinations, and in the same week as the class test for Social and Developmental Psychology). In addition to data from these tests, which were taken in the academic year 1994–1995, we also include in these analyses data from the Research Methods Part I course, which ran in the academic year 1995–1996. This course was very similar to the original 1994–1995 course but included some new materials replacing material originally taught, and a slightly changed lineup of lectures delivered the teaching sessions. We refer to Part I and Part II of the courses from the year 1994–1995 as Research Methods(1) and Research Methods(2), respectively, and to the Part I course from 1995–1996 as Research Methods(3).

### Materials

Each class test contained, notionally, 68 multiple-choice questions (MCQs), 4 for each of the 17 content lectures. In practice, some lectures, which included demonstrations and the like, did not give rise to appropriate MCQs, and consequently, the actual number of questions varied with individual courses. Number of questions per course are listed in Table 1, presented later. Each multiple-choice question contained three options, and students were forced to select an option. Failure to select an option resulted in a “fail” on that question, and students were encouraged to guess rather than not answer. The questions were based on information presented directly, either verbally or visually, in a lecture.

All the courses were team taught, and individual lecturers generated four MCQs for each of the lectures they had delivered. The lecturers were also asked to vary the difficulty level so that for any given lecture, two of the MCQs were relatively easy and two more difficult. The construction of tests was coordinated by Martin A. Conway, who also screened the difficulty levels of the MCQs. In practice, it proved too difficult to ensure constant difficulty levels across all MCQs, either within a single course or across courses, and consequently this variable does not feature in the results reported later. Each MCQ contained only one correct answer, and the two incorrect alternatives were taken from information presented in the same lecture in which the target had been presented. Incorrect alternatives were selected on the basis of high plausibility and strong conceptual relatedness to the target. Examples of the MCQs for each of the four lecture courses and for the research methods courses are listed in the Appendix.

Systematically varying the difficulty level of MCQs for the class tests in the research methods courses proved impossible. Similarly, mapping all questions onto information presented in the classes was difficult. Instead, a different strategy was adopted. For Research Methods(1), half the questions were based on information explicitly and directly presented in the classes, either visually or verbally. The remaining questions required students to make an inference on the basis of what they had been taught. In the MCQs for Research Methods(2), all the questions were of this latter type. In the class test for Research Methods(3) from 1995–1996, no attempt was made to balance questions in this way, although because of changes in the course, most of the questions referred directly to items specifically presented by the lecturers in the teaching sessions. Note also that many of the questions for the test in 1995–1996 were new questions not previously presented in 1994–1995.

### Procedure

For each course the MCQs were presented in a printed booklet in which the cover page listed instructions. For each item the question appeared first with the three choices listed below, followed by the four memory-awareness categories: remember, know, familiar, and guess. Students first selected an answer and then selected a memory-awareness category. Failure to indicate either an answer or one of the four memory-awareness categories led to the question being failed. Answers with more than one option marked on either the answer options or memory categories also led to a fail on that question. Standard instructions were provided for the remember, know, familiar, and guess categories (cf. Gardiner, 1988; Dewhurst & Conway, 1994; Rajaram, 1993; Tulving, 1985a). Participants were instructed that they may have selected an answer because 1. You remembered a specific episode from one of the lectures. In this case you might have images and feelings in mind relating to the recalled information. Perhaps, you
virtually “hear” again or “see” again the lecturer presenting some item of information. Alternatively, you might have a specific memory of reading or talking about the topic. Answers such as these are called REMEMBER answers.

2. You might “just know” the correct answer and the alternative you have selected “stood out” from the three choices available. In this case you would not recall a specific episode and instead you would simply know the answer. Answers with this basis are called KNOW answers.

3. It may be, however, that you did not remember a specific instance, nor do you know the answer. Nevertheless the alternative you have selected may seem or feel more familiar than any of the other alternatives. Answers made on this basis are called FAMILIAR answers.

4. Finally, you may not have remembered, known, or felt the choice you selected to have been familiar. In which case you may have made a guess, possibly an informed guess, e.g., some of the choices look unlikely for other reasons so you have selected the one that looks least unlikely. This is called a GUESS answer.

The full instructions for each of the four response categories were projected on a large screen at the front of the lecture theater in which the class tests took place. Note that this was always the lecture theater or laboratory in which they had taken the course on which they were being tested. The lecturer verbally described the response categories and answered any questions concerning their use. Each test took approximately 40-45 min to complete. Once a test commenced, no talking was permitted, and candidates raised their hand to attract the attention of an invigilator (proctor) if a query arose. No one was permitted to leave the room until a test had been completed by all candidates. Candidates were advised to adopt the following strategy in completing the MCQ:

Go through the paper answering all questions which you can immediately answer. Then make a second pass working on those questions which you find difficult. When “5 minutes remaining” is announced mark your best guess for each question still outstanding.

Finally, it was explained to the students that the purpose in collecting the memory-awareness ratings was to attempt to gauge how learning was taking place and to use the data to fine-tune future MCQs. This was the 1st year these tests had been used, and students were urged to be as accurate as possible in their memory-awareness judgments because the resulting psychological data would be of great value in assessing the effectiveness of the courses. A revision day was also arranged in which the results of the memory-awareness responses would be reported. Results on overall level of performance, and degree class, were posted within 1 week of a class test. In order to proceed in their studies, students had to pass the class test, which was compulsory and contributed 20% to the award of credit points for each course, the remainder being made up by the course work essay (20%) and final (summer term) examination (60%). All three forms of assessment had to be passed if credit points were to be awarded. However, it was explained to students that the purpose of the class tests was primarily to provide them with immediate feedback on what knowledge they had acquired from the lectures and thus give them some means of gauging their state of understanding at that point.

Participants

Three types of students took first-year psychology. Single-
honors psychology students took all four lecture courses, the two Research methods courses, and the laboratory practicals. In total this accounted for two thirds of their first-year course, with the remainder being taken in a cognate discipline. Joint-honors students took two lecture courses, Research Methods(1) or (in 1995–1996) Research Methods(3) plus half of the laboratory classes, accounting for one third of their first-year courses. A further third was taken in the joint discipline (sociology, philosophy, or zoology), and the remaining third in a related topic. Subsidiary students, from degree tracks in the faculties of science, social science, and arts, took only the lecture courses, again accounting for one third of their first-year courses. There were 60 single-honors students, 30 joint-honors students, and 142 subsidiary students. These numbers varied with each course depending on the particular track followed by a student (see Table 1) and also upon whether or not they were present for the class test. Data from students who completed the class test at a later date are not included in the study.

Results

The design had three main variables: courses (with eight levels) degree class (with four levels), and memory awareness (four levels). University degrees in the United Kingdom are divided into four classes, from highest to lowest: first, upper second, second, and third; hence the four levels of this variable. Courses and degree classes were treated as grouping factors, and memory awareness was treated as a within-subjects factor. To anticipate our main finding, all three factors contributed to a complex but interpretable three-way interaction. We approached this interaction in stages, reporting first analyses of Courses X Memory Awareness followed by analyses of Degree Class X Memory Awareness, culminating in a section on the interaction of Course X Degree Class X Memory Awareness. This approach allowed us to focus on important details of the data that were not readily accessible in the full design and also to make contrasts between levels of variables that would not be possible with the full data set. In all the analyses, we examined the data expressed in two different ways, which focus on two different aspects of the data: quantity of correct answers and accuracy of correct answers. Koriat and Goldsmith (1994) have drawn an important distinction between the amount or quantity remembered compared to the accuracy or quality of what is remembered. For instance, in our class tests a student may have assigned very few answers to a particular response category, perhaps they have a low amount of know responses, but these were always correct. Such a student would have a low quantity of just know responses, but ones that were highly accurate. The quantity and accuracy analyses examined such differences. In the quantity analyses, the data were represented as a priori probabilities and this was accomplished by calculating the proportions of correct answers falling in each of the four memory-awareness categories for each student. In the accuracy analyses correct scores were represented as a posterior or posterior probabilities (Murdock, 1974). Posterior probabilities addressed the question: Given that a response is assigned by a student to one of the four memory-awareness response categories, what is the probability that the response is correct? In order to calculate posterior probabilities, the proportion of correct answers from the total of answers given in each memory awareness category
individually for each student was calculated. In each of the three sections that follow, quantity analyses are reported first followed by the corresponding accuracy analyses. Finally, interactions because of their complexity are presented in bar charts; however, Tables 1 and 2 present the mean probabilities, mean number correct, number of students taking a test, and number of questions in a test for quantity and accuracy, respectively. In these tables, the data are shown at the Level of Course × Degree Class × Memory Awareness, and all figures are derivable from them.

Part 1: Variations in Memory Awareness With Type of Course

Each course was taken by different numbers of students, and some students took all courses (single-honors psychology students), whereas the others, the majority, took a mixture of courses appropriate to their degree program (Table 1). For the purposes of these analyses, we treated courses as a grouping factor. The pattern of findings did not differ if the analyses were restricted to only those students

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Probabilities of Correct Answers Within Degree Class for Each of the Memory-Awareness Response Categories</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Course and response categories</th>
<th>Degree class</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First</td>
<td>Upper second</td>
<td>Second</td>
<td>Third</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>M</td>
<td>p</td>
<td>M</td>
<td>p</td>
</tr>
<tr>
<td>Introduction to psychology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Qs = 64)</td>
<td>N = 31</td>
<td>N = 74</td>
<td>N = 71</td>
<td>N = 25</td>
<td></td>
</tr>
<tr>
<td>Remember</td>
<td>.46</td>
<td>.24</td>
<td>.38</td>
<td>.18</td>
<td>.32</td>
</tr>
<tr>
<td>Know</td>
<td>.21</td>
<td>.11</td>
<td>.19</td>
<td>.9</td>
<td>.19</td>
</tr>
<tr>
<td>Familiar</td>
<td>.19</td>
<td>.10</td>
<td>.28</td>
<td>.13</td>
<td>.27</td>
</tr>
<tr>
<td>Guess</td>
<td>.14</td>
<td>.7</td>
<td>.15</td>
<td>.7</td>
<td>.22</td>
</tr>
<tr>
<td>Physiology psychology (Qs = 68)</td>
<td>N = 20</td>
<td>N = 71</td>
<td>N = 49</td>
<td>N = 56</td>
<td></td>
</tr>
<tr>
<td>Remember</td>
<td>.42</td>
<td>.24</td>
<td>.33</td>
<td>.16</td>
<td>.30</td>
</tr>
<tr>
<td>Know</td>
<td>.21</td>
<td>.12</td>
<td>.25</td>
<td>.12</td>
<td>.21</td>
</tr>
<tr>
<td>Familiar</td>
<td>.25</td>
<td>.4</td>
<td>.27</td>
<td>.13</td>
<td>.28</td>
</tr>
<tr>
<td>Guess</td>
<td>.12</td>
<td>.7</td>
<td>.15</td>
<td>.7</td>
<td>.21</td>
</tr>
<tr>
<td>Cognitive psychology (Qs = 60)</td>
<td>N=20</td>
<td>N=63</td>
<td>N=53</td>
<td>N=44</td>
<td></td>
</tr>
<tr>
<td>Remember</td>
<td>.49</td>
<td>.24</td>
<td>.39</td>
<td>.17</td>
<td>.39</td>
</tr>
<tr>
<td>Know</td>
<td>.18</td>
<td>.9</td>
<td>.23</td>
<td>.10</td>
<td>.18</td>
</tr>
<tr>
<td>Familiar</td>
<td>.24</td>
<td>.12</td>
<td>.24</td>
<td>.11</td>
<td>.29</td>
</tr>
<tr>
<td>Guess</td>
<td>.08</td>
<td>.4</td>
<td>.14</td>
<td>.6</td>
<td>.16</td>
</tr>
<tr>
<td>Social and developmental psy-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chology (Qs = 60)</td>
<td>N = 31</td>
<td>N = 52</td>
<td>N = 36</td>
<td>N = 45</td>
<td></td>
</tr>
<tr>
<td>Remember</td>
<td>.42</td>
<td>.21</td>
<td>.41</td>
<td>.18</td>
<td>.39</td>
</tr>
<tr>
<td>Know</td>
<td>.26</td>
<td>.13</td>
<td>.18</td>
<td>.8</td>
<td>.18</td>
</tr>
<tr>
<td>Familiar</td>
<td>.22</td>
<td>.11</td>
<td>.25</td>
<td>.11</td>
<td>.26</td>
</tr>
<tr>
<td>Guess</td>
<td>.01</td>
<td>.5</td>
<td>.16</td>
<td>.7</td>
<td>.16</td>
</tr>
<tr>
<td>Research methods(1) (Qs = 90)</td>
<td>N = 27</td>
<td>N = 27</td>
<td>N = 19</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Remember</td>
<td>.23</td>
<td>.18</td>
<td>.24</td>
<td>.16</td>
<td>.27</td>
</tr>
<tr>
<td>Know</td>
<td>.53</td>
<td>.41</td>
<td>.40</td>
<td>.26</td>
<td>.39</td>
</tr>
<tr>
<td>Familiar</td>
<td>.16</td>
<td>.12</td>
<td>.23</td>
<td>.15</td>
<td>.29</td>
</tr>
<tr>
<td>Guess</td>
<td>.08</td>
<td>.6</td>
<td>.14</td>
<td>.9</td>
<td>.14</td>
</tr>
<tr>
<td>Research methods(2) (Qs = 50)</td>
<td>N = 7</td>
<td>N = 19</td>
<td>N = 23</td>
<td>N = 11</td>
<td></td>
</tr>
<tr>
<td>Remember</td>
<td>.14</td>
<td>.7</td>
<td>.14</td>
<td>.6</td>
<td>.15</td>
</tr>
<tr>
<td>Know</td>
<td>.48</td>
<td>.24</td>
<td>.36</td>
<td>.16</td>
<td>.40</td>
</tr>
<tr>
<td>Familiar</td>
<td>.24</td>
<td>.12</td>
<td>.32</td>
<td>.14</td>
<td>.25</td>
</tr>
<tr>
<td>Guess</td>
<td>.14</td>
<td>.7</td>
<td>.18</td>
<td>.8</td>
<td>.20</td>
</tr>
<tr>
<td>Research methods(3) (Qs = 90)</td>
<td>N = 25</td>
<td>N = 22</td>
<td>N = 11</td>
<td>N = 16</td>
<td></td>
</tr>
<tr>
<td>Remember</td>
<td>.23</td>
<td>.18</td>
<td>.23</td>
<td>.16</td>
<td>.30</td>
</tr>
<tr>
<td>Know</td>
<td>.50</td>
<td>.38</td>
<td>.39</td>
<td>.27</td>
<td>.40</td>
</tr>
<tr>
<td>Familiar</td>
<td>.21</td>
<td>.16</td>
<td>.23</td>
<td>.16</td>
<td>.21</td>
</tr>
<tr>
<td>Guess</td>
<td>.07</td>
<td>.5</td>
<td>.15</td>
<td>.10</td>
<td>.10</td>
</tr>
<tr>
<td>Retest of introduction (Qs = 64)</td>
<td>N = 15</td>
<td>N = 19</td>
<td>N = 18</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Remember</td>
<td>.15</td>
<td>.8</td>
<td>.13</td>
<td>.6</td>
<td>.15</td>
</tr>
<tr>
<td>Know</td>
<td>.48</td>
<td>.25</td>
<td>.42</td>
<td>.20</td>
<td>.39</td>
</tr>
<tr>
<td>Familiar</td>
<td>.23</td>
<td>.12</td>
<td>.28</td>
<td>.13</td>
<td>.27</td>
</tr>
<tr>
<td>Guess</td>
<td>.13</td>
<td>.7</td>
<td>.17</td>
<td>.8</td>
<td>.19</td>
</tr>
</tbody>
</table>

Note. Within each degree class for each course the four probabilities total to 1. p = probability; M = mean number correct; Q = questions; — = not obtained.
who took all courses and courses treated as a within-subjects factor. Note that a problem with the data was that of dependency between the memory awareness categories. When a student selected one of the memory-awareness options (e.g., circles Remember), then obviously they could not select a further option, and in this sense the memory-awareness categories were dependent upon one another. In fact, this is a fairly common problem when more than one judgment is made of a remembered item (Rajaram & Roediger, in press). The approach taken here was to analyze the data for all variables together and then to conduct further analyses on each response category separately.

Figure 1 shows the mean a priori probabilities of correct responses assigned to the four memory-awareness response categories for each course individually. The data were entered into an ANOVA with course as a between-groups factor with eight levels and memory awareness as a within-subjects factor with four levels. A significant two-way interaction of Course × Awareness was found, $F(21, 2976) = 20.3, MSE = .03, p < .01$. It can be seen from Figure 1 that this interaction arises chiefly from variations in the remember and just know categories across courses. For the lecture courses, remember is the dominant category, whereas for research methods and for the retest for Introduction to Psychology know is the dominant response category. The familiar and guess categories also vary across courses but not systematically and not to the same extent as the remember and know categories. These differences were investigated further by conducting one-way ANOVAs on responses in each of the awareness categories separately with courses as the single (grouping) factor. Differences between means were examined by use of the Tukey-Kramer test with alpha < .05. There was a reliable main effect of course upon remember judgments, $F(7, 992) = 16.9, MSE = .71, p < .01$, and contrasts between means found that the three research methods and the retest of the introductory course differed reliably from all other courses with the exception of physiological psychology. Additionally, Research Methods(2) differed significantly from the physiological psychology course as did the retest, which in turn differed reliably from Research Methods(3). Taken together, these findings show that the three lecture courses, Introduction, Cognitive, and Social and Developmental, were associated with a significantly higher rate of remember responses than the courses in research methods and the retest of the introductory course. The course in physiological psychology, which had a slightly lower mean probability for remember responses, fell midway between the two types of courses. For know responses, the reverse pattern was observed and comparisons between the means in the significant main effect, $F(7, 992) = 41.5, MSE = 1.2, p < .01$, found significantly more know responses to the courses in research methods and the retest of the introductory course. The course in physiological psychology, which had a slightly lower mean probability for remember responses, fall midway between the two types of courses. For know responses, the reverse pattern was observed and comparisons between the means in the significant main effect, $F(7, 992) = 41.5, MSE = 1.2, p < .01$, found significantly more know responses to the courses in research methods and the retest of the introductory course. For both these variables, however, there were few significant differences between means (compared to the remember and know variables), and they did not form any systematic pattern. These findings demonstrate that episodic memory plays a critical role in the acquisition of knowledge on lecture courses but appears to play no central role in the acquisition of knowledge on courses in research methods. The findings also demonstrate a shift in the dominant response class for the Introduction to Psychology course,
which initially was associated with a preponderance of correct remember responses but when retested at the end of the academic year was dominated by correct just know responses. One further point that might be noted here is that because of the timing of the class tests, students took lecture courses in parallel with research methods courses and (the same) students were simultaneously producing different memory awareness profiles on the lecture as compared to the research methods courses.

Figure 2 shows the posterior probabilities for awareness categories by courses. Note that all these means are reliably above chance. An identical ANOVA to that conducted on the a priori probabilities found a significant interaction between courses and memory awareness, $F(21, 2971) = 4.9$, $MSE = .075$, $p < .01$. In the separate analyses of response classes, reliable main effects were found for all classes. For know, familiar, and guess, there were, however, few significant differences between means and no systematic pattern was present. Possibly, these few unsystematic differences reflect very specific variations between the courses and tests that are beyond the level of detail of the present study. For remember judgments there were more reliable differences between the means and some suggestion of an emerging systematic pattern. Significantly more accurate remember responses were produced on the tests of the introduction, physiological, and cognitive courses than on the three courses in research methods, demonstrating a decline in the accuracy of remember responses on these latter courses. This is of some interest because, in contrast, it can be seen from Figure 2 that know responses do not systematically vary in their accuracy across courses, with the exception of know responses on Social and Developmental, which were significantly lower than on all other courses. This contrasts with Figure 1 and the analyses of the quantity responses in which fewer know responses were found to be given on lecture courses. An implication is that regardless of the number of know responses given on a test of a course, they have a generally high probability of being correct. The same is not generally true of remember responses because as these fall in quantity, they also diminish in accuracy, or at least this is the case when the courses in research methods are compared to the first three lecture courses.
Part 2: Degree Class and Memory Awareness

The previous analyses demonstrated that memory awareness varies with the type of course tested and also can shift over time from a profile initially dominated by recollective experience to one dominated by "just knowing." A persistent and high degree of accuracy was also found, across all courses, for those responses allocated to the just know category. In this section we investigate whether these general characteristics of memory awareness following different types of learning are also present when the degree class of the student is taken into account.

The lecture courses. For the analyses of both the lecture courses and courses in research methods, two variables were of interest: the degree class obtained for overall mark and the distributions of correct and incorrect responses to the four memory awareness categories. In order to determine degree class on each course, the total number correct was adjusted for guessing by use of a high-threshold model. Total correct was then mapped onto a 21-point scale used to assign degree class. Marks between 20 and 15 were classed as a First, 14 to 12 as an Upper Second, 11 to nine as a Second, eight to six as a Third, and five and below as various levels of fail. The variable degree class formed the grouping variable in the analyses. The distributions of students to degree classes for each course are shown in Table 1. Note that for Research Methods(1) and the retest for Introduction to Psychology there were very few students with thirds and thus this degree class was omitted and the data included with the data of students with seconds. There were a few fails on some of the courses, and data from these students were included in the third-class degree class. The analyses were the same as those conducted previously, with degree class replacing courses.

In an initial analysis that included the four lecture courses as an additional grouping variable no differences were found between lecture courses ($F < 1$) nor did the variable lecture course interact with the other variables. Consequently, individual analyses of each lecture course are not reported here, and instead the data are collapsed over lecture courses. In the quantity analyses there was a reliable interaction of degree class with memory-awareness categories, $F(9,$
MEMORY AWARENESS AND LEARNING

In the comparisons between means for the remember responses, $F(3, 737) = 19.3$, $MSE = .77$, $p < .01$, it was found that all the means differed reliably from each other with the exception of upper seconds and seconds. Thus, correct responses, which students recollectively experienced reliably, declined with degree class, and students with firsts had more remember responses across all courses than students with lower degree classes. There was no significant main effect of degree class upon know responses, which did not differ across degree classes and, instead, remained at a constant level. There was a reliable effect in the familiar class of responses, $F(3, 737) = 12.5$, $MSE = .17$, $p < .01$, and again all means differed significantly from each other with the exception of upper seconds versus seconds. This shows a reliable and steady increase in correct responses associated with a feeling of familiarity with decreasing degree class: Students who obtain thirds across all courses had more correct answers accompanied by feelings of familiarity than students with higher degree classes. Finally, there was an effect of degree class in guesses, $F(3, 737) = 22.6$, $MSE = .31$, $p < .01$. Firsts and upper seconds did not differ in the relatively small numbers of correct guesses that they made, whereas seconds and thirds did not differ in their larger number of correct guesses. All other comparisons were, however, significant demonstrating an increase in correct guessing with decreasing degree class.

In the accuracy analyses, using the same ANOVA design, there was a significant interaction of degree class with memory awareness, $F(9, 2211) = 2.6$, $MSE = .07$, $p < .03$, and Table 2 shows the mean posterior probabilities for each lecture course. In the comparisons between means on the remember responses, $F(3, 737) = 42.1$, $MSE = 1.4$, $p < .01$, firsts and upper seconds did not differ reliably, but all other comparisons were significant showing that accuracy of remembered answers decreased with degree class. For know responses, $F(3, 737) = 26.1$, $MSE = .91$, $p < .01$, thirds had reliably poorer accuracy than all other degree classes who in turn did not differ in the accuracy of their know responses, which generally were at a high level. For the remaining two response classes of familiarity, $F(3, 737) = 105.1$, $MSE = 1.8$, $p < .01$, and guesses, $F(3, 737) = 26.1$, $MSE = .91$, $p < .01$, all contrasts were significant for both variables showing a reliable decrease in accuracy with decreasing degree class; but note that this rarely led to performances significantly below chance (see Table 2).

In summary the pattern of findings for correct answers shows a remarkable consistency across the four lecture courses: Students who do well in examinations administered immediately at the conclusion of a lecture course do so because they remember more than students who perform at lower levels. In turn, students with lower marks showed no distinctive pattern in their correct responses in that they were just as likely to remember, know, or have a feeling of familiarity as they were to guess a correct answer. Although not reported here, it should be noted that it makes no substantive difference to the findings if degree class is determined by the overall mark awarded for the whole of the first year performance, or if degree classes from one course are used as the grouping factor on another course. In other words the students who do well on one course by and large do well on all courses, and in terms of the four lecture courses reported in this section, they do so because they are able to remember more than their colleagues with lower grades. As regards accuracy of responses, this was found to systematically decline across degree classes for all awareness categories (see Table 2).

Courses in research methods. As with the lecture courses it was found that the research courses when entered as an additional grouping factor in the ANOVA produce no significant effects. In this section then we report ANOVAs identical in design to those used for the lecture courses collapsed over the three research methods courses. Research Methods(3) from 1995–1996 was included because we were concerned that the analyses were sampling fewer research methods courses than lecture courses. It attests to the robustness of the pattern of findings that the same effects were observed for this later course, which was taught by different lecturers, covered slightly different topics, and featured a new class test and a new cohort of students. In the analyses of the quantity data a significant interaction of degree class with memory awareness was observed, $F(9, 609) = 2.7$, $MSE = .12$, $p < .01$. It can be seen from Table 1 that this arises from the dominance of know responses in the higher degree classes, which declines with falling degree class. Comparison between means in each individual memory-awareness category found no reliable main effects or contrasts for remember and familiar responses. Thus, responses in these two categories did not vary with degree class. There was a significant main effect of know responses, $F(3, 203) = 5.5$, $MSE = .23$, $p < .01$, and all comparisons were significant with the exception of upper seconds versus seconds. There was also a significant main effect of guess responses, $F(3, 203) = 4.2$, $MSE = .08$, $p < .01$. Firsts had reliably fewer correct guesses than all other degree classes. In striking contrast to the lecture courses, the class tests in research methods were then dominated by know rather than remember responses. This was the case for all degree classes (Table 1), although it was most evident for students with first-class marks who outperformed their colleagues because they knew more. These findings are in marked contrast to the same students’ performance in tests for the lecture courses in which remember was the dominant category overall and the category that differentiated higher from lower grades. The accuracy analyses found no reliable interaction between degree class and memory awareness. Instead there was a main effect of memory awareness, $F(3, 609) = 65.9$, $MSE = .05$, $p < .01$, and the mean probabilities for each category

---

1 In the academic year 1995–1996, the same courses were taken again by a new cohort of students. All the courses had been modified to various degrees, and all the tests contained new materials. Despite these changes, the findings are virtually identical to those reported in this article, contain no substantial differences, and have highly comparable levels of performance, which provides strong support for the stability of these effects.
were: remember, .69; know, .85; familiar, .63; and guess, .44. Least square means comparisons found significant differences on all pairwise contrasts. Thus, overall know responses were significantly more accurate than all other responses, remember responses were more accurate than familiar and guess, and familiar was more accurate than guess. Analyses of each variable individually found that only know responses yielded a significant main effect over degree class, $F(3, 203) = 10.4, MSE = .18, p < .01$. In the Tukey-Cramer contrasts, firsts had more accurate know responses than seconds and thirds but did not differ from upper seconds, and upper seconds had more accurate know responses than seconds and thirds, who in turn did not reliably differ.

A potential problem with these differences between the patterns of performance on the lecture courses versus the courses in research methods is that the MCQs for the latter contain some items that require answers that may not have been directly presented in the classes. This leaves open the possibility that the different patterns of responding may be related more to question type than change in underlying knowledge representations. Fortunately, for Research Methods Part(1) taught in 1994–1995 half the questions in the class test explicitly required reasoning in order to select a correct answer, whereas the remainder of the questions could be answered by recalling knowledge directly presented in the classes. We had hoped that this division into conceptual versus episodic questions would produce a pattern of responding in which answers to conceptual questions were dominated by know responses and answers to episodic questions by remember responses. These predictions were not supported. Answers to both types of questions were dominated by know responses and for this reason we do not report the full analyses of this variable here. However, the fact that half the questions could be answered by recalling a specific episode from a class allows us to test the conjecture that it is question type rather than underlying knowledge representation that produces the differing patterns of memory awareness in the research courses compared to the lecture courses. If question type is the critical variable then questions that can be answered by recalling a learning episode, rather than drawing on conceptual knowledge, should be dominated by remember responses rather than know responses. For the episodic questions from Research Methods(1) there was a main effect of memory awareness, $F(3, 210) = 23.2, MSE = .77, p < .01$, and the mean a priori probabilities were: remember, .21; know, .34; familiar, .15; and guess, .09. There were significantly more know responses than remember responses, $t(210) = 3.95, p < .01$, and overall, this response category was the dominant response category for these episodic questions (an identical pattern to that observed for the conceptual questions on the same test). These findings show that question type is not the critical variable mediating different patterns of memory awareness between the two types of courses (lecture versus research), instead, and as we elaborate later, it is the type of knowledge used to answer questions that gives rise to the different patterns of memory awareness.

Retest of the introduction to psychology course. As part of an optional revision day students were invited to retake any of the class tests and compare their current performance with past performance. The revision day took place 4 days prior to commencement of the summer examinations, and it was not possible to directly contact all students who had taken all courses. Instead the revision day was advertised on a first-year notice board and ultimately attracted only a small number of students consisting entirely of single- and joint-honors psychology students. Moreover, because of pressures on the time available, only data for the introductory course were usable (data for other courses were collected, but the examination conditions used previously were not adhered to and, consequently, these data were deemed unusable). Fifty-two students retook the original class test for Introduction to Psychology during the first session in the revision day and was conducted under the same conditions as the original test. The retest was administered 25 weeks after the original test had been taken.

In the quantity analyses significant effects of degree class, $F(2, 49) = 146, MSE = .01, p < .01$, and memory awareness, $F(3, 147) = 35.7, MSE = .02, p < .01$, were observed and the mean probabilities are shown at the bottom of Table 1. The striking aspect of these means, compared to the means from the original test (top of Table 1), is that students with higher degree classes now assign most of their correct answers to the know rather than remember category. In the individual analyses of response categories only know produced a significant effect, $F(2, 49) = 4.1, MSE = .02, p < .025$, and comparisons between means established that significantly more know responses were given by firsts than seconds. In the accuracy analyses, similar effects were observed, and because of the similarity, we do not report these further. The changed pattern of correct responses from test to retest shows the same shift from remember to know responses found earlier in the analyses of the courses in research methods (Table 1). In order to examine this directly, an additional analysis was conducted. In this analysis, students' scores (correct answers expressed as a priori probabilities) on the original test were paired with their scores on the retest, and test–retest formed a within-subjects variable in the ANOVAs that were otherwise the same as those already reported (degree class was determined by performance on the retest). All three main effects were significant: degree class, $F(2, 49) = 62.3, MSE = .01, p < .01$; memory awareness, $F(3, 147) = 11.3, MSE = .021, p < .01$; and test–retest, $F(1, 49) = 7.7, MSE = .01, p < .01$. All three main effects featured in two-way interactions. The interaction of degree class with test–retest, $F(2, 49) = 3.9, MSE = .01, p < .03$, arose because only students with firsts and upper seconds had higher scores on the retest compared to the original test. The means for the interaction of memory awareness with test–retest, $F(3, 147) = 36.9, MSE = .02, p < .01$, are plotted in Figure 3. It can be seen from Figure 3 that the critical differences lie in the patterns of remember

\footnote{We thank David Pillemer for drawing this point to our attention.}
and know responses across the two tests. Separate analyses found that these differences were significant for remember, F(1, 49) = 42.5, MSE = .03, p < .01, and know, F(1, 49) = 57.4, MSE = .05, p < .01, but not for familiar and guess (F < 1 in both cases). Thus, the critical change in memory awareness as learning progresses is from remember to know. Note that, as the three-way interaction was not significant (F < 1) and the two-way interaction between degree class and memory awareness also failed to reach significance, (F < 1), then the shift from remembering to knowing is general and present at all levels of degree class, although quite clearly this shift was more marked for students with firsts than students with lower grades. Finally, the accuracy analyses were conducted on the test-retest data, and Table 2 shows the mean probabilities. There were reliable effects of degree class and test versus retest, but as these parallel the findings from the quantity analyses we do not reported them further here. The only other significant effect was a main effect of memory awareness, F(3, 147) = 58.1, MSE = .04, p < .01. Mean accuracy probabilities were remember, .76; know, .85; familiar, .67; and guess, .49. Least square means comparisons found that all four means differed reliably, thus the rank ordering in terms of accuracy across both test and retest was know > remember > familiar > guess.

The major finding to arise from the retest for Introduction to Psychology was the shift in the dominant response category for correct answers from remember to know, whereas the response rates remained comparable in the familiar and guess categories. As we hypothesized earlier, this may reflect a shift in long-term memory representations that are initially predominantly episodic in nature to a more conceptual style of representation. This shift is most marked in students with firsts, and students with seconds did not show any enhancement of performance on the retest. These findings suggest that students with lower grades also develop more conceptual or semantic types of long-term memory representations but that these do not support an increased level of overall performance.

Part 3: Type of Course, Degree Class, and Memory Awareness

The preceding analyses examined the relations between course and memory awareness and separately between degree class and memory awareness. In this final section, we briefly consider the three-way interaction of Course × Degree Class × Awareness. In this ANOVA, the grouping factors are course with two levels, lecture courses versus research methods courses plus the retest of the introductory course, and degree class with four levels: first, upper second, and third. Degree class is assigned on the actual degree class the students obtained on the first-year course overall; for Research Methods(3), the mark from 1995–1996 was used. The within-subjects factor is memory awareness. In the quantity analyses, all main effects and interactions were significant, and here we focus on the Course × Degree Class × Memory Awareness interaction, F(9, 2976) = 3.1, MSE = .03, p < .01. The mean probabilities plotted in Figure 4 indicate two main influences on memory awareness.

The type of course (lecture, research methods, and retest) determines which awareness category is dominant. For the tests of lecture courses, remember is the dominant response category, whereas in the tests of research methods and in the retest, know is the dominant response category. This effect of course is a function of the degree class attained such that for remember responses in the test of lecture courses, the effect declines with level attained and is not present in students obtaining the lowest grade. For know responses, too, in the tests of research methods and the retest, the effect declines with failing degree class but is still present, albeit in diminished form, for students with the lowest degree class. It is particularly notable that for the tests of the lecture courses, know responses do not vary with degree class, whereas in...
the tests of research methods, it is the remember responses that do not vary. Paralleling these systematic differences are complementary effects upon familiarity and correct guesses, both of which increase with decreasing degree class for both types of course (see Figure 4). The increase in familiarity judgments is especially evident in the tests of the lecture courses, whereas for third-class students it eventually exceeds remember responses. In our view, these differences may largely reflect a reliance on different types of knowledge in answering the test questions on the two sets of courses. In responding to tests on lecture courses, when these are taken immediately after a course is completed, then the high-scoring students rely very extensively on episodic memories of learning episodes in which the sought-for knowledge had previously been encountered and is now retained. In responding to tests of research methods and in the retest of the introductory course, responding is mediated largely by conceptual rather than episodic knowledge. We
believe that this reflects a remember-to-know shift, and we develop this suggestion in the next section.

In the accuracy analyses the Course × Degree Class × Memory Awareness interaction was again significant, \( F(9, 2976) = 2.2, MSE = .04, p < .02 \). Figure 5 shows the mean probabilities (Table 2 shows these for individual courses), and it can be seen that in the tests of the lecture courses, the accuracy of remember and know responses was high. The accuracy of familiar and guess responses was lower but usually above chance (see Table 2). In the tests of research methods and the retest know responses were the most accurate. Thus, know responses were consistently accurate over all the courses tested. Over both sets of courses accuracy declined with falling degree class. Of note here is that for students with thirds, remember responses were the most accurate over all courses (Figure 5) but were not the
most frequent class of responses for these students (Figure 4). From this it seems that what little these students do remember, they remember fairly accurately.

General Discussion

In this study we addressed four central questions: Do episodic memories play a role in this type of long-term learning? Do remembering and knowing vary with overall level of performance? Do states of awareness change during the process of knowledge acquisition? And, do states of awareness vary with different courses? In answer to the first and second of these questions the findings show that when knowledge acquisition is tested immediately after a lecture course has been completed, students with higher marks outperform their colleagues because they remember more (Table 1). In the case of the third question, it was found that when students were retested after a filled retention interval during which additional learning took place, the dominant response category shifted from remembering to knowing (Figure 3). In answer to the fourth question, differences between courses were observed, and in contrast to the dominance of remember responses in the tests of the lecture courses, we found that when knowledge was acquired in a prescribed knowledge domain and in a more active learning schedule, as in the laboratory-based teaching of research methods, then all the better students performed at a higher level because they knew more (Table 1 and Figures 1 and 4). In addition to these three main sets of findings, it was also found that knowing was independent, in terms of accuracy, of the number of correct answers (Figure 2). Accuracy of remember responses over courses was, however, found to systematically decline with diminishing number of correct answers (Table 2). The findings have various implications for accounts of the relation between states of memory awareness and long-term knowledge, and also for practical issues relating to teaching and assessment. These implications are considered in the discussion that follows.

Why Should Students Know More, and Remember Less, at Retest?

The second finding of particular interest to us was that of a remember-to-know shift, referred to hereafter as the R-to-K shift. This was most strikingly present in the retest of the introductory course. The R-to-K shift was, however, most probably occurring continuously on all courses but at different rates in different courses. In the lecture courses the rate was slow compared to the courses in research methods. Furthermore, the R-to-K shift, although present in all degree classes, was stronger in those students who attained the highest marks. It seems to us that the R-to-K shift probably reflects two processes: first, the fairly rapid loss of the ability to retrieve specific episodic memories and, second, the increasing availability of more extensive semantic or conceptual representations. By this account, the relative preponderance of remember responses in an initial test, and know responses in a later retest, reflects a decrease in the accessibility of episodic memories. At the same time as this occurs there may, however, be a corresponding development of schematized knowledge representations to which access becomes increasingly automatic. Such structures may develop in order to impose organization on the remaining nonepisodic knowledge that would have been organized previously in terms of relations between different memories (cf. Conway, 1996). Indeed, something like this must have occurred in order for students to be able to pass tests such as the MCQ on Research Methods(2) and, of course, in their summer examinations that were essay-based and that required answers that demonstrated a level of conceptual understanding. It is notable that poorer answers in the summer examinations were those that simply listed facts and concepts, whereas answers attracting higher grades showed conceptual organization of knowledge. In other words, if all a student could bring to an examination were an unrelated set of concepts arising from degraded episodic memories then they would be unlikely to achieve any but the lowest of grades. One additional point that should be noted here is that the cohort of students whose responses contributed to this data set were explicitly instructed (repeatedly and at length) that higher grades were awarded to work that showed good evidence of insight and understanding, and they were directly warned of the dangers of the listing approach to essay writing. This, in itself may have been sufficient to motivate students to active attempts to organize and associate concepts in a coherent manner.

The two processes we have postulated to bring about the R-to-K shift, loss of episodic details and the emergence of conceptual organization, are linked at least insofar as they both result from repeated encounters with much the same
information in rather different contexts. The students encounter the same concepts and facts repeatedly across lectures and laboratory classes, and because of this repetition there may be some strengthening of knowledge common to (episodic) representations in long-term memory. The learning contexts in which the same concepts are presented, however, vary more in their episodic details (i.e., occur at different times of the day, in different lecture theatres, possibly with different lecturers, different overhead projections, and different numbers of students; the student may sit in a different seat for each lecture next to different students, and so on). Given that access to episodic details is fairly rapidly lost, then the effects of repeatedly encountering the same concepts and facts may further promote a change in the accessibility of the underlying representations: a change from easy access to episodic memories to easy access of semantic knowledge, free of any information about the learning contexts in which the knowledge was acquired. Thus, an important factor leading to a preponderance of know responses may be the extent to which knowledge has been schematized, and a first step in schematization may be this loss of episodic details.

Consistent with this account, Gardiner and Java (1991) investigated forgetting in recognition memory as measured by remember and know responses over retention intervals ranging up to 6 months and found that remember responses declined much more rapidly than know responses, particularly over shorter retention intervals (see Knowlton & Squire, 1995, for related findings). Know responses in that study (which correspond with familiarity in the present study) declined slowly and mainly at the longer retention intervals. Evidence of schematization is less direct, but as we argued earlier schemas are complex semantic representations that encompass many different types of knowledge in a proscribed domain and do so in such a way as to facilitate access and problem solving (cf. Anderson, 1987). One way to think about this is to view individual memories and concepts as being like books in a library and schemas as representing the library indexing system. Without an index, books can only be accessed by reading their titles, and related volumes require direct and laborious comparison to establish their relatedness. The benefit of an indexing system (schemas) is that it supports fast and content-free access to individual volumes (memories or concepts) and to subject areas (subdomains of knowledge). The organizational properties of an index or schema make usable a library of many volumes or a domain of many concepts. But an index does not simply arise from the fact that a library contains many volumes. Instead, an index must be effortfully constructed. In like manner, a learner must make a sustained and active attempt at learning if schematic knowledge structures are to be generated. Repeated exposure to facts, rules, concepts, and so forth may lead to representations of concepts that are semantic in nature (i.e., devoid of episodic details: the books in our library analogy), but repetition will not in itself lead to extensive schematization. Furthermore, despite the obvious advantages of an index to a library or schemas to a complex knowledge base, there are costs entailed in creating either form of organization, and if the number of volumes in the library or the number of concepts recently acquired is small, then the costs may outweigh the benefits. At some point in the process of acquisition a learner reaches a point where the requirement for schematic organization is pressing. Just as a library beyond a certain number of volumes becomes unusable without an effective indexing system, so does unorganized semantic knowledge; the students taking the lecture courses are, we suggest, at this point. They have acquired knowledge of a number of different subdomains of psychology. Much of this knowledge is in the form of episodic memories (hence the dominance of remember responses in the tests on these courses), but as these memories increase, they rapidly become inaccessible and now some type of organization is required if the already acquired knowledge is to support problem solving and further learning. At the close of a lecture course then, the student has the opportunity to begin the process of schematizing their knowledge: They have sufficient knowledge to make schematization a benefit, and this knowledge, because of frequency of exposure, is in the process of shifting into a semantic form.

### Why Should Students With Higher Marks in Tests Following Methods Courses Know More, Instead of Remembering More?

The point at which schematization of knowledge becomes possible in the courses in research methods would appear to occur while the courses are actually being studied rather than after the whole course has been completed. This is suggested by the finding that these students show a predominance of know responses in their class tests. There may be several reasons for this earlier schematization of knowledge in the courses on research methods compared to the lecture courses. As we hypothesized earlier, it may be that this area lends itself more readily to schematization than the materials covered in the lecture courses (cf. Neisser, 1984, and Conway et al., 1991, for related proposals). One reason for this, apart from any qualitative differences between the two types of courses, was that the lecture courses covered more topics than the courses in research methods. Thus, as a number of subdomains of knowledge had to be schematized in the lecture courses (e.g., history of psychology, normal and neuropsychology of memory and perception, abnormal

---

4 The research methods courses were taught in 12 three-hour laboratory classes, whereas the lecture courses were taught in 17 one-hour lecture slots (the 18th being reserved for the class test). Thus, a course in research methods had 24 hr of teaching, whereas a lecture course had only 17 hr. Perhaps this discrepancy in total time taught is sufficient to account for the differences in the patterns of memory awareness between the two courses? The differences in total time taught are, however, somewhat misleading. Students also took tutorials that ran concurrently with a lecture course and focused almost exclusively upon the lecture course currently being attended and on the upcoming course essay. There were no tutorials for the courses in research methods. Thus the differences in total time taught are far less than the course timetables suggest, and quite possibly more hours of formal teaching were devoted to the lecture courses compared to the courses in research methods.
psychology, and areas of social and developmental psychology, in the introductory course alone), the schematization process was slowed compared to the same process in the acquisition of knowledge of research methods. In addition to this, research methods as a self-contained knowledge domain may inherently lend itself to schematization. For example, specific experimental designs logically entail selection of only certain statistical tests, whereas the concept of "the unconscious" does not logically entail the concept of "repression," although the two are conceptually related in psychoanalytic theory. In this way the knowledge to be acquired in research methods may have predisposed the learner to rapidly form indices between concepts.

An additional factor that may have led to the preponderance of know responses in research methods, in all degree classes, is that on these courses the students encountered the to-be-acquired knowledge repeatedly across classes, often paired with different concepts across contexts, and in different problems. It is possible that because of the repetition of material across multiple contexts and the interactive problem-solving orientation of the learning, the schematization process is facilitated and an R-to-K shift occurs during, rather than after, the course. If the student accepts the opportunity provided by the repetition and active learning schedule and schematizes their knowledge, then a predominance of know responses becomes evident in the tests on these courses. The lower part of Table 1 shows that this was indeed the case for the research methods courses and was most strikingly evident for students who obtained the higher grades. However, the emergence of the just know state as the dominant state was also present for students with lower grades. This perhaps suggests that a learning schedule that repeatedly presents knowledge in different contexts in active learning exercises has a general beneficial effect in facilitating the schematization of knowledge and this effect is at least partly independent of the cognitive effort expended by the learner and of initial differences in the amount of recollection. Put another way, although high initial levels of recollection may facilitate schematization, schematization may not depend on high levels of recollection and indeed it may take place in the absence of much episodic memory at all (cf. Tulving, Hayman, & MacDonald, 1991).

Despite the findings supporting the proposal of a change in memory awareness dependent upon the schematization of underlying knowledge representations, it is clear that this was not an absolute change. The findings do not show a complete shift from remembering to knowing across the two types of courses but rather indicate a change in the dominant response category. In the tests of research methods, for example, all groups had some correct answers designated as remember responses and, similarly, in the tests of the lecture courses there was a notable rate of know responses. Thus, although the change from remember to know is marked across courses, both response categories are set in the context of a pattern of responding that encompasses all response categories. This accords well with the earlier discussion of the schematized nature of long-term knowledge acquired through the process of formal education (Conway et al., 1991; Bahrick, 1984; Neisser, 1984; Semb & Ellis, 1994) and with the notion of an embedding of episodic and semantic memory (Tulving, 1985b), and it lends some support to the schema-plus-episodic view discussed earlier. By these views, students may continue to remember the first time they encountered the notion of a within-subjects design while also developing and maintaining a complex schematic knowledge structure for experimental design as a whole. As learning proceeds and schematic knowledge structures are formed, these structures may maintain indices to some episodic memories of specific learning episodes, much as in Linton's (1986) account of her development of a script for academic meetings coupled with retention of a few specific memories for the (many) meetings at which she had been present. The important point suggested by our findings is that as knowledge becomes progressively schematized then the just know state comes to dominate and characterize the phenomenal experience that accompanies knowledge access. In summary, there are both marked regularities and marked R-to-K shifts in the nature of memory awareness associated with different courses and with the same course as learning proceeds. In these respects, the type of memory awareness that dominates at any particular time can be considered an index or diagnostic of the nature of the underlying memory representations.

Memory Awareness and Accuracy

Unlike the analyses that focused on number of correct answers (Table 1), the accuracy analyses (Table 2, Figures 2 and 5) did not find that remember responses differentiated the better from the poorer students on the lecture courses. Instead, accuracy of remember and know responses was at a similar high level, and both types of response declined with falling degree class. It is also notable that accuracy did not fall below chance for any degree class in the remember and know categories (Table 2). This shows that despite the lower number of correct answers assigned to the know category compared to the remember category (Table 1) in the tests of the lecture courses, accuracy between the two classes of response was comparable within degree classes and was generally high. In contrast, accuracy of remember responses was lower in the four tests in which know responses dominated (lower part of Table 2, Figures 2 and 5). This suggests that if, as we propose, schematic knowledge structures mediate performance, there will be a high level of accuracy that occurs regardless of the number of responses mediated by schematic knowledge. In other words, just know responses, whether infrequent or frequent, are highly likely to be correct. Remember responses only attain the same levels of accuracy when they are frequent. If, as we suggested earlier, part of the process of schematization involves a loss of access to episodic detail, then the falling accuracy and falling numbers of remember responses may reflect a widespread diminution of the clarity and accessibility of episodic memories of specific learning experiences. And, consequently, in the efficacy of this type of memory representation to support correct responses. Thus, it is only when recollection dominates responding that accuracy is high. The inescapable conclusion is that the just know state
is a generally more reliable indicator of an accurate answer than is the recollective state.

Another important aspect of the accuracy analysis is the above-chance guessing rates attained by students with first-class degrees in six out of the seven tests. This outcome contrasts markedly with findings from recent laboratory studies, which show that when in recognition memory participants report guessing, as well as remembering and knowing, their guess responses show no memory for the study episode, that is, guess responses to studied items do not significantly exceed guess responses to nonstudied items (Gardiner, Java, et al., in press; Gardiner, Kaminska, Dixon, & Java, in press). In those studies, of course, studied and nonstudied items were counterbalanced, whereas such counterbalancing is not possible in this kind of study. It seems likely, therefore, that accurate guess responses in this study are based on an educated or informed guessing strategy, in which students can use their general background knowledge to work out which of the alternatives is the more likely, and which the least likely. This would explain why, by and large, accurate guessing was achieved by only those students with high grades. It should be possible for these students to do this successfully, at least on some occasions, because the incorrect alternatives were taken from information presented in the same lecture in which the target was presented, and the incorrect alternatives were selected to be plausible and conceptually related to the target. The abler students would be better placed to reject the incorrect alternatives, even when they do not know the correct answer. Finally, familiarity responses were generally more accurate than guess responses and less accurate than remember and know responses, and they showed a similar trend of decreasing accuracy as a function of decreasing degree class. These responses may reflect partial access to episodic memories of learning experiences, access to memories that have lost much of their detail, or both. They provide the basis for some correct answers but also give rise to many incorrect answers, especially for students with low overall marks on the tests of the lecture courses (Figure 4).

Memory Awareness and Performance: Some Educational Considerations

The shift from remembering to knowing is exactly the change that educators surely want to see in their students, and this shift was more pronounced in students with higher grades. The implication is that those students who are able to engage episodic memory more effectively than their colleagues have some advantage in the educational process, and further studies of individual differences in this respect could have important educational implications. For example, it would be of interest to determine whether these differences reflect specific motivational factors associated with studying, or with differences in study skills, or whether these differences might reflect more general differences in mnemonic ability that would be apparent in most situations, not just in the context of studying for an undergraduate degree.

Another key issue for educators is the speed with which schematic knowledge structures can be developed. Once a coherent conceptual structure is in place, this can be used in the acquisition of further knowledge by providing a means for imposing organization on newly presented knowledge and will facilitate encoding (Anderson, 1987). Encoding using already schematized knowledge can also support widespread accuracy in responding (Table 2 and Figures 2 and 5). Thus, the faster these structures are established the sooner they can be used in the accumulation of knowledge that is detailed and durable. One important finding of the present study was that the research methods courses always lead to a dominance of know responses, even at immediate test in the last session of each course, and that recollective experience was at a low level that did not vary with degree class (Table 1). As proposed earlier it may be the case that knowledge of research methods is more conducive to schematization and conceptual representation than knowledge from other areas. More generally, it must be the case that some types of knowledge lend themselves to schematization more readily than other types of knowledge, but these differences are often confounded with level of difficulty of the to-be-learned materials, for example learning a route from one's department to the library versus learning how to use the library's on-line catalogue. In the case of this study, the research methods course was undeniably the most difficult one students studied in their first year (a fact attested to by formal student feedback from all courses), and yet the findings indicate that it was the most rapidly schematized. A suggestion arising from this aspect of the findings is that courses designed with schematization in mind and taught in interactive sessions, with repetitions of concepts in different contexts, are most likely to promote the formation and retention of organized long-term conceptual knowledge.

Finally, consider the issue of the timing of assessment tests and how this could relate to changes in memory representations reflected in shifts of states of memory awareness. Consider what might have occurred if knowledge acquisition on the lecture courses was only assessed prior to the underlying change from episodic to schematic knowledge, indexed by the R-to-K shift. For the courses sampled in this study, if there had been no summer examinations of the four lecture courses and if the class test and course work had been the sole means of assessing and determining level of learning, then for many students (particularly subsidiary students who were not pursuing further courses in psychology) learning could cease shortly after completion of the lectures. Our findings suggest that at this point knowledge is still in the process of shifting from an episodic to a conceptual form and, therefore, subject to rapid forgetting. It is not until some time later, following many further exposures to the course material and protracted independent study, that knowledge fully shifts to a schematic and more stable and durable form. Thus, assessments that occur close in time to the presentation of new material and that once taken eliminate the requirement of further study virtually ensure that knowledge is not primarily represented conceptually and in a durable form. Quite clearly an undesirable outcome for any course in higher education.
References


Appendix

Some Examples of Questions From the Tests

Introduction to Psychology

In Pavlov's famous experiments hungry dogs learned to salivate to the presence of a light. In these studies what was the UNCONDITIONED RESPONSE (UCR)?

(a) food    (b) the light    (c) salivation
Remember    Know    Familiar    Guess

Patients with neurological damage to Broca's area can no longer:

(a) produce fluent speech
(b) understand spoken language
(c) correctly perceive spoken words

Remember    Know    Familiar    Guess

Physiological Psychology

Sound pressure waves are converted into movements of the ossicles (bones of the inner ear) by the:

(a) tympanic membrane    (b) patella    (c) sustentaculum
Remember    Know    Familiar    Guess

The cortical layer which is most densely innervated by sensory fibres from the thalamus is:

(a) layer I    (b) layer IV    (c) layer VI
Remember    Know    Familiar    Guess

Research Methods (Part 1)

In a study of anti-depressant drugs and memory, what is the group taking the drug called?

(a) placebo    (b) control    (c) experimental
Remember    Know    Familiar    Guess

When a test measures what it claims to measure, it is:

(a) reliable    (b) valid    (c) sensitive
Remember    Know    Familiar    Guess

Received March 12, 1996
Revision received September 10, 1996
Accepted February 19, 1997