# I'd Know that Face Anywhere!

by

Vincenza Gruppuso
M.Sc., University of Victoria, 1995
B.A., McMaster University, 1992
Dip.C.S., Institute of Child Study, University of Toronto, 1987
B.Sc., University of Toronto, 1985

A dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

We accept this dissertation as conforming
to the required standard

Dr. D. S. Lindsay, Supervisor (Department of Psychology)

Dr. M. E. J. Masson, Departmental Member (Department of Psychology)

Dr. H. Kadlec, Departmental Member (Department of Psychology)

Dr. L. Saxon, Outside Member (Department of Linguistics)

Dr. B. Whittlesea, External Examiner (Department of Psychology, Simon Fraser University)

Supervisor: Dr. D. Stephen Lindsay


## Abstract

The empirical studies reported used the remember/know paradigm to assess the effects of manipulating the number of exposures, delay, and context on the phenomenology of face recognition. In this paradigm, participants classified recognized faces according to the type of memorial awareness for prior occurrence. If recognition was based on the retrieval of episodic information such as context information, then participants indicated that they *remembered* the face. In contrast, if recognition was based on a feeling of familiarity without the recall of specifying information (i.e., an undifferentiated feeling of familiarity), then participants indicated that they *knew* the face.

Dual-process approaches to understanding remember and know states of awareness and the memorial processes (i.e., recollection (R) and familiarity (F)) that buttress them include those in which there exists (a) an exclusive relationship between processes (i.e., R processes underpin remember responses, and F processes underpin know responses); and, (b) an independent relationship between processes (i.e., remembering is a function of R, and, knowing is a function of F in the absence of R). In contrast, the single-process perspective explains response differences in terms of differences in trace strength of familiarity. Initial increases in strength of familiarity may be sufficient to recognize a face and to state that one knows it. If additional specifying information becomes available, familiarity for the face becomes stronger and a remember response is provided.

The model of recognition promoted in these studies includes aspects of the above approaches. The studies were designed to evaluate predictions following from a

functionalist account of recognition memory. This model of recognition memory is based on the notion of independent memory attributes. When retrieval of a particular piece of encoded memory information can fulfil the goal of a task (e.g., identify source), that particular attribute contributes to an estimate of R. If it fails to do so, but elicits a feeling of oldness, then the information contributes to an estimate of F. Thus, retrieved information can contribute to either R or F but never to both within a particular task. Across tasks, memory attributes are free to contribute to the same or different process. Thus, in the functional view, R and F are post hoc classifications. In addition, it also suggests that, in general, processes that contribute to R may not be qualitatively different from those that contribute to F.

In Experiments 1 and 2, delay between study and test was manipulated to test the prediction that retroactive interference would contribute to the disruption of integrated memory attributes. This type of memory information would likely contain target face information bound to context information (e.g., information about the study room). While the retrieval of face-plus-context information on an immediate test would contribute to an estimate of R, the retrieval of face-only information would contribute to an estimate of F.

Context was manipulated in Experiments 3 and 4. Each face was studied with a unique context photograph. At test, target faces in Experiment 3 were presented with either a studied or new context. Experiment 4 included an additional condition in which target faces were paired with switched contexts. In the studied context condition, memory processes that encoded face and context information would likely be re-enacted, promote a subjective experience of remembering, and, thus, contribute to an estimate of

R. In contrast, in the switched and new context conditions, retrieved information about the face only may contribute to a feeling of oldness and an estimate of F.

The results for know responding in Experiments 2, 3, and 4 provided support for the functional model. In Experiment 1, F for repeated items was unaffected by the delay, and in Experiment 2 it was reduced by the delay. The latter result suggested that there were more items on the immediate test that contributed to an estimate of F and were then forgotten on the delayed test than were items that initially contributed to an estimate of R and later contributed to an estimate of F. In the final two experiments, the change in context at test had no effect on estimates of F. While this latter result does not provide definitive support for the functional model of recognition memory it, as well as the reduction in estimates of R, does support the notion of independent processes.

Examiners:

Dr. D. S. Lindsay, Supervisor (Department of Psychology)

Dr. M. E. J. Masson, Departmental Member (Department of Psychology)

Dr. H. Kadlec, Departmental Member (Department of Psychology)

Dr. L. Saxon, Outside Member (Department of Linguistics)

Dr. B. Whittlesea, External Examiner (Department of Psychology, Simon Fraser University)

Table of Contents

## List of Tables

List of Figures

# Acknowledgments

I believe that "it takes a village to raise a child." Likewise, I also believe that the "village" was pivotal to the completion of this dissertation. It was only with the support of family and friends that I was able to persevere in my atypical dissertation journey. They kept insisting that I not only "should" complete this project, but that I was highly capable of doing so.

A special thank-you to Steve Lindsay for understanding my family needs, and for his highly valued, and generous, assistance and direction within, and outside of, the department of psychology.

To my committee members, Helena Kadlec, Michael Masson, and Leslie Saxon, for their unfaltering patience and guidance in completing the dissertation.

To Valerie Gonzales for listening and for inspiring me to push my often self-imposed limits.

To Tanya Berry, David Sulz, Laurie-Anne Keith, Ingrid Friesen, Ann Auld, Debbie Ball, Deborah Douglas, Enid Elliott, Lucille MacKay, Jack Lalonde, Joanne Fraser, Bev Kissinger, Glen Bodner, Anna-Lisa Cohen, Michelle Arnold, Cindy Bukach, Leora Dahl, Iris van Rooij, Denise Broekman, Joshua Goldberg, and many, many dear friends for their diverse and indispensable roles in making the dissertation a reality.

To my families in Toronto, Ontario and Goldsboro, North Carolina for their infinite support and encouragement.

To my husband, Tom Goldbach, who unquestioningly and lovingly took the journey with me and without which I would not have completed.

To my son, Marc Goldbach, who is my greatest teacher.

# Dedication

To Marc for his courage and tenacity in rising above life's challenges. And, without a doubt, for just being such a cool kid.

I'd Know that Face Anywhere!

Introduction

The subjective assessment of memorial information is central to the recognition of an item or event as one experienced in the past. The products of memory sometime specify a previous episodic event and thus promote a state of awareness that differs from that which arises when the products of memory support only an undifferentiated feeling of oldness. Both memory products and the emergence of different states of awareness, or phenomenal experiences, are a result of cognitive processes. Hence, the study of experiential aspects of memory should advance knowledge associated with process models of recognition memory.

Much research on the phenomenology of recognition memory is based on Tulving's (1985) original remember/know (RK) paradigm. In this paradigm, recognition of an item that is accompanied by conscious recollection of information that explicitly supports previous experience with the item (e.g., contextual information) is additionally classified as a *remember* response. However, if recognition is based on a subjective feeling of oldness (i.e., familiarity) without any conscious recollection of specifying information, then it is simply classified as a *know* response. Mandler's (1980) classic description of encountering a known individual in an unusual context (e.g., seeing the butcher form your supermarket on the bus) illustrates the concept of familiarity in the absence of recollection. Most people have experienced the powerful (and nagging) feelings of familiarity of having met this person earlier when specifying or context information has not yet been retrieved. Faces may be particularly likely to produce feelings of knowing in the absence of recollection, in part because humans attend closely to faces (and there may be innate, Fantz, 1961, and specialized neuronal mechanisms for encoding them, Damasio, Tranel, & Damasio, 1990) and in part because they can be highly complex and distinctive stimuli. These characteristics would support the familiarity of previously seen faces, but they would not necessarily support recollection of the episodic details of the prior encounters. If so, compared to other sorts of stimuli, faces would be particularly likely to give rise to know responses (i.e., familiarity in the absence of recollection). Within the laboratory the subjective experience of recognizing previously encountered individuals can be studied by using faces as stimulus items in remember/know paradigms.

A search of the memory literature revealed many studies on issues such as what makes faces so memorable (e.g., distinctiveness of features) but few that studied the phenomenal experience of remembering faces or the processes (i.e., familiarity and recollection) that support this experience (Mäntylä, 1997; Mäntylä & Cornoldi, 2002; Parkin, Gardiner, & Rosser, 1995). The paucity of studies in this area seems odd because we rely on phenomenal experiences to guide everyday decisions about people we encounter. For example, we may refuse to approach an individual because the absence of a phenomenal experience of remembering has led us to classify the individual as a "stranger." However, if a phenomenal experience of remembering arises we may approach the "friend." My studies broadened the research by assessing the effects of manipulating the number of exposures, delay, and context on the phenomenology of face recognition. These variables were chosen because they enable tests of hypotheses regarding the relationship between different bases for recognition. In addition, they represent ecologically valid factors that likely affect face recognition in daily life.

The empirical studies reported were designed to evaluate predictions following from a functionalist account of recognition memory. A full review of this model is provided later in the introduction. Briefly, this model of recognition memory is based on the notion of independent memory attributes. The retrieval of particular pieces of encoded memory information may or may not allow completion of the goal of a task (e.g., identify source). When it can fulfil the goal of the task, the particular attribute contributes to an estimate of recollection (R) processes. If it fails to do so, it is not necessarily the case that the item will be rejected and classified as unstudied. Information that comes to mind during a memory test may give rise to an undifferentiated feeling of oldness, and this information therefore contributes to an estimate of familiarity (F) processes. Thus, a particular piece of information can contribute to either R or F but never to both within a particular task. Across tasks, memory attributes are free to contribute to the same or different process. Thus, in the functional view, R and F are post hoc classifications. In addition, it also suggests that processes that contribute to R are not necessarily qualitatively different from those processes that contribute to F.

In Experiments 1 and 2, delay between study and test was manipulated to test the prediction that retroactive interference would contribute to the disruption of integrated memory attributes. This type of memory information would likely contain target face information bound to context information (e.g., information about the study room, position of stimulus on monitor, etc.). While the retrieval of face-plus-context information on an immediate test would contribute to an estimate of R, the retrieval of face-only information would contribute to a feeling of oldness and, consequently, to an estimate of F. Such movement from R to F would support a functional view of recognition memory.

Context was specifically manipulated in Experiments 3 and 4. Each face was studied with a unique context photograph. At test, target faces in Experiment 3 were presented with either a studied or new context. Experiment 4 included an additional condition in which target faces were paired with switched contexts. In the studied context condition, memory processes that encoded face and context information would likely be re-enacted, promote a subjective experience of remembering, and, thus, contribute to an estimate of R. In contrast, in the switched and new context conditions, retrieved information about the face only may contribute to a feeling of oldness and an estimate of F. As with the delay manipulation, the contribution to either R or F depending upon the presence or absence of the study context would support the functional account of recognition memory.

*Dual-Process Theory of Recognition Memory*

Dual-process theories of memory propose that recognition memory task performance is a product of two independent cognitive processes (Jacoby & Dallas, 1981; Mandler, 1980): recollection and familiarity. Recollection (R) is consciously controlled. It is a relatively slow process and is performed with intention, awareness, and effort. In contrast, familiarity (F) is automatic. It is a relatively fast process and proceeds without intention or awareness (Jacoby, 1991; Jacoby, Yonelinas, & Jennings, 1997). In addition, there is assumed to be a qualitative difference between processing that supports R and that which supports F.

Despite these differences, both processes can produce a subjective experience of "oldness." For recollection, the subjective experience of remembering is supported by the search and retrieval of specific episodic information. The memorial products of familiarity, in contrast, cannot be attributed to a particular study episode and hence an undifferentiated feeling of oldness is generated. This latter "sense of knowing," according to Mandler (1980), relies on processes that integrate intraevent attributes such as sensory and perceptual aspects of the episode. In addition, Jacoby (Jacoby & Dallas, 1981; Kelley & Jacoby, 1998) posits that this feeling of oldness comes as a result of fluent (i.e., easy or fast or both easy and fast) processing of the item that makes the item appear to "jump out."[1] Within the context of a recognition memory test, the participant is likely to attribute this fluency to the past and a feeling of oldness or recognition ensues. Mandler (1980) provided a compelling example of the two processes when he described what it was like to encounter a known individual in an unusual context. Initially, specifying or context information has not yet been retrieved and there is simply a nagging feeling of having met this person earlier (familiarity). This classic example of F in the absence of R may be a result of attributing fluent processing of perceptual attributes (e.g., the colour and shape of the person's eyes) to the past. If a search process results in the retrieval of

---

1

The sense of fluency associated with an item may arise because of an integration mechanism as described by Mandler (1980) but this is not the only mechanism that gives rise to fluency. For example, fluency may be induced by varying the physical characteristics of stimulus items that were not previously studied. Whittlesea, Jacoby, and Girard (1990) varied the clarity of a to-be-identified target word and showed that a target that was less occluded was easier to identify and more likely to be classified as a repeated item regardless of its actual repetition status.

specific context information, then the individual is identified (e.g., "He's the butcher from the supermarket!") (recollection).

The process-dissociation (PD) procedure developed by Jacoby (1991) serves as a methodological tool for deriving estimates of the contributions of these two independent processes to measures of memory such as recognition memory. Jacoby proposed that consciously controlled processes may be indexed by differences between responses when people are trying *to* versus trying *not to* do something. Thus, the PD procedure compares responses in a facilitation condition (typically called the inclusion condition), in which participants try to select a particular set of items, with responses in an interference condition (typically called the exclusion condition), in which participants try to select against a particular set of items. In the inclusion condition, responses may be based on controlled (i.e., R) or automatic (i.e., F) or both influences (i.e., $p$("Old") = R + F - RF). In the exclusion condition the extent to which the participant is successful in trying not to do something is a reflection of controlled responding. In contrast, selecting not-to-be-selected items in the exclusion condition arises when automatic processes are unopposed by controlled processes (i.e., $p$("Old") = F(1 - R) = F - RF). The difference between responding in the inclusion condition and responding in the exclusion condition provides an estimate of the contribution of R to recognition memory judgments (i.e., (R + F - RF) - (F - RF) = R). This estimate of R can be used to solve the inclusion or exclusion equation to obtain an estimate of the contribution of F processes to responding.

Evidence to support the independence between the two processes has been amassed through studies that have manipulated variables to produce theoretically sensible dissociations (Jacoby, 1998; Jacoby et al., 1997). For example, it is expected that dividing attention at test would likely impair the intentional retrieval of episodic information. Consequently, the extent to which R contributes to responding at test would be less in a divided attention condition than in a full attention condition. Because F is said to be automatic one would not expect divided attention to affect its contribution to recognition memory performance. This is exactly what Jacoby (1991) reported. A similar dissociation was obtained when attention was divided during study (Jacoby, Toth, & Yonelinas, 1993). Independence is also supported by the finding that for elderly participants R processes are less likely to contribute to responding than for young participants, whereas the contribution of F processes to responding does not differ between these two groups (Jennings & Jacoby, 1993).

Curran and Hintzman (1995) questioned the PD procedure's assumption of independence because they found an unexpected dissociation between memory processes in a word-stem completion task. This dissociation, termed paradoxical, involved changes in both controlled and automatic processes where only a change in the controlled component was anticipated. In support of their claim for non-independence was an analysis of item responses that showed strong positive correlations between estimates of controlled and automatic processes. Regardless of the relationship between processes (i.e., independent, redundant, or exclusive), controlled processes dominate responding. In those instances where both controlled and automatic processes co-exist, responding will be attributed to estimates of controlled processing. Consequently, estimates of automatic processing will tend to be underestimated and estimates of controlled processing will be greater by an amount equivalent to the extent of the underestimation and hence show a paradoxical dissociation (Jacoby, 1998).

Other researchers have indicated that the PD procedure does not account for response bias and its effects on parameter estimates (Buchner, Erdfelder, & Vaterrodt-Plunnecke, 1995; Graf & Komatsu, 1994; Roediger & McDermott, 1994). They note that a difference in the amount of guessing or the false-alarm rate between inclusion and exclusion tests, manipulations, or experimental conditions, strongly signals a change in the response criterion. Such changes may force derived estimates of R (i.e., controlled) or F (i.e., automatic) or both processes to be over- or under-estimated.

Jacoby and colleagues (Jacoby 1998; Jacoby, Begg, & Toth, 1997; Yonelinas, Regehr, & Jacoby, 1995) acknowledge that application of the PD procedure is not without limits. If direct retrieval instructions are not ensured and participants rely on a generate-recognize strategy (i.e., the relationship between processes becomes redundant), then estimates of automatic processes will be underestimated and produce paradoxical dissociations. If floor and ceiling effects are not avoided, then the response criterion across tasks will differ significantly, and use of the procedure will be invalidated (Jacoby, 1998).

Despite the absence of "signatures" (i.e., floor and ceiling effects; differing false-alarm rates; generate-recognize strategy rather than direct retrieval strategy) indicating inappropriate use of the PD procedure, Bodner, Masson, and Caldwell (2000) observed conflicting results on a stem-completion task when independence between controlled and automatic processes was assumed. They extended the procedure used by Toth, Reingold, and Jacoby (1994) in which participants were asked to either read items aloud or generate items from a phrase. Bodner et al.'s additional encoding condition asked participants to provide related or associated words to critical items that were read aloud. This encoding condition relied on both conceptually- and perceptually-driven processes.

The authors used an inclusion-exclusion paradigm to obtain estimates of controlled (i.e., recollection) and automatic (i.e., familiarity) processes. In the exclusion condition, participants were asked to complete stems (e.g., piz- - ) without using words from the study list. Reporting of study-list responses on the exclusion test was expected to occur if automatic influences of memory are unopposed by controlled or conscious uses of memory. Although they found no obvious evidence to suggest that independence was not valid (e.g., no difference in baseline rates) their estimate of automatic processes for the associate condition was significantly lower than that for the read condition although both conditions provided participants with the same perceptual information. The authors suggested that the influence of automatic processing in the associate condition was underestimated because of prior conceptual encoding that resulted in the unintended use of a generate-recognize strategy. In other words, participants automatically generated the word for the stem-completion test and then appropriately excluded this item once they recognized it from the study list (i.e., conceptual processing at study aided recognition of the item at test). Thus, the final estimate of the contribution of automatic processing on responding was reduced by the amount of successful exclusion of items.

Bodner et al. (2000) suggested that this underestimation of automatic processes invalidates the use of the PD procedure for this task and consequently the assumption of independence. The stem-completion task appears to better fit a model with redundancy between processes in which a completion is first automatically generated and then a conscious recognition process evaluates whether or not it was present in a previous study episode. Consequently, the PD procedure may not be the appropriate methodological tool for the stem-completion task (Curran & Hintzman, 1997) or cued-recall tasks in general.

*Dual-Process Signal-Detection Model*

In the model of recognition memory proposed by Yonelinas (1994), familiarity and recollection are independent processes governed by different principles. While recollection is a threshold process with an all-or-none criterion, familiarity is described by signal detection theory. Recognition performance can be based on the specific retrieval of information about a target episode (i.e., recollection). In addition, an assessment of the strength or familiarity of an item can also contribute to performance. Familiarity varies along a continuum of quantitative values ranging from very weak to very strong. Performance is influenced by values of familiarity that exceed some criterion.

In the signal detection model, non-studied items have base values of familiarity that are expected to be normally distributed. After study, this normal curve retains its shape but is shifted toward the very-strong end of the continuum of familiarity values. Thus, both non-studied and studied curves have equal variance. The criterion set by the individual typically maximizes recognition of studied items and minimizes "recognition" of non-studied items (see Figure 1). This strategy is also influenced by item, study, or both item and study, manipulations that influence the degree to which the non-studied and studied curves are separated.

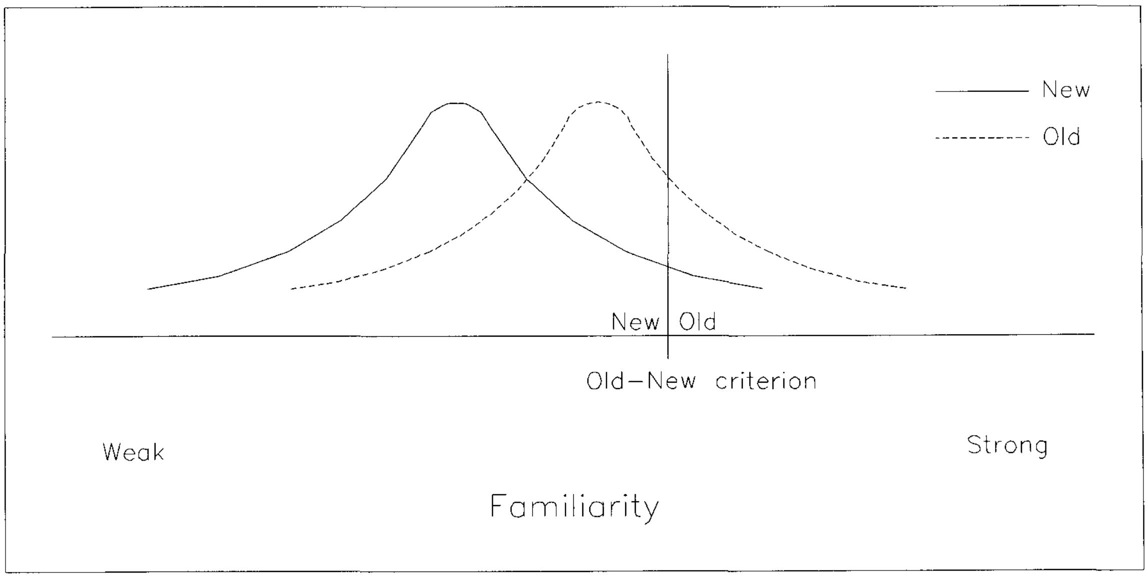


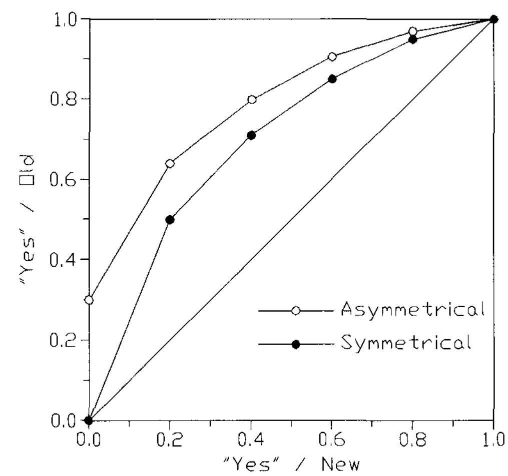

*Figure 1.* Distribution of familiarity of new (non-studied) and old (studied) items.

In a typical recognition experiment that uses signal detection theory, participants are asked to rate the confidence with which they recognize a particular item (e.g., very confident of non-recognition, mostly confident of non-recognition, slightly confident of non-recognition, slightly confident of recognition, mostly confident of recognition, very confident of recognition). The cumulative probabilities of correct recognitions (i.e., hits on the y-axis) are plotted against the cumulative probabilities of incorrect recognitions (i.e., false alarm, or "recognition"of a non-studied item on the x-axis) for each level of confidence beginning with the most confident judgments. This plot is the receiver-operating characteristic or ROC. When the ROC is a symmetrical curve along the positive diagonal, then recognition is considered to be based solely on familiarity. However, when the curve becomes asymmetrical an additional factor may be influencing responding. For most recognition memory experiments, resultant ROC curves are asymmetrical rather than symmetrical. If recollection is the additional factor, an increase in the probability of very confident hits is expected. Consequently, the ROC curve will become asymmetrical with the larger portion of the curve on the left side of the x-axis (see Figure 2a).

Yonelinas (1994) transformed each point on the ROC curve to a z-score. The resultant transformed curve plotted on z coordinates becomes linear. The intercept of this line (i.e., where z-score for false alarms equals zero) provides a measure of sensitivity, *d'* (i.e., a measure of the distance between the means of the non-studied and studied item distributions of familiarity). A symmetrical ROC curve produces a line with a slope of 1.0, and an asymmetrical curve produces a slope of less than 1.0 (see Figure 2b).

To test the signal-detection hypothesis, Yonelinas (1994) extended an earlier

(a)




(b)

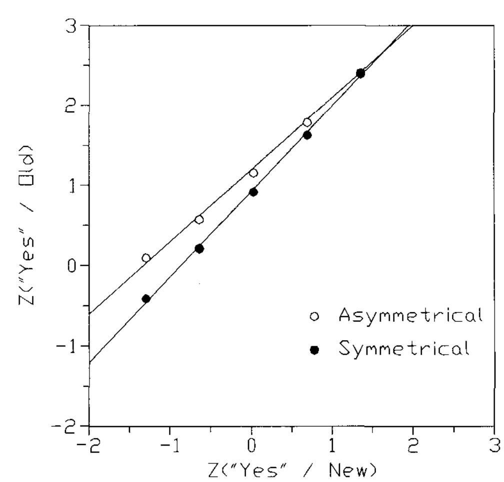



*Figure 2.* (a) Symmetrical and asymmetrical receiver-operating characteristics (ROC). Hypothesized criteria placements for a 6-levels-of-confidence experiment are included. The y-axis represents the probability of responding yes to a studied item, and the x-axis represents the probability of responding yes to a non-studied item. (b) Hypothetical plot of z-scores associated with symmetrical and asymmetrical ROCs. See text for details. These graphs are adapted from Yonelinas (1994, p. 1342).

paradigm that included the study of long and short lists of words (Yonelinas & Jacoby, 1994). In that experiment Yonelinas and Jacoby showed that recollection was less likely to contribute to responding in the recognition of long lists than in the recognition of short lists. In contrast, familiarity was not affected by list length. Yonelinas (1994) found that the slopes of the z-transformed ROC curves for both long and short lists were less than 1.0. This was to be expected because the earlier study showed that recollection contributed to the recognition of both list types. However, the slope for the long list was significantly larger than that for the short list. If there is a greater contribution of recollection for short lists than for long lists, then the skew observed in the ROC curve for shorts lists should be greater than the one produced for the long lists. In other words, the z-transformed line for short lists would have a reduced slope than the line for long lists. The opposite would be true for *d'*, and, in fact, *d'* for the long list was smaller than *d'* for the short list. This also implies that items in the short list were more easily identified among a list of studied and non-studied items than were items in the long list.

Evidence to support Yonelinas' dual-process signal-detection model comes from additional studies that looked at tasks or populations that relied almost exclusively on R or F. For example, Yonelinas (1997) compared item and associative learning (i.e., single words vs. word pairs) and found that the ROC for item recognition was curvilinear while that for associative recognition was linear. As indicated above, the curvilinear (and skewed) ROC for item learning was expected for a task that makes use of both R and F processes. In the associative recognition task, all words in test pairs were previously studied; some were tested in their original studied form while others were rearranged.

Thus, a non-specifying F process would be irrelevant to this task, and participants would be forced to rely on R processes (and hence the retrieval of specifying information) to decide whether or not a word pair at test was a pair presented at study. The linear ROC indicates that recollection-based responding was not affected by changes in false alarm rates (i.e., changes in criterion). This supports the hypothesis that recollection is a threshold process.

In a study that compared performance of individuals with amnesia and healthy controls, Yonelinas, Kroll, Dobbins, Lazzara, and Knight (1998) found symmetrical ROCs for the amnesics and asymmetrical ROCs for control participants. Slopes of the z-transformed ROCs (zROC) for the group with amnesia were not significantly different from 1.0, while slopes for the control groups were significantly less than 1.0. Estimates of recognition processes were derived by assuming independence between processes and, compared to healthy controls, individuals with amnesia showed deficits in both recollection and familiarity. However, impairment of recollection processes was greater than impairment of familiarity processes. Consequently, individuals with amnesia would more likely rely on familiarity processes when responding. This hypothesis was supported by the finding that the ROC was curvilinear and symmetrical, and the zROC was linear for these individuals.[2] In addition, this also supports the notion that familiarity can be

---

2

Yonelinas et al. (1998) noted that the poor performance of individuals with amnesia as compared to controls may have influenced the shape of the ROC curve. However, when performance of a second control group was manipulated to parallel that for amnesics, the shape of the ROC curve for controls continued to reveal asymmetry and the slope of the zROC slope was significantly less than 1.0.

described as a signal-detection process.

*Functionalist Account of Recognition Memory*

Using a two-list study session and Jacoby's (1991) PD procedure, Gruppuso, Lindsay, and Kelley (1997) found evidence to support an alternate view of memory processes and offered functional definitions for recollection and familiarity in recognition memory. At study, participants used a single encoding task (e.g., What is the monetary value of the item?) to study all first list items. For the second list, the encoding task varied between one of two tasks for each item (i.e., What is the monetary value of the item? or, How often have you seen this item in the past month?). One of the two tasks in the second study list was always the same as the task used in the first list. For the recognition memory test, all participants were instructed to include items from the first list and exclude items from the second list. Manipulation of similarity of the encoding task across the two lists yielded an estimate of the contribution of R processes to memory in the dissimilar condition that was greater than the estimate observed in the similar condition. In contrast, the estimate of F processes was greater in the similar condition than in the dissimilar condition.

Gruppuso et al.'s explanation of the inverse effect of similarity on these memory processes was based on the notion of independence of attributes encoded for the items during study. At test, some stimulus attributes or features may be retrieved independently of other attributes. Within a particular task (e.g., exclusion test) the retrieval of a particular attribute will contribute to the estimate of R or F but not both. Across tasks, however, this same attribute can contribute to the same parameter estimate (e.g., contribute to the estimate of R in two different tasks) or a different parameter estimate (e.g., contribute to the estimate of R in one task and the estimate of F in another task). The defining factor within and across tasks is the context of the task: If a particular piece of memory information provides a basis for excluding to-be-excluded information, then retrieval and use of that information will contribute to the estimate of R. If, in another context, the information does not provide a basis for exclusion, it will contribute to F. Thus, F is functionally defined as the retrieval of information about a studied item which enables one to recognize that item as old but does not enable one to answer the question set by the situation (e.g., identify the source of the item and hence correctly exclude to-be-excluded items). R is the retrieval of information that allows one to successfully complete the goal of the task. Additional support for Gruppuso et al.'s assumption of independence between memory attributes was obtained in the form of theoretically sensible dissociations between estimates of R and F for similar and dissimilar items as a function of full versus divided attention at test (Experiment 2) and at study (Experiment 3).

Other researchers have obtained comparable results. Within the source-monitoring framework, Dodson and Johnson (1996) showed that participants were more likely to make errors about the source of the test item when to-be-included and to-be-excluded items were similar than when they were dissimilar. Thus, participants were less likely to exclude to-be-excluded items in a similar condition than in a dissimilar condition. Mulligan and Hirshman (1997) described a diagnostic context model of recognition that makes a distinction between recollection of information that enables the participant to identify the source (i.e., diagnostic recollection) and recollection that does not permit identification of source (i.e., nondiagnostic recollection). Similarly, Yonelinas and Jacoby (1996) reported that participants can recall information that is irrelevant to the task (i.e., noncriterial recollection). Consequently, this *recollected* information cannot be used as a basis for exclusion and this type of recollection will contribute to an estimate of F rather than an estimate of R. These studies also support the view advanced by Gruppuso et al. (1997) that R and F are post hoc classifications that reflect, within a particular context, the retrieval of information that does or does not permit exclusion of target items. Furthermore, these results suggest that, within limits, the type of processing and the products thereof that contribute to estimates of R may not be qualitatively different from those which contribute to estimates of F (although in many contexts there may be such qualitative differences between the kinds of memory information that do versus do not enable exclusion).

*Memory Awareness*

The paradigms used in the above experiments provide a third-person or indirect account of processes that underlie the state of being aware that one is re-experiencing an event (i.e., processes that support recognition memory). For example, PD experimenters use the participant's positive and negative responses to different item types (e.g., to-be-selected versus not-to-be-selected items) and different formulae to infer the extent to which a particular process contributed to the participant's performance or remembering at test. In addition to third-person accounts, Gardiner (1988; Gardiner & Java, 1993) proposed that first-person or direct methods for measuring states of awareness be used to contribute to our knowledge of recognition memory.

The method advocated by Gardiner is based on Tulving's (1985) remember/know (RK) paradigm. In this paradigm, as in other recognition memory tests, participants indicate at test whether or not they recognize an item from the study episode. In addition, participants in the RK paradigm are instructed to provide a judgment of the type of awareness that supported each affirmative recognition decision. An item is said to be "remembered" if memory for the item is accompanied by conscious recollection of associative information such as information about what one was thinking at the time the item was initially presented, an image of the item as it appeared in the study list, or other association. For other recognized items, there is no accompanying conscious recollection of information that specifically associates the test item with the study session, but the participant has a subjective feeling of having encountered this item previously. When this occurs, participants indicate that they "know" that the item was studied; knowing is a feeling of familiarity in the absence of conscious recollection.

As defined above the relationship between remember and know responding must be exclusive; when participants recognize a test item they either consciously recollect accompanying information and respond "remember" or they do not consciously recollect any accompanying information and respond "know" (i.e., familiarity in the absence of conscious recollection). Gardiner and Parkin (1990) take this further and suggest that the relationship between the underlying mental processes (i.e., recollection and familiarity processes) or systems (i.e., episodic and semantic systems) that support these responses is exclusive. In general, recollection processes are identified with remember responses and familiarity processes are identified with know responses.

Studies using the RK procedure have observed different effects of variables on remember and know responding. For example, remember responding varied and know responding remained invariant when level of processing at study was manipulated (Gardiner, 1988). Participants were less likely to respond "remember" at test when they provided rhymes for each word on a study list than when they provided semantic associates. Gardiner and Parkin (1990) found that dividing attention at study reduced the probability of remember responding at test and had no effect on the probability of know responding. In contrast, in a study that addressed aging and remember/know responding, Parkin and Walter (1992) found that elderly participants provided more know responses than remember responses and that the pattern of responding among young participants was reversed (i.e., more remember responses than know responses). Similarly, using words and nonwords as stimulus items, Gardiner and Java (1990) found that words were more likely to be remembered than nonwords and that nonwords were more likely to be known than words. Gregg and Gardiner (1994) found that across conditions that did or did not maintain the same modality at test, the probability of know responding increased dramatically for items in the same-modality condition, whereas remember responding was not affected by the modality manipulation. Finally, in a study using memory for Polish melodies, Gardiner, Kaminska, Dixon, and Java (1996) found that increasing the number of study trials increased the probability of both remember and know responding.

Gardiner and colleagues (e.g., Gardiner, Kaminska, et al., 1996; Gregg & Gardiner, 1994) indicate that interpretation of results obtained with the RK procedure directly address only the effects of variables on phenomenal states of awareness. The type of relationship that may exist between processes that support responses does not direct analysis of the data. However, researchers can apply post hoc interpretations to support particular relationships (i.e., exclusive, independent, redundant). Gardiner favours an exclusive relationship (e.g., Gardiner & Parkin, 1990) between processes but admits that there exists the possibility that the relationship may be independent or redundant (Gardiner, Kaminska, et al., 1996). In contrast, Jacoby's (1991) model-based PD procedure necessarily assumes an independent relationship between processes (Richardson-Klavehn, Gardiner, & Java, 1996). Thus, in the PD procedure data are analyzed with the a priori assumption that the relationship between processes is independent.

*Independence Remember/Know (IRK)*

It is indisputable that at the level of the individual item there is an exclusive relationship between remember and know *responding* because the participant is allowed to choose only one of these two alternatives. As indicated above, this does not preclude an independent relationship between their underpinning processes (Jacoby et al., 1997). To understand how independent processes produce different states of awareness, consider the following scenarios. Independence implies that one of the following conditions exists for each item on a recognition test: 1. Conscious recollection is present and familiarity is absent; 2. Conscious recollection is absent and familiarity is present; 3. Both conscious recollection and familiarity are present; or 4. Both conscious recollection and familiarity are absent. In addition, Jacoby assumes that in the presence of conscious recollection, the participant will respond Remember regardless of the presence or absence of familiarity processes (i.e., remember responding dominates over know responding). Thus in conditions 1 and 3 the participant will provide remember responses. Know responses are a function of condition 2 (even though familiarity also occurs in condition 3). Recognition failure will occur in condition 4.

The descriptions provided in the above paragraph indicate that there are two conditions that include the presence of familiarity processes (i.e., conditions 2 and 3). Only condition 2 leads to a know response. In contrast, condition 3 leads to a remember response. Thus, according to Jacoby et al. (1997) simply using the probability of know responding (as proposed by Gardiner, 1988) to measure the influence of familiarity processes will tend to underestimate the true value. To offset this effect, Jacoby and his colleagues propose that the estimate of familiarity be assessed by using only the proportion of responses for which it was possible to provide a know response (i.e., $F = K / (1 - R)$). In other words, participants provide know responses when the following conditions co-occur: the presence of familiarity processes and the absence of recollection processes (i.e., condition 2: $K = F (1 - R)$). When this equation is solved for the estimate of familiarity we get the above equation (i.e., $F = K / ( 1 - R )$). Jacoby et al. (1997) refer to this method of estimating recollection and familiarity processes as the Independence Remember/Know (IRK) procedure.

Yonelinas and Jacoby (1995) investigated the nature of the relationship (i.e., exclusive or independent) between recollection and familiarity by studying the effect of size congruency on estimates of the two processes in each of the paradigms described above (i.e., the PD, RK, and IRK procedures). The authors expected to find that recollection and familiarity would be negatively affected by altering (i.e., increasing or decreasing) the size of studied objects (i.e., randomly produced line drawings of geometric shapes) on a subsequent recognition memory test. Thus, perceptual mismatch at test should attenuate the probability that recollection and familiarity processes contribute to responding. Estimates of these processes using the PD and IRK procedures were reduced by size incongruency. As expected, the RK procedure produced an estimate of recollection that was greater in the size congruent condition than in the size incongruent condition. However, the estimate of familiarity (as indexed by know responses in the RK procedure) was greater in the size incongruent condition than in the size congruent condition. It seems highly unlikely that a size-altered (i.e., perceptually mismatched) test object would be more familiar than a size-preserved (i.e., perceptually matched) one. Rather, Yonelinas and Jacoby argued that know responses increased for incongruent items merely because remember responses decreased, thereby increasing opportunities for familiarity to occur in the absence of recollection. These results more likely support the view that recollection and familiarity are independent rather than exclusive processes.

The results obtained with the RK procedure are not unique to the Yonelinas and Jacoby (1995) study. Rajaram (1996) used line drawings of familiar objects in her study (Experiment 2) and reported a value for know responding that was greater for items perceptually mismatched (i.e., size incongruent) than perceptually matched (i.e., size congruent) at test. As another example, Lindsay and Kelley (1996) found that reports of familiarity in the absence of recollection were greater for non-studied than for studied items. In addition, Jacoby et al. (1997) pointed to other inconsistencies obtained with the

RK procedure. As described earlier, Parkin and Walter (1992) found that their elderly participants provided more know responses than young participants. By the exclusivity assumption, this implies that for elderly participants familiarity processes are more likely, rather than just as likely, to contribute to responding. It is unclear why aging would lead to an increase in the availability or accessability of familiarity. Jennings and Jacoby (1993) used the PD procedure and found that elderly participants were just as likely as younger participants to rely on familiarity processes when responding. Jacoby et al. (1997) indicate that this result is similar to effects obtained by other researchers studying memory and aging.

It does seem reasonable that a decrease in the availability or accessibility of recollection processes would lead the elderly participant to more often base recognition of a studied item on the products of familiarity processes. Consequently, the elderly participant would provide more know than remember responses. The reverse is true for young participants (i.e., more remember than know responses) (Parkin & Walter, 1992). Recall that assuming independence implies that some remember responses are a product of both recollection and familiarity processes. Because young participants have a greater number of remember responses, the extent to which know responses underestimate the true value of familiarity would be greater for the young than for the elderly. Correcting for this underestimation would likely result in similar familiarity values for both young and elderly participants. Thus, independence can account for subjective (e.g., those obtained by a first-person RK procedure) and objective (e.g., those obtained by a third-person PD procedure) results.

Researchers have commented on the possibility of a similarity between know responding in the IRK procedure and exclusion errors in the PD procedure (Jacoby, Jones, & Dolan, 1998; Richardson-Klavehn et al., 1996). As indicated above, Jacoby (Jacoby et al., 1997) described know responding as a function of familiarity in the absence of recollection and captured this relationship with the equation, $K = F(1 - R)$. Note also that false alarm responding in the exclusion condition of the PD procedure occurs when familiarity processes are unopposed by recollection processes (i.e., $E = F(1 - R)$). Because of the similarity between the two equations, the effect of manipulations on know responding or exclusion performance should be similar. Thus similar outcomes from different approaches that both assume independence would support the view that the processes are independent. However, Richardson-Klavehn et al. (1996) pointed to experiments by Gardiner (1988) and Jacoby (1991) that in their view do not support the prediction that know responding and exclusion performance should be similar. Gardiner showed that the type of encoding task (i.e., generating or reading words at study) did not affect the proportion of know responses. A review of Jacoby's (1991) Experiment 2 in which participants in a full attention condition (i.e., a condition similar to that provided by Gardiner, 1988) were instructed to exclude both types of words, found the false alarm value (i.e., an exclusion value) for generate words to be less than that for read words (i.e., .30 versus .35). However, this difference was not significant, $t(40) = 1.21$. In Experiment 3, this condition was replicated and generate words were "slightly less likely to be called old than were read words (.29 versus .37), $t(40) = 1.79, p < .08$" (p. 528). Although the latter result is weak, the results of Jacoby's Experiments 2 and 3 parallel Gardiner's results (i.e., generating or reading words at study does not affect know responding and does not affect exclusion responses). This contradicts Richardson-Klavehn et al.'s (1996) assertion that these experiments do not support the notion that findings from an exclusion condition should parallel those from know responses (p. 137).

Jacoby et al. (1998) specifically addressed the predicted similarity between know and exclusion responding in a word repetition experiment. The authors found that number of repetitions (i.e., 1, 2, or 3 repetitions) did not affect the proportion of know responses. Hence, according to an exclusivity assumption there was no effect of repetition on familiarity. In contrast, Jacoby et al. found a positive effect of repetition on estimates of familiarity when derived with the IRK equation. When they varied the experiment to accommodate the inclusion/exclusion paradigm (i.e., participants read one list of words and heard another list), the proportion of items that participants failed to exclude was the same for all three repetition conditions. This experiment showed parallel effects on know and exclusion responding. Because effects on know responding have been shown to parallel those on exclusion responding and exclusion responding is defined as a function of independent familiarity and recollection processes (i.e., $E = F(1 - R)$), there exists some evidence to support Jacoby's notion that know responding is also a function of independent familiarity and recollection processes (i.e., $K = F(1 - R)$).

*Single-Process Model of Recognition Memory*

According to Donaldson (1996), data from the remember-know paradigm can be accounted for in terms of differences in trace strength of a unitary memory process. He, as well as other researchers (e.g., Hirshman & Master, 1997; Inoue & Bellezza, 1998), proposed a single-process model of remember and know responding that in their view is more parsimonious than the dual-process model. The single process is best described with signal-detection theory and produces a continuum of values of an item's memorial strength (e.g., information that can be used to assess an item's oldness). At one end of the continuum items are very weak (e.g., very little information) and at the other they are very strong (e.g., much information). In addition, this model includes two decision criteria. The first criterion is that for yes-no recognition. Consider an experiment that consists of two distributions of items: old and new. The strength of new items should be weaker than that of old items. If the strength dimension is imagined to increase along a horizontal line (that grows from left to right) then the distribution of new items will be displaced to the left of old items (see Figure 3). At test, strength values that fall to the right of an old-new criterion result in an old decision and those values to the left of the criterion result in a new decision. The second criterion is the remember criterion and is placed to the right of the old-new criterion. Items with strength values to the right of the remember criterion are judged to be remembered. Those items with values that fall to the left of the second criterion and to the right of the first criterion are judged as known.

The following examples from Hirshman and Master (1997) will help illustrate how a single-factor model can account for some dissociations between remember and know responding. The graphs in Figure 4 are adapted from Hirshman and Master (p. 348). In the top panel, the old-new criterion and the remember criterion are placed so that to the right of the remember criterion there is more area under the Condition 2 distribution than under the Condition 1 distribution. Hence there are more remember responses in

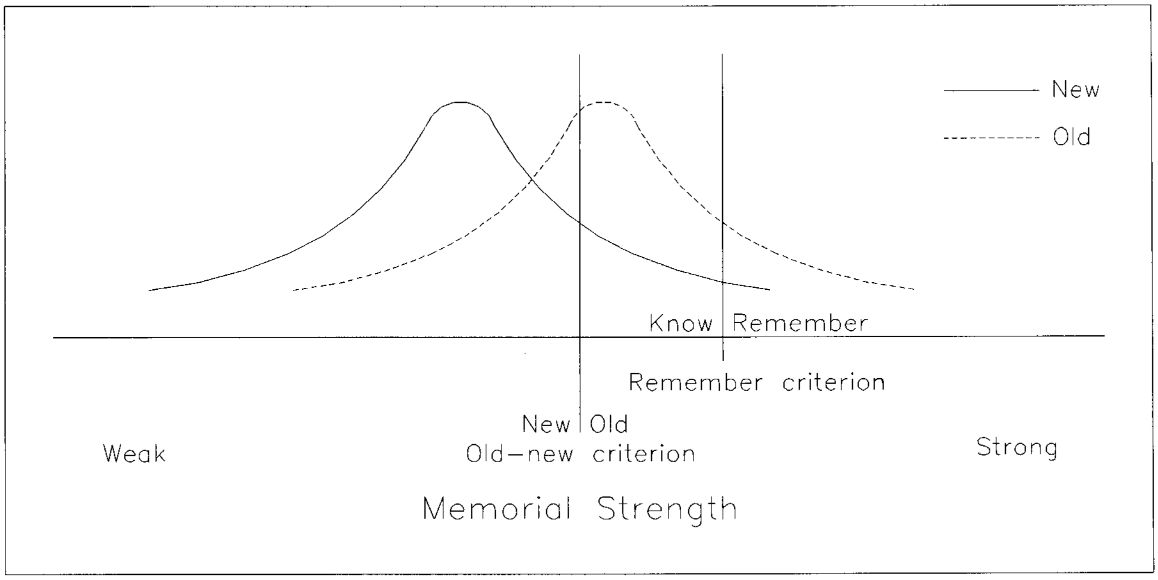



*Figure 3.* Diagram representing the distribution of new and old items as a function of memorial strength. The two response criteria define old-new responses and remember-know responses. See text for details.

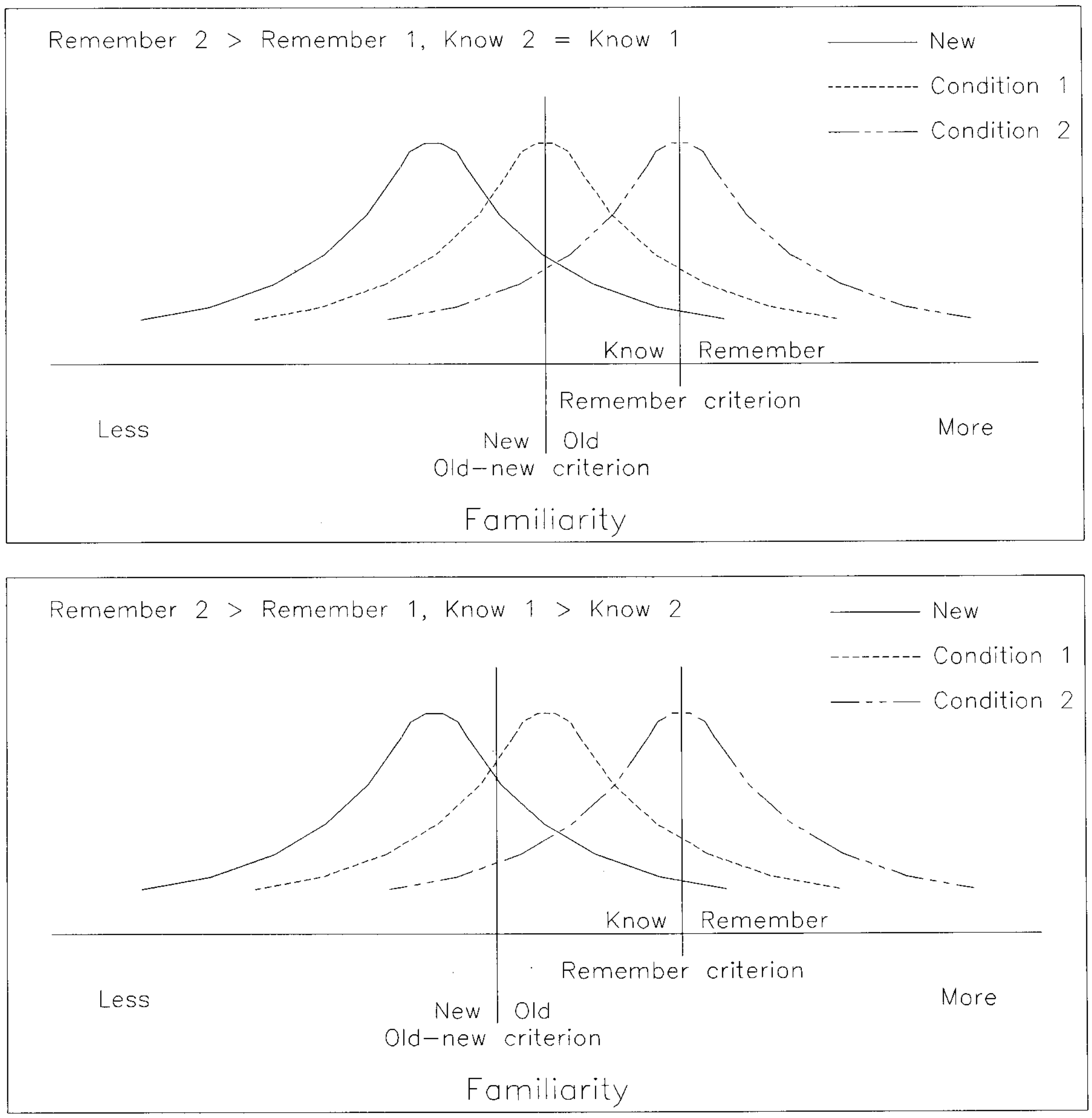



*Figure 4.* Distribution of different sets of items and placement of criteria. The second panel shows a shift of the old-new criterion (C). This shift to the left implies more liberal responding. Changes in criteria in these single-process models can explain some remember-know response patterns. See text for details.

Condition 2 than in Condition 1. The area between the two criteria corresponds to know responses and shows equal areas under the Condition 2 and Condition 1 distributions (i.e., equal know responding in both conditions). This type of result is similar to Gardiner's (1988) finding in which participants were more likely to provide remember responses to items in a semantic condition (i.e., Condition 2) than to items in a rhyme condition (i.e., Condition 1), but the probability of know responding did not differ between conditions. In the bottom panel of Figure 4, the remember criterion remains in the same position as that in the top panel but the old-new criterion is displaced to the left. This shift in criterion produces a greater area under Condition 1 than under Condition 2. This shift implies that know responding (identified with the area between criteria) is greater for items in Condition 1 than in Condition 2. This is what Gardiner and Java (1990) found when they used words (i.e., Condition 2) and nonwords (i.e., Condition 1) as stimuli.

Despite the fit between the single-process model and empirical evidence described above, Jacoby et al. (1998) provided indirect support for the view that the RK procedure reflects the influence of dual processes rather than a single process. They found that as the number of study presentations increased, response deadlines (i.e., in one condition responses were accepted only before a deadline and in another condition responses were accepted only after a delay) in an exclusion test produced opposite effects on the number of exclusion errors (i.e., the deadline condition produced an increase in exclusion errors and the delay condition produced a decrease in exclusion errors). Because the authors previously showed that effects on exclusion errors in the PD procedure paralleled the effects on know responses in the RK procedure, they reasoned that the response deadline manipulation results should be similar for know responses. A single-process model, according to Jacoby et al. (1998), could not account for these hypothesized opposite effects on know responding either by changing the distance between the means of the distributions or by changing the response criteria or both.

Gardiner, Richardson-Klavehn, and Ramponi (1998) argued that a single-process model is too simple and could not account for the presence of qualitatively different phenomenal states of awareness.[3] They also indicated that support for a dual process view of memory is provided by studies in other domains such as physiological psychology. For example, Duzel, Yonelinas, Mangun, Heinze, and Tulving (1997), in an effort to map states of awareness onto neural activity, found that event-related potentials (ERPs) for remember responses differed significantly from ERPs for know responses.

Although Donaldson (1996) has provided evidence to support a single-process view of remember/know responding, he also reported findings that invalidate this view. Specifically, he revealed in his meta-analysis of remember/know studies and in a follow-up study, that measures of memory associated with know responding (i.e., *A*' for know responding) were correlated with criterion measures (i.e., $B''_D$). If the single-process view is based on a model of memory that uses signal-detection theory, such correlations between sensitivity and bias violate a basic premise of the theory: independence between these two measures (Macmillan & Creelman, 1991). In the presence of a correlation, it

---

3

Donaldson (1996) also acknowledged that the signal detection model does not capture the phenomenal distinction between remember and know responding. However, he also argued that the "introspective [RK procedure] fails to capture the distinction" (p. 532).

would be possible to observe a greater number of know responses to new items than to old items (as in Lindsay & Kelley, 1996) when the response criterion (i.e., yes-no criterion) became very liberal (i.e., moved to the left). Consequently, the use of know responses as a basis for measuring memory would produce the odd result that memory was greater for new items than for old items. When responding was modeled upon know responding in the absence of remember responding as advocated by Jacoby (i.e., K/(1 - R), Jacoby et al., 1997), Donaldson (1996) noted that measures of memory were not correlated with response criteria. Thus, it is possible that researchers have succeeded in describing many remember/know results with a single process because the influence of a second, additional process (i.e., recollection) was minimal.

*Levels-of-Processing and Repetition Effects on Memory Processes*

Research shows that level-of-processing manipulations do not always produce the same effects on estimates of familiarity across different experiments. Gardiner (1988) found that semantic versus non-semantic encoding conditions and generate versus read encoding conditions had no effect on the probability of know responses. In addition, in the generate/read conditions this was true regardless of the time of testing (i.e., after one hour or after one week). Gregg and Gardiner (1994) also reported that know responding did not differ between conditions that varied level of processing at study (e.g., "Is the word associated with a 4-footed animal?" vs. "Is the word clearly visible?"). Such findings suggest that knowing or familiarity is unaffected by level of processing at study. In contrast, in a study reported by Rajaram (1993) semantically encoded items were less likely to be known than non-semantically encoded items. Jacoby's (1991) Experiment 3 (which used the process dissociation procedure, rather than remember/know judgments) produced estimates of familiarity that were greater for words studied as anagrams than for words simply read at study.

The effect of repetition on knowing or familiarity has also been inconsistent across studies. In a study of the effect of repetition on recognition of music, Gardiner, Kaminska, et al. (1996) found that participants were just as likely to know pieces of classical music that were repeated once at study as they were to know pieces of classical music that were repeated three times at study. In contrast, for unfamiliar music such as Polish music, participants provided fewer know responses for once-studied pieces than for three-time-studied pieces. As in the classical music study reported above, Jacoby et al. (1998) found that the proportion of know responses to words repeated three times at study was equal to that obtained for words repeated once. However, when estimates of familiarity were calculated with the independence remember/know formula (i.e., $F = K/(1 - R)$), these authors found that the contribution of familiarity processes to responding was greater for words in the three-time study condition than in the one-time study condition.

Jacoby et al. (1997) suggested that inconsistencies (i.e., in effects of levels-of-processing or repetition manipulations) obtained with the RK procedure are an artifact of the exclusion assumption: Large effects on remember responding constrain the effects on know responding and any corresponding effects may be missed (i.e., a Type II error) (cf. Gardiner, Java, & Richardson-Klavehn, 1996). Thus, one purpose of the following experiments was to assess the effect of a processing variable such as repetition on estimates of memory processes.

Given that familiarity is based on prior experience, it seems logical to expect the probability of familiarity to increase with number of exposures. Support for this claim can be found in earlier research on repetition priming effects in indirect tests of memory (e.g., word identification, lexical decision).[4] These tests primarily reflect the influence of implicit, automatic, or unconscious forms of memory (Kelley & Lindsay, 1996; Masson, 1989; Richardson-Klavehn & Bjork, 1988; Roediger & McDermott, 1993; Schacter, 1987; Toth, 1996; Toth, 2000).[5] If familiarity, as Mandler (1980) suggested, is based on intraitem integration, then repetition will be important for this process insofar as it provides for the opportunity to organize "elements of the target event" (p. 255). Thus, in

---

4

The term "indirect" is used in the sense that although the participant may become aware (i.e., either because the instructions indicated so, or he or she naturally noticed) that some of the words were presented in an earlier phase of the experiment, such awareness is not important to completion of the task (Richardson-Klavehn et al., 1996).

5

The notion that indirect tests of memory (e.g., word identification, lexical decision, word-stem completion, word-fragment completion) rely on unconscious processes and direct tests of memory (e.g., recall, recognition) rely on conscious processes (i.e., assume process-purity of tasks) may be difficult to uphold (e.g., Dunn & Kirsner, 1989; Jacoby, 1991). The impetus behind the development of the process-dissociation procedure (Jacoby, 1991; Jacoby et al., 1993) was this awareness of the possible influence of both types of processes on a particular task. Instead of assuming that tasks were process-pure, the process-dissociation procedure acknowledged the potential influence of both processes and used an opposition procedure (i.e., exclusion task; the influence of one process opposes the influence of the other process) in addition to a facilitation procedure (i.e., inclusion task; reliance on either a conscious process or unconscious process or both processes leads to successful performance) to obtain estimates of the influence of either process on responding.

general, repetition would be expected to augment the influence of automatic forms of memory on responding. For example, under the pretense that the researchers were interested in speed of reading, Jacoby and Dallas (1981) in their Experiment 5 had participants study words presented either once or twice. At test, participants were asked to identify studied and non-studied words briefly flashed on a screen (i.e., 35 ms.). Their results showed that the probability of identification was greater for repeated words than for non-repeated words. Similarly, in a lexical decision task (i.e., participants indicate whether or not a string of letters forms a word), Masson and Freedman (1990; Experiment 1) reported that participants were faster at identifying repeated words than non-repeated words. These results suggest that knowing (or, more properly, the memory mechanisms that underlie knowing) should increase with repetition.

*Memory Processes over Time*

During recognition memory study sessions, participants are provided with the opportunity to process and encode a variety of attributes for each item. Some of the cognitive processes performed when studying words on a list include noting the shape of the word, identifying the letters in the word, noting the meaning of the word, and thinking about when the word was last seen. In addition, the role of some processes is to bind different pieces of information (e.g., semantic information may become bound to information about the last non-experimental encounter) (e.g., Johnson & Chalfonte, 1994). These binding operations will describe a target item more wholly, uniquely, and distinctively. At test, retrieval conditions that generate cognitive processes that mimic processes performed during study will tend to provide the participant with a feeling that he or she is re-experiencing a past episode (as indicated by the encoding specificity principle described by Tulving & Thomson, 1973, the concept of transfer appropriate processing described by Morris, Bransford, & Franks, 1977, and the fluency hypothesis described by Jacoby & Dallas, 1981, and Kelley & Jacoby, 1998). Furthermore, enacting study operations at test will make it likely that encoded information is retrieved and hence provide a basis for a remember response.

The introduction of a delay between study and test makes it possible for other memory operations to interfere with retrieval of processes performed within the study session (i.e., retroactive interference). With increases in delay, re-enactment of encoding processes for individual attributes (e.g., shape information or letter information) will become less likely and, importantly, so will the re-enactment of those processes that bound or integrated individual attributes (e.g., bound target-context information). For the RK paradigm, a delay will make it less likely that source-specifying episodic information (e.g., bound semantic-source information) will be retrieved and hence the participant will less likely provide remember responses. Retrieved, unbound information (e.g., semantic information alone) may provide a basis for a feeling of oldness (and hence contribute to an estimate of familiarity or to know responses). Other variables, such as dividing attention at study that interfere with the performance of binding operations at study should also interfere with the generation of bound information needed for remembering at test.

Thus, it was predicted that in a paradigm that includes immediate and delayed test conditions some memory information that in an immediate test contributed to remember responding may in a delayed test contribute to know responding or the estimate of familiarity. This type of result would support the notion of independence among memory attributes and contradict both Jacoby's (e.g., Yonelinas & Jacoby, 1996) and Gardiner's (e.g., Gardiner & Java, 1991) notion of familiarity or know responding. Their views rest on the belief that the feeling of oldness associated with familiarity rests on qualitatively different processes than does recollection. For Jacoby such movement between remember and know responding would imply that the two recognition memory processes are not independent mechanisms. An exclusive relationship between these two constructs, as advocated by Gardiner, would also not predict the above movement between remember and know responding.

Gardiner and Java (1991) considered the possibility of such movement in their discussion of forgetting in recognition memory (p. 621). For the experiments reported, they measured remember and know responding after 10 minutes, 1 hour, 1 day, 1 week, 4 weeks, or 6 months, and found no evidence for an increase in know responding even after 6 months. Instead know responding exhibited a slow and steady decline after the immediate test that paralleled the decline in remember responding after the one-day interval.[6]

The authors briefly discussed (p. 621) an additional study that used a test-retest paradigm (tests separated by one week) that in their estimation was a sound basis for observing recollection becoming familiarity over time if such movement was possible. They failed to find a significant number of know responses on the second test that were

---

6

They found the rate of decay for remember responding to be steep between the immediate and the one-day tests, after which there was a slow decline.

initially remember responses on the first test. They dismissed the idea that over time recollection responses become know responses and further suggested that such movement would not support an exclusive relationship between two (separate and distinct) memory processes. Gardiner and Java did not provide any additional details for this particular study. However, if the participants were instructed to memorize the words for later tests of memory (i.e., intentional instructions) as in their Experiments 1 and 2, then participants may have used the first test as an opportunity to assess their memory for the words and strengthen memory associations. Thus, it is possible that their paradigm could not permit any observation of movement between recollection and familiarity estimates.

The effects of a delay on estimates of recollection and familiarity were also assessed in the following experiments. When the hypothesized relationship between the two processes and among attributes is independent, evidence of "downward migration" was expected: Some information that initially supported remember responding would later contribute to know responding.

*Remembering and Knowing Faces*

As mentioned earlier, Mandler (1980) illustrated the concept of familiarity in the absence of recollection with the everyday example of encountering a known individual in an unusual context (e.g., seeing the butcher from your supermarket on the bus). The example is compelling because most people have experienced the powerful (and nagging) feelings of familiarity that often arise in such situations. Faces may be particularly likely to produce feelings of knowing in the absence of recollection, in part because humans attend closely to faces (and there may be innate, Fantz, 1961, and specialized neuronal mechanisms for encoding them, Damasio et al., 1990) and in part because they can be highly complex and distinctive stimuli. These characteristics would support the familiarity of previously seen faces, but they would not necessarily support recollection of the episodic details of the prior encounters. If so, compared to other sorts of stimuli, faces would be particularly likely to give rise to know responses (i.e., familiarity in the absence of recollection). Within the laboratory the subjective experience of recognizing previously encountered individuals can be studied by using faces as stimulus items in remember/know paradigms.

To date, only three studies using the RK procedure reported the use of faces as stimulus items. Parkin et al. (1995) found that dividing attention at study reduced the probability of a remember response but had no reliable effect on the probability of a know response. According to the authors, dividing attention affects effortful encoding important for recollection processes but not familiarity processes. In their second experiment, they reported that participants were more likely to respond remember to repeated faces that were spaced throughout a study list than they were to faces that were massed within the list. The opposite was true for know responses (i.e., more know responding to massed repetition items than to spaced repetition items). The authors suggested that the spaced repetition condition provided the participant with the opportunity for elaborative and effortful encoding and increased recollection whereas massed repetition may have only promoted additional maintenance rehearsal and increased familiarity. They noted that these results are similar to data obtained with word stimuli, and thus they "increase the ecological validity of the experiential [RK] approach and the generality of its database" (p.

393).

Parkin et al.'s (1995) findings in Experiment 2 were replicated by Mäntylä and Cornoldi's (2002) first experiment: There were more remember responses in the spaced repetition condition than in the massed condition, and the opposite pattern of results for know responses (i.e., more know responses in the massed condition than in the spaced condition). Mäntylä and Cornoldi proposed that increased remember responding (and hence recollection processes) for spaced repetition was a result of elaborative encoding that specifically increased the distinctiveness of the face. Because they were primarily interested in showing that distinctiveness affected recollection processes, the effect on know responding (or familiarity) was not specifically addressed except to indicate that it was similar to other researchers' findings.

Mäntylä and Cornoldi (2002) extended the paradigm and included a perceptual change in their second experiment. The second presentation of the face was either an exact copy or a mirror image. For those participants whose study session included a mirror image, the face presented at test was either the original or mirror version. Participants were more likely to remember a face in the spaced condition rather than the massed condition regardless of any perceptual change present during study. According to the authors this result supported their view that any type of processing (i.e., conceptual or perceptual) that enhances the distinctiveness of an episodic event is important for recollection processes. Thus, different views of a stimulus item encourage encoding of distinct attributes that are important for recollection processes. This contradicts the notion that recollection relies on conceptual processing and familiarity relies on perceptual processing (e.g., Rajaram, 1993).

Unlike the findings of their first experiment, Mäntylä and Cornoldi (2002) found that know responding in their second experiment was not affected by the type of repetition when the second face was an exact copy of the first. In addition, perceptual change (whether tested with the mirror image or original) did not affect know responding (i.e., know responding to spaced condition items was equivalent to that for massed condition items). In the second experiment, the authors encouraged participants to be strict in choosing a yes or no response because they wanted to reduce the amount of guessing. The authors reasoned that this strict criterion succeeded in reducing the probability of guessing and the concomitant inflated know responding that would have followed. In other words, know responding in Experiment 1 was greater for spaced repetition items because it included guess responses that were encouraged by a lenient criterion. They further suggested that these results do not affect their interpretation of findings with remember responding which supported the distinctiveness hypothesis.

In an earlier experiment with faces, Mäntylä (1997) reported opposite effects on remember and know responding. Participants in his first experiment encoded two sets of photographs of faces. One set of items was rated on the distinctiveness of faces (i.e., how easy it was to pick out a particular face in a crowd). For the other set, participants were encouraged to look at the similarities among faces by sorting them into 4 different personality-type categories. Items rated for distinctiveness were more likely to be remembered than items rated for similarity. In contrast, know responses were more often given to items rated for similarity than for items rated for distinctiveness. According to

Mäntylä, these results indicated that study tasks that emphasized the encoding of distinct attributes of stimulus items increased the contribution of recollection processing to responding at test. In contrast, those tasks that emphasized similarities among items increased the contribution of familiarity processing to responding at test.

The above studies with faces and the remember/know paradigm (Mäntylä, 1997; Mäntylä & Cornoldi, 2002; Parkin et al., 1995) have demonstrated that dividing attention reduces the contribution of recollection processes to responding and that spaced repetition and distinctiveness encoding (as opposed to relational encoding) increase recollection processes. The effects of these variables were compatible with explanations that elaborative processing and the product(s) of this processing (i.e., improved distinctiveness) are important for recollection processes. In other words, tasks that encouraged processing of memory attributes that specified source tended to contribute to estimates of recollection.

The effects of these variables on know responding were less consistent. In Mäntylä's (1997) Experiment 1, relational encoding produced more know responding than distinctiveness encoding. Dividing attention had no effect on know responding in Parkin et al.'s (1995) first experiment. There were more know responses in the massed repetition condition than in the spaced repetition condition in Mäntylä and Cornoldi's Experiment 1 (2002) and in Parkin et al.'s Experiment 2 (1995). However, in Mäntylä and Cornoldi's second experiment which included a replication of their first experiment there was no effect of type of repetition on know responding. Because these experiments used know responding as a direct index of familiarity they may have under- or over-estimated the true contribution of familiarity to responding depending upon the demands of the task (e.g., restriction of know responding because participants "remember" most items) (e.g., Jacoby et al., 1997).

*Current Studies*

The above experiments were based on a dual-process view of recognition memory in which the relationship between processes was assumed to be exclusive. In addition to using remember response probabilities as a direct estimate of recollection processes, know response probabilities were used as a direct estimate of familiarity processes. The following experiments have in common the goal of testing recognition memory theory. Specifically, they were used to test the notion of independent, dual memory processes (i.e., recollection and familiarity), and the idea that these two memory experiences arise via multiple, independent attributes in particular contexts. Thus, recollection processes were expected to contribute to responding with or without the presence of familiarity processes and familiarity processes were expected to contribute to responding with or without the presence of recollection processes. In addition, memory for a particular attribute may contribute to remember responses (recollection) in one context but to know responses (familiarity) in another (and vice versa).

In my view, memory relies on the retrieval of a myriad of independent attributes, and the response difference in the RK paradigm relates to the extent to which the retrieved information uniquely specifies a particular source of a past experience. If retrieved information identifies a previous episodic event (e.g., the study phase of an experiment), then the participant provides a Remember response. In contrast, when retrieved information does not identify a previous episodic event but indicates a past experience, then the participant provides a Know response. For example, if the retrieved information specifies where you last saw an individual (e.g., context information such as the counter at a butcher shop) then you indicate that you remember the person. However, retrieved information may at times only provide a nagging feeling of having encountered someone previously because one is unable to link this information with a previous occurrence. Thus, you may recall that you had encountered this person after work or that it was on a particularly cold day and you can confidently state, "I'd know this face anywhere!" However, you don't have the information (i.e., the ideal context information) that allows you to "remember" or uniquely identify the person and you hence just "know" the person.

The experiments used face stimuli and the remember/know paradigm to derive knowledge about processes that give rise to phenomenal experiences of remembering. This paradigm measures first-person, subjective reports and derived estimates of the cognitive processes believed to underlie those experiences (i.e., recollection and familiarity). Specifically, the experiments evaluated the effects of repetition, delay, and context variations on reports of remember and know and on estimates of recollection and familiarity on face recognition.

Memory for a face repeatedly presented at study will more likely include a greater, and perhaps more varied, number of associations than memory for a non-repeated face. I predicted that participants would more likely say that they remember repeated faces than non-repeated faces, and that this would restrict know responses. Thus, participants would more likely know non-repeated faces than repeated faces. The effect of delay (and hence interference) would be to reduce the estimate of recollection with some prior bases for recollection to become bases for familiarity (i.e., R-to-F). In addition, interference was also expected to reduce some bases for familiarity to below threshold levels (i.e., F-to-New). Hence, if bases for recollection can truly become bases for familiarity (i.e., R-to-F), and if that phenomenon more than offsets F-to-New, then familiarity should increase with delay.

The experience of encountering known individuals in unusual versus usual contexts can be simulated in the laboratory by presenting photographs of faces in background contexts that may or may not be present at test. In a studied-context condition, processing of face and context information at test would match that performed at study, and the retrieved information would support remembering. In a new-context condition the retrieval cues do not necessarily elicit the encoding operations that would have, for example, bound the portrait with a particular context. Retrieved information about the face may only support a feeling of oldness. Without the retrieval of the appropriate context information the likelihood of a remember response is diminished and face information would contribute to know responding and, ultimately, an estimate of familiarity. Thus, face information in the studied-context condition was expected to contribute to recollection processes. In contrast, the same information in a new-context condition would more likely elicit a know response and contribute to an estimate of familiarity. Hence context manipulations also provide an opportunity to observe movement between remember and know responding, and to assess the notion of independent attributes and independent processes.

## General Method

*Dependent Variables and Derivation of Estimates of Processes.* The dependent measures (for all experiments) included the probabilities of responding remember or know to critical test items. Parameter estimates (i.e., recollection, familiarity, *d'*, *A'*, and $B''_D$[7]) were derived from these probabilities. Researchers differ in their opinion as to which method of correction provides more meaningful results. Thus, data were analyzed according to different methods of correcting biased responding and different theories of recognition memory. The probability of remember responding was used as the estimate of recollection. For the exclusive dual-process model of recognition memory (e.g., Gardiner & Parkin, 1990), know probabilities were used to estimate familiarity. Remember and know probabilities were also used to derive estimates of familiarity according to Jacoby et al.'s (1997) independence remember/know (IRK) equation.

One method of correcting remember and know hit rates used Gardiner and Java's (1991) approach of subtracting false alarm from hit rates (a 2-equal-threshold method, Yonelinas et al., 1996). Another method was suggested by Yonelinas et al. (1998). According to these authors, one cannot truly remember a new, non-studied item, and the presence of these erroneous responses reduces the opportunity to make valid remember responses. Thus, they proposed that remember responses be adjusted with a high-

---

7

*d'* and *A'* provide an index of discrimination or sensitivity; a participant's ability to discriminate between old and new items. $B''_D$ provides an index of the participant's tendency (bias) to provide one response over another. Both *A'* and $B''_D$ are distribution-free indices whereas *d'* relies on a Gaussian or normal distribution of some sensory event such as familiarity (Green & Swets, 1974; Macmillan & Creelman, 1991; Snodgrass & Corwin, 1988).

threshold correction that conditionalizes corrected hit rates on the opportunity to make a remember response (i.e., [Remember $_{old}$ - Remember $_{new}$]/[1 - Remember $_{new}$]). Negative values obtained with either of the above corrections were replaced with zeroes.

Corrected IRK estimates of familiarity were obtained by subtracting the IRK value for new items from that for old items. Yonelinas et al.'s (1996) method of assessing familiarity by analyzing *d'* values (i.e., sensitivity) was also implemented. The estimates of familiarity derived for old and new items (Yonelinas et al., 1998) were used to obtain a *d'* value for each participant (i.e., using a *d'* look-up table, Elliott, 1964). The values .5/N and 1 - .5/N were substituted for floor and ceiling familiarity estimates (i.e., hit rate values of 0 or 1), respectively (Kadlec, 1999).

Finally, the data were transformed according to Donaldson's (1992, 1996) single-factor model. In contrast to *d'*, Donaldson (1992) uses a distribution-free, non-parametric measure of sensitivity, *A'* (i.e., *A'* = .5 + [(H - FA)(1 + H - FA)]/[(4H(1 - FA)], where H = probability of a hit, and FA = probability of a false alarm) (see Macmillan & Creelman, 1996, for an alternate view). *A'* values range from 0 to +1, and an *A'* value of .50 represents chance performance. Donaldson has also derived a non-parametric measure of response bias, $B''_D$ (i.e., $B''_D$ = [(1 - H)(1-FA) - HFA]/[(1 - H)(1 - FA) + HFA]).[8] $B''_D$ values range from -1 to +1. Negative values represent liberal responding and positive values represent conservative responding.

---

8

For below chance (BCh) performance (i.e., false alarms greater than hits) sensitivity and bias were derived with equations provided by Aaronson and Watts (1987): $A'_{BCh}$ = .5 - [((FA - H)(1 + FA - H)]/[4FA(1 - H))], and $B''_{BCh}$ = [(FA(1 - FA)) - (H(1 - H)]/[(FA(1 - FA)) + (H(1 - H))].

*Analyses.* For ease of presentation, the following analyses are presented in text for Experiments 1, 2, and 4: (1) analyses with the high-threshold correction (i.e., (H - FA)/(1 - FA)) for recollection (e.g., Yonelinas, 1994; Yonelinas et al., 1996); (2) analyses with the 2-equal threshold correction (i.e., (H - FA)) for know responding (e.g., Gardiner & Java, 1991); (3) analyses with IRK estimates of familiarity corrected by subtracting that estimate associated with false alarms; and, (4) analyses with *d'* estimates of familiarity. These analyses examine the most informative of the measures summarized in the preceding section; all other analyses are presented in tables in Appendix C. However, similarity and differences among all analyses will be discussed in text as appropriate.

Measures of effect size include eta-squared (i.e., $\eta^2$) for analysis of variance (ANOVA) effects and the point-biserial correlation-squared (i.e., $r^2$) for effects between two means. Both values provide a measure of the proportion of the variation accounted for by the independent variable for sample data.

Analyses for Donaldson's (1996) predictions based on a single-factor model of memory are presented in the *Sensitivity and Bias* sections of each experiment. If recognition memory is based on a single process, then Donaldson (1996) predicts that measures of memory (i.e., *A'*) obtained with recognition (i.e., yes-no) responses should not differ from those measures obtained with remember responses because sensitivity should be independent of response criterion (i.e., *A'*recognition = *A'*remember).

Hit and false alarm rates for know responses are based on areas *between* the yes-no criterion and the remember criterion (see Figure 3). For experimental designs with more than one criterion, measures of memory such as *A'* (or *d'*) for each criterion are based on cumulative probabilities (Macmillan & Creelman, 1991). Thus, as Donaldson (1996) pointed out, a measure of memory for know responding (i.e., *A*'know which is expected to represent area *between* criteria) is meaningless according to signal detection theory. However, he also indicated that while these *A*'know values may not provide sensible information about memory, there should be some regularity in the relationship between *A*'know and the yes/no recognition criterion in the single-factor model of recognition memory. Specifically, he predicted a positive correlation between measures of memory for know and placement of the recognition criterion. For example, when responding is conservative (i.e., criterion is positive and is displaced to the right; see Figure 5, top panel) the probability of a hit for know items (i.e., between the two criteria) will be greater than the probability of a false alarm for know items (i.e., between the two criteria) and a positive *A*'know value is derived. As responding becomes more liberal (i.e., criterion becomes more negative and is displaced to the left), the probability of a hit is reduced and may equal the false alarm rate (see Figure 5, middle panel). This will reduce the *A*'know value. If responding becomes very liberal (see Figure 5, bottom panel), then the probability of a hit becomes less than the probability of a false alarm (i.e., below chance performance). This will produce a negative *A*'know value. Thus, the *A*'know measure of memory is expected to vary directly with the placement of the recognition criterion. To assess this prediction, analyses for each experiment included derivation of a value for the correlation between *A*'know and $B''_D$ recognition. If this prediction is valid, then the correlation between these two values should be positive and significant.

As Donaldson suggested (1996), correlations between rescaled or adjusted know

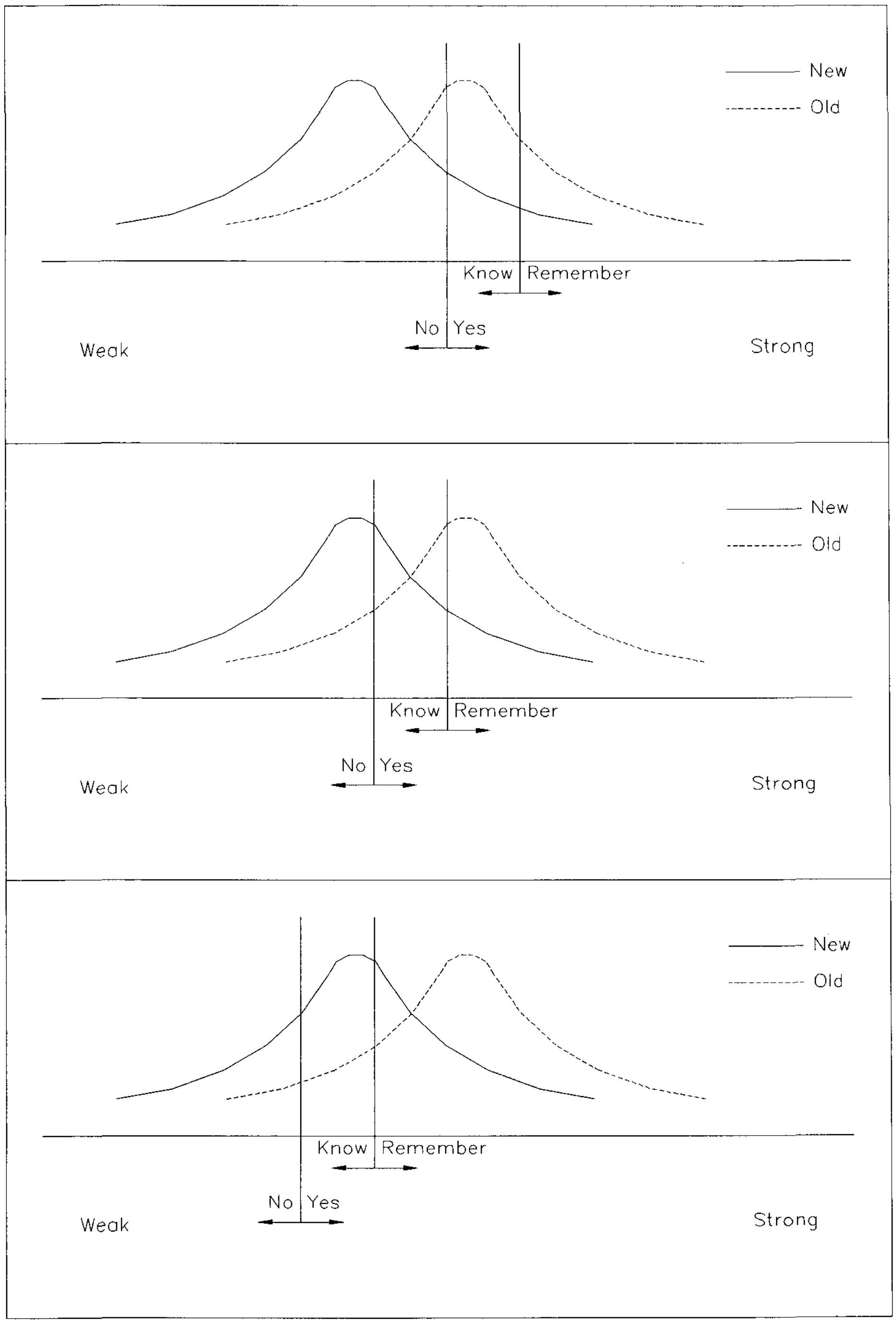



*Figure 5.* Changes in criteria and concomitant changes in probabilities of hits (old items) and false alarms (new items). Responding becomes more liberal as criteria are displaced to the left. See text for details. This figure is adapted from Donaldson (1996, p. 524).

responses and bias estimates were also derived (i.e., between *A*'know-IRK and $B"_D$ recognition). In rescaling, know responses are essentially adjusted with Jacoby's IRK equation (i.e., K/(1 - R), Jacoby et al., 1997). This results in the removal of that portion of the distribution on the right side of the remember criterion (see Figure 3). Predictions about measures of memory (i.e., *A*'know) are based on these new total distributions. In removing information about remember responses, the area associated with know responding is no longer between criteria, and derived *A*'know measures of memory will be meaningful. Thus, there should be no significant correlation between *A*'know-IRK and $B"_D$ recognition. Such a result is consistent with signal detection theory's prediction that there should be no correlation between measures of memory and criterion (Macmillan & Creelman, 1991).

## Experiment 1

Participants studied target faces using an encoding task. The number of exposures (i.e., one presentation vs. three presentations) was manipulated within subjects, and delay (i.e., none vs. 3-day) was manipulated between subjects. The contribution of recollection processes to responding as measured by the proportion of remember responses was predicted to be greater for repeated items than for non-repeated items. The opposite was expected for know responding (i.e., more know responses to non-repeated items than to repeated items). However, the estimate of familiarity as calculated with the IRK equation (i.e., F = K/(1 - R)) was expected to be greater for repeated items than for non-repeated items. That is, I predicted that repetition would enhance familiarity as well as recollection. Remember responding was anticipated to be less likely in the delay condition than in the immediate condition, and this was expected to be more evident for non-repeated than repeated items. Most importantly, it was predicted that know responses and the estimate of familiarity in the delay condition would be greater than that in the immediate condition, as memory information that initially supported recollection came to support familiarity instead.

*Method*

*Participants and design.* Eighty-eight University of Victoria undergraduate students participated for optional extra credit in an introductory psychology course or for a $5 payment. Data from 5 participants in the immediate test condition were not included because of experimenter error (3 participants) or computer error (1 participant) or failure to understand the test instructions (1 participant). Three participants in the delayed test condition did not return for the test session. Thus, data from 40 participants in each delay condition were analyzed.

Participants were told that they would be making judgments on a series of faces. They were not informed about the final recognition memory test. During the first phase of the experiment all participants indicated the sex of the individual portrayed in each photograph. They were informed that some of the faces would be presented more than once and that repeated faces were never altered from their original view. After this study phase, participants in the delay condition were dismissed and reminded to return in 3 days; Participants in the immediate condition completed an unrelated 20-minute filler task. In the final phase of the experiment, all participants completed the memory test. This test included studied and non-studied faces. Participants indicated whether or not the face was one presented in the first phase of the experiment (i.e., the phase in which judgments about sex were made). In addition, participants were instructed to provide remember or know responses for recognized faces. The experimenter recorded the participants' verbal study, and test, responses.

*Materials.* A pool of 120 colour photographs of unfamiliar adults (portraits included head and shoulders)[9] was created with an equal number of female and male faces. The pool was divided to create a set of critical faces (i.e., 96 faces) and a set of non-critical faces (i.e., 24 faces). Critical items included faces of individuals from different estimated age groups (i.e., 20 to 30 years of age, 30 to 40 years of age, 40 to 50 years of age, and 50 to 60 years of age) as determined by the experimenter. Each age group included an equal number of photographs. In addition, each age group consisted of an equal number of female and male faces. Critical items did not include faces with eye glasses, moustaches, or beards. There was a tan-coloured backdrop in each photograph.

A separate PCX file was created for each photograph. Each file consisted of a photograph centred on a black background that filled the screen (see figure 6). The file resolution was 28.35 pixels/cm and the photograph of the face measured 160 pixels (5.64 cm) in width and 240 pixels (8.47 cm) in height.

Critical items were divided into 2 sets (A and B) of 48 faces. Participants saw either A items or B items at study. Each set was further divided (A1 and A2; B1 and B2)

---

9

I wish to thank Dr. J. Bruno Debruille and the Centre de recherche Fernand-Seguin (Hôpital Louis-H Lafontaine, 7331 rue Hochelaga, Montréal, Québec Canada H2N 3V2) for their permission to use photographs from Med Bank (i.e, MultiPurpose Bank of European Descent faces).

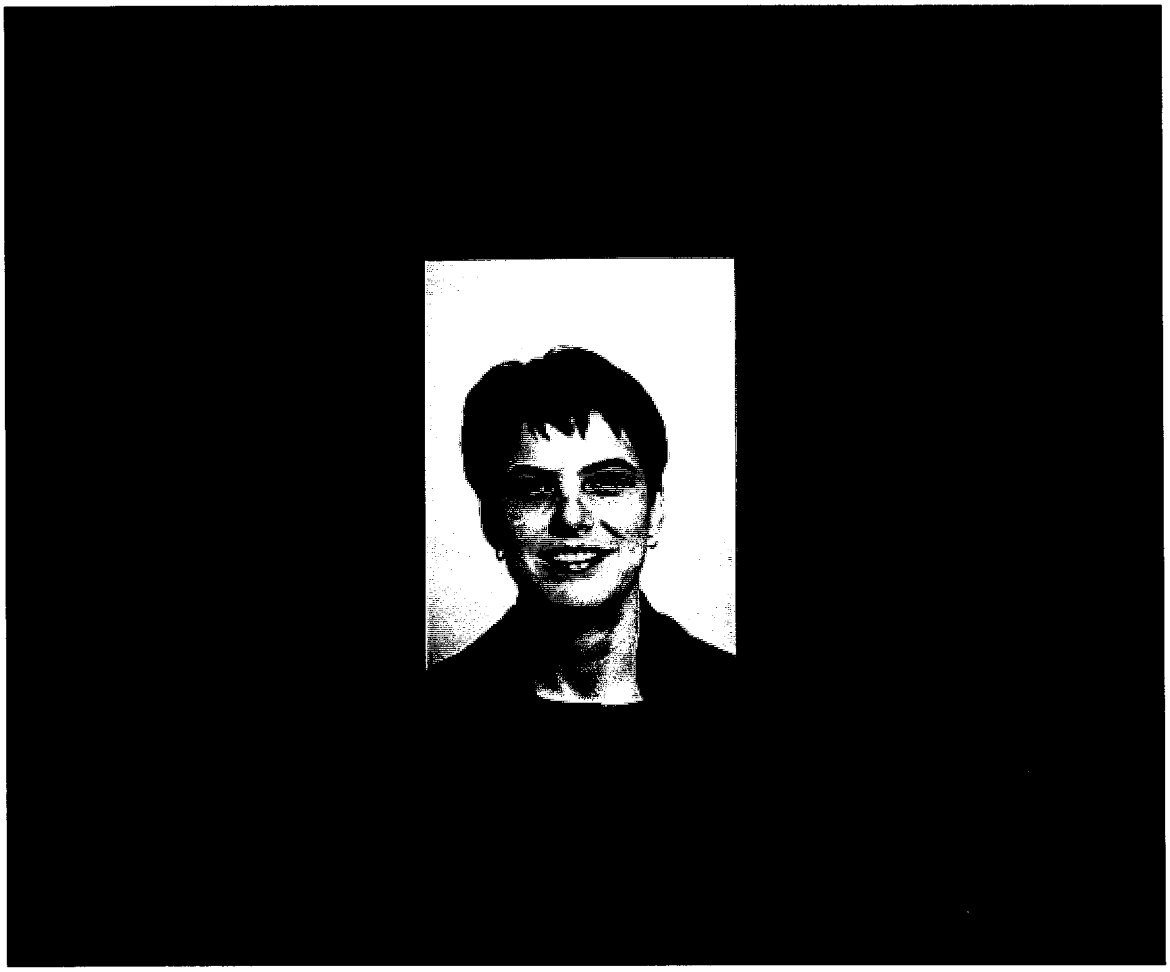


*Figure 6.* Example of stimulus item and presentation used in Experiments 1 and 2.

to accommodate the repetition manipulation. For example, for half of the "A" participants, A1 items were not repeated and A2 items were repeated three times. A1 items were yoked with B1 items, and A2 items were yoked with B2 items. For example, if A1 items were studied then B1 items served as the corresponding lures on the test. A1, A2, B1, and B2 items were rotated through all item types (i.e., repeated, non-repeated, lures for repeated items, lures for non-repeated items) to form 4 different study lists. Across lists, the item in each study list position differed but the item type did not. For example, in study list position 1 there was always a to-be-repeated item and it was always of a female judged to be 40 to 50 years of age. Study lists were divided into 3 blocks. Each block included one presentation of repeated items and a third of the non-repeated items. Each age group and each sex were equally represented in each third of the study list. There were 8 buffer items at the beginning, and 8 buffer items at the end, of the study list. These items did not change across counterbalancing study lists. Age and sex factors of items were counterbalanced for each study and test list.

There was one test for all participants. The test included all 96 critical items. Each half of the test included an equal number of A1, A2, B1, and B2 items. In addition, male and female faces were equally represented in each half. There were 8 (4 new and 4 old) buffer items at the beginning, and 8 (4 new and 4 old) buffer items at the end of the test list. The final criterion restricted the position of items on the test such that no more than three consecutive same correct recognition answers (yes or no) were allowed.

*Procedure.* Participants were tested individually. The experiment was conducted on an IBM-compatible personal computer using the Micro-Experimental Laboratory

Professional software package (version 2.0; Schneider, 1995).

At study, participants were instructed to provide a judgment about the sex of the individual in each photograph (i.e., male/female judgments). They were informed that some items would be repeated and that the same photograph of the individual was used on each repeated trial (i.e., photographs were never altered for the different trials). Each study trial began with the presentation of a face for 1s. The face was erased from the screen and participants provided verbal encoding judgments. The experimenter recorded each participant's responses.

After the initial encoding task, participants in the immediate condition completed an unrelated filler task for approximately 20 minutes. They were informed that they would continue with a face judgment task after completion of the filler task. Participants in the delay condition were dismissed after they were told that they would continue with a face judgment task in three days. Neither group of participants was told that their memory for the faces would be tested.

At test, participants were asked to indicate whether or not (i.e., yes or no) they recognized a face as one presented during the initial encoding task. In addition, for those faces that were recognized, they indicated whether they remembered or knew that the item was studied. Instructions for remember and know responding were printed on a computer monitor. Briefly, they were instructed to respond "remember" if recognition of the face was accompanied by specifying information that associated the face with the encoding phase. In contrast, participants were to respond "know" when they recognized the face and this recognition was not accompanied by any specifying information that associated the face with the encoding phase. The instructions also included an example of a remember and know response. The experimenter provided an additional verbal review of the instructions. Participants were encouraged to ask questions. A full description of the instructions is provided in Appendix A. All participants were given three practice trials before testing began.

*Results*

Each parameter estimate (i.e., remember responding, know responding, familiarity, *d'*) was analyzed in a separate 2 (delay: none, 3-day) X 2 (item type: non-repeated, repeated) mixed-model analysis of variance (ANOVA), with delay as the between-subjects variable and item type as the repeated variable.

*Lures.* A main effect of delay, $F(1, 78) = 5.21$, $MSE = .006$, $p < .03$, $\eta^2 = .051$, for remember responses to non-studied faces confirmed a change in response criterion for this response type (see Table 1). Participants were more likely to respond remember to new items in the immediate (i.e., no delay) condition ($M = .06$) than to new items in the delay condition ($M = .03$). For know responding to lures, only the interaction between repetition and delay approached significance, $F(1, 78) = 3.70$, $MSE = .004$, $p = .058$, $\eta^2 = .045$: On the immediate test, there was a tendency for know responding to be greater for repeated items than non-repeated items, $t(39) = 1.92$, $p = .062$, $r^2 = .086$ (see Table 2).

*Remember Responses.* Analysis of conditionalized responses (i.e., (H - FA)/(1 - FA), see Table 1) revealed a main effect of item type, $F(1, 78) = 194.39$, $MSE = .015$, $p < .001$, $\eta^2 = .71$, a main effect of delay, $F(1, 78) = 5.28$, $MSE = .035$, $p < .03$, $\eta^2 = .027$, and a nonsignificant interaction, $F < 1$. As predicted, repeated items ($M = .38$) were more

Table 1

*Mean Remember Response Probabilities in Experiment 1*

| | No Delay | | 3D Delay | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Non-repeated | | | | |
| Targets | .19 | .02 | .10 | .02 |
| Lures | .06 | .01 | .03 | .01 |
| (H - FA) | .13 | .02 | .08 | .01 |
| (H - FA)/(1 - FA) | .14 | .02 | .08 | .01 |
| Repeated | | | | |
| Targets | .44 | .03 | .36 | .03 |
| Lures | .06 | .01 | .03 | .01 |
| (H - FA) | .39 | .03 | .33 | .03 |
| (H - FA)/(1 - FA) | .41 | .03 | .34 | .03 |

*Note.* 3D = 3 days; *M* = Mean; *SEM* = standard error of the mean; H = hits; FA = false alarms.

Table 2

*Mean Know Response Probabilities in Experiment 1*

| | No Delay | | 3D Delay | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Non-repeated | | | | |
| Targets | .22 | .02 | .16 | .02 |
| Lures | .11 | .01 | .11 | .02 |
| (H - FA) | .13 | .02 | .07 | .01 |
| Repeated | | | | |
| Targets | .22 | .02 | .23 | .02 |
| Lures | .14 | .02 | .10 | .02 |
| (H - FA) | .11 | .02 | .14 | .02 |

*Note.* 3D = 3 days; *M* = Mean; *SEM* = standard error of the mean; (H - FA) = hits minus false alarms.

likely remembered than non-repeated items ($M = .11$). In addition, delaying the test for 3 days reduced the probability of providing remember responses (immediate $M = .28$; delay $M = .21$).

Table C1, found in Appendix C, lists the results for analyses with uncorrected responses and for responses corrected using the 2-equal threshold method (i.e., H - FA). The results from these latter two analyses did not differ from those using conditionalized responses listed above.

*Know Responses.* Corrected know response probabilities (i.e., H - FA, see Table 2) showed a significant interaction, $F(1, 78) = 8.13$, $MSE = .009$, $p < .01$, $\eta^2 = .094$, and a main effect of item type that fell short of statistical reliability, $F(1, 78) = 3.32$, $MSE = .009$, $p = .072$, $\eta^2 = .041$. Additional analyses of the interaction found that non-repeated faces were less likely known if they were tested after 3 days than if they were tested immediately following study, $t(78) = 2.50$, $p < .02$, $r^2 = .074$. There was no significant difference in know responding to repeated items across the delay factor, $p = .20$. Item type was also examined within each test condition: Know responding to non-repeated items did not differ from that to repeated items in the immediate test, $t < 1$; and, in the delayed test condition, participants were more likely to know repeated items than non-repeated items, $t(39) = 3.66$, $p < .001$, $r^2 = .26$.

Similar to the above analyses the interaction was significant when uncorrected know responses were analyzed. Unlike the above analysis, the main effect of item type was significant. However, post hoc tests revealed results identical to those found with corrected know probabilities (see Table C2 in Appendix C).

*IRK Familiarity.* Corrected IRK familiarity estimates (i.e., H - FA, see Table 3) revealed a main effect of item type, $F(1, 78) = 54.02$, $MSE = .016$, $p < .001$, $\eta^2 = .41$, and a significant interaction effect, $F(1, 78) = 6.45$, $MSE = .016$, $p < .02$, $\eta^2 = .076$. Additional analyses found that familiarity was more likely to contribute to responses to repeated items than to non-repeated items in the immediate test, $t(39) = 3.20$, $p < .005$, $r^2 = .21$, and the delayed test, $t(39) = 7.50$, $p < .001$, $r^2 = .59$. The contribution of familiarity to responses to non-repeated items decreased after 3 days, $t(78) = 3.03$, $p < .005$, $r^2 = .11$. In contrast, familiarity associated with repeated items was not significantly affected by the delay, $t < 1$, although this nonsignificant tendency for familiarity to increase across the Delay factor was in the predicted direction.[10]

---

10

Estimates of familiarity derived with the independence remember/know (IRK) equation and corrected remember and corrected know probabilities were also analyzed. Similar to the results with IRK estimates of familiarity obtained as a difference between old familiarity and new familiarity there was a main effect of item type, $F(1, 78) = 16.33$, $MSE = .015$, $p < .001$, $\eta^2 = .17$, and a significant interaction, $F(1, 78) = 9.47$, $MSE = .015$, $p < .005$, $\eta^2 = .11$. For non-repeated items, delaying the recognition test reduced the estimate of familiarity (immediate $M = .14$; delay $M = .08$), $t(78) = 2.69$, $p < .01$, $r^2 = .085$. There was no reliable effect of delay on repeated items, although the trend was in the predicted direction (immediate $M = .16$; delay $M = .22$), $t(78) = 1.59$, $p = .12$, $r^2 = .031$. The contribution of familiarity to responding on the immediate test was not affected by the repetition variable, $t < 1$. However, on the delayed test there was more familiarity associated with repeated items than with non-repeated items, $t(39) = 5.47$, $p < .001$, $r^2 = .43$.

Despite the ease of the method using corrected probabilities, there exists the possibility of restriction of values. Corrections to remember and know probabilities are affected by below chance performance (i.e., false alarm rates greater than hit rates) because negative values are raised to 0.00. For remember responses this was a rare occurrence in the immediate condition only: once for non-repeated items and twice for repeated items. However, corrections to know

Table 3

*Mean IRK and d' Estimates of Familiarity in Experiment 1*

| | No Delay | | 3D Delay | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Non-repeated | | | | |
| Targets | .27 | .02 | .18 | .02 |
| Lures | .12 | .02 | .11 | .02 |
| (H - FA) | .16 | .02 | .08 | .02 |
| *d'* | .62 | .08 | .33 | .08 |
| Repeated | | | | |
| Targets | .40 | .03 | .39 | .03 |
| Lures | .14 | .02 | .11 | .02 |
| (H - FA) | .26 | .03 | .29 | .03 |
| *d'* | .83 | .11 | 1.09 | .09 |

*Note.* IRK = independence remember/know; 3D = 3 days; *M* = Mean; *SEM* = standard error of the mean; (H - FA) = hits minus false alarms.

---

responses required frequent re-setting of obtained values for below chance performance. In the immediate condition, there were 6 conversions to zero for non-repeated items, and 11 conversions for repeated items. The pattern of conversions for the delayed condition was different: 9 for non-repeated items and 3 for repeated items. Because these conversions would likely artificially, and significantly, restrict the range of values for parameter estimates, the method of correcting remember and know responding before deriving estimates of familiarity will be more vulnerable to low levels of performance. The difference method which provides an estimate of familiarity beyond that experienced for new items would less likely suffer from an artificial restriction of range of values. In this experiment familiarity for new items was never greater than the familiarity associated with old items.

Analyses with uncorrected familiarity estimates produced identical results. These analyses are listed in Table C3 in Appendix C.

*d' Estimate of Familiarity.* With one minor difference, analysis of the data with signal detection theory was similar to the above results (see Table 3). There was a significant main effect of item type, $F(1, 78) = 36.33$, $MSE = .26$, $p < .001$, $\eta^2 = .32$, and an Item Type x Delay interaction, $F(1, 78) = 11.79$, $MSE = .26$, $p < .01$, $\eta^2 = .13$. Additional analyses revealed that unlike IRK familiarity, the difference in sensitivity between repeated and non-repeated items in the immediate test condition was not reliable, $t(39) = 1.83$, $p = .075$, $r^2 = .079$. In the delayed test condition, sensitivity to repeated items was significantly greater than that to non-repeated items, $t(39) = 6.72$, $p < .001$, $r^2 = .54$. There was a detectable reduction in familiarity (i.e., sensitivity) for non-repeated items tested after a 3-day delay, $t(78) = 2.47$, $p < .02$, $r^2 = .073$. For repeated items, there was a tendency for the $d'$ estimate of familiarity to be greater after the delay than when tested immediately, $t(78) = 1.89$, $p = .062$, $r^2 = .044$.

*Sensitivity and Bias.* Non-parametric estimates of sensitivity (i.e., $A'$) and bias (i.e., $B''_D$) are presented in Table 4. In the immediate condition, Donaldson's (1996) first prediction (i.e., $A'$recognition = $A'$remember) was true for repeated items, $t(39) = 1.45$, $p = .16$, $r^2 = .051$, but not for non-repeated items, $t(39) = 2.98$, $p < .005$, $r^2 = .19$. The reverse was found for items in the delay condition: the two measures of memory differed for repeated faces, $t(39) = 4.31$, $p < .001$, $r^2 = .32$, but, for non-repeated faces the two measures were not reliably different, $t < 1$.

In this experiment, correlations between measures of memory for know responding

Table 4

*A' and $B''_D$ Estimates in Experiment 1*

| | No Delay | | 3D Delay | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| *A'* | | | | |
| Non-repeated | | | | |
| Recognition | .72 | .01 | .65 | .02 |
| Remember | .67 | .02 | .63 | .02 |
| Know | .64 | .02 | .58 | .02 |
| Know-IRK | .67 | .02 | .59 | .02 |
| Repeated | | | | |
| Recognition | .83 | .01 | .83 | .01 |
| Remember | .81 | .01 | .80 | .01 |
| Know | .59 | .03 | .67 | .02 |
| Know-IRK | .71 | .02 | .75 | .01 |
| $B''_D$ | | | | |
| Non-repeated | | | | |
| Recognition | .65 | .05 | .78 | .05 |
| Repeated | | | | |
| Recognition | .28 | .08 | .56 | .07 |

*Note.* 3D = 3 days; *M* = mean; *SEM* = standard error of the mean; IRK = independence remember/know, K/(1 - R).

were positively correlated with response bias (re: Donaldson's second prediction). In the immediate condition the correlation was .45, $p < .005$, for non-repeated items, and .59, $p < .001$, for repeated items. The same pattern occurred for items in the delay condition: .42, $p < .01$, for non-repeated items and .47, $p < .005$, for repeated items.

Rescaling (i.e., removing that portion of the distribution associated with remember responses; K/(1 - R)) reduced the correlation for both repeated and non-repeated items. The reduction for repeated items (i.e., revised correlation of .04, $p = .80$, in the immediate condition and -.19, $p = .24$, in the delay condition) was more pronounced than that for non-repeated items (i.e., revised correlation of .33, $p < .04$, for the immediate condition, and .33, $p < .04$, in the delay condition). The correlations between rescaled *A*' and the recognition criterion for repeated items were not significant. In contrast, the revised *r* values for the non-repeated items continued to be significantly correlated even after adjusting know responses.

Measures of memory associated with hit and false alarm rates of remember responses (i.e., *A*'remember) were not always equivalent to those measures obtained with overall hit and false alarm rates (i.e., *A*'recognition). *A*'recognition values were always greater than *A*'remember values, suggesting that this measure of memory reflects an additional process associated with know responding (e.g., Gardiner & Conway, 1999; Gardiner, Gregg, Mashru, & Thaman, 2001). Memory measured with know responses was at times correlated with response bias (i.e., $B''_D$), suggesting that measures of memory were not independent of criterion placement as expected with a strength-based single-process model. However, when know responses were adjusted as suggested by Jacoby et al. (1997) in his dual-process model of memory, correlations between memory (i.e., $A$'know) and response bias (i.e., $B''_D$) were reduced, and for repeated items they were statistically eliminated. Thus, these results suggest that the dual-process model rather than the single-process model adequately describes recognition memory in the remember/know task.

*Discussion*

In summary, repeating items at study increased the contribution of recollection to responding at test regardless of any, or the type of, applied correction. In addition, delaying the recognition memory test reduced remember responding to faces. Neither of these findings is surprising. Jacoby et al. (1998) found no difference in know responding between words studied once and those studied three times. In this experiment this was true for faces in the immediate test but not for faces in the delayed test: On the latter, repetition increased knowing. Because of repetition, the amount of information encoded with items was expected to be greater and more varied than for items that were not repeated. For either type of item, after process interference caused by delay, some information would not be accessible or available, and other information may no longer be bound to information that specifically identifies source (i.e., the study list). However, given that repeated items rather than non-repeated items have more information associated with them, then repeated items rather than non-repeated items would more likely have information accessible or available to be recalled at test (albeit not necessarily source-specifying information). This information would not be sufficient to produce a remember response but would be sufficient to produce a know response. In other words, information for repeated items that on the immediate test contributed to remember responses could only contribute to know responses on the delayed test.

When familiarity was derived with the independence remember/know equation, as predicted, it was more likely to contribute to responses to repeated items than to non-repeated items in both test conditions. A reduction in recollection after a delay was expected to be accompanied by an increase in familiarity (whether measured as IRK familiarity or $d'$). The data for non-repeated items contradicted this prediction because the decrease in recollection was associated with a decrease in familiarity. For repeated items, the decrease for recollection and increase for familiarity fit the prediction but the change for familiarity was not significant. Because overall recognition was low (immediate test: non-repeated $M = .41$, repeated $M = .66$; delayed test: non-repeated $M = .26$, repeated $M = .59$), especially for non-repeated items, even a reduction in remember responding would not necessarily be a product of an increase in know responding (i.e., the increase in know responding may have also been restricted). If there were increases in know responding, the decrease in overall responding would make it difficult to observe any meaningful changes in know responding and ultimately in estimates of familiarity.

Yonelinas (Yonelinas, Regehr et al., 1995; Yonelinas et al., 1998) suggested that failure to incorporate response bias may lead to conclusions that differ greatly from those obtained from methods that take into account criterion changes. Despite the change in response criterion (as indicated by the analysis of lures for remember responses) between the immediate and delayed tests, final conclusions based on uncorrected and corrected probabilities or parameter estimates did not differ. Although corrected know probabilities and familiarity estimates revealed a nonsignificant tendency toward the predicted increase in familiarity after the delay, $d'$ estimates of familiarity appear to be more sensitive to change than IRK estimates because the estimate of power associated with the former was larger than that for the latter: The increase for repeated items was associated with a power of .16 (one-tailed; obtained with GPOWER software program, Erdfelder, Faul, & Buchner, 1996) with IRK familiarity (hits minus false alarms) and a power of .59 (one-tailed) with $d'$ familiarity.

## Experiment 2

The initial low estimates of recollection and familiarity for non-repeated items in Experiment 1 indicate that the encoding task only provided a minimal opportunity for the encoding of processes necessary for long-term retrieval. Judgments about sex are quick and automatic, and consequently memory for the once-presented faces was poor even on the immediate test. To better assess the effects of repetition and delay on memory for unfamiliar faces, a "deeper" encoding task was implemented in Experiment 2. Rating the attractiveness of the face rather than making a judgment about sex should increase the amount of, and variety of, information initially encoded. To avoid ceiling effects with remember responses, a one-week, rather than a 3-day, delay between study and test was imposed. This longer delay would allow sufficient interference with memory processes (e.g., encoding processes that bound different attributes), and significantly decrease the estimate of recollection.

The purpose of Experiment 1 was to show that memory information that on an immediate test contributed to estimates of recollection would come to contribute to estimates of familiarity on a delayed test. Such migration of responding would be due to the "unbinding" of information that comes about because of delay. The results for repeated items in Experiment 1 were encouraging in this respect because estimates of familiarity (either IRK or $d'$) were greater after delay. However, these increases were not statistically reliable. As suggested above, it may be that the very shallow encoding task used in Experiment 1 left participants with very poor memories of the faces, such that the rate of decline of any recognition of studied faces exceeded the rate of migration of memory influences from recollection to familiarity (R-to-F) across delay. That is, the rate of forgetting (F-to-New) may have been greater than the rate of "unbinding" (R-to-F).

In Experiment 2, participants provided recognition (yes/no) responses, and remember and know responses (when appropriate) immediately following the study session for half of the studied items, and one week later for the remaining items. The use of a different set of the same type of item (i.e., faces) on the second test avoided the potential effect(s) of additional study provided by the first test. Responses to item types on the initial test were compared to responses to these same item types on the final test. It was predicted that overall recognition of repeated items would be greater than overall recognition of non-repeated items. In addition, remember responses on the second test should be less than those on the first test, whereas know responses should be greater on the second test than on the first test. Estimated familiarity was also expected to increase with delay. Such changes would support the notion of independent attributes because information encoded at study would be shown to contribute to different independent processes depending upon when the items are tested (i.e., recollection processes on the first test, and familiarity processes on the second test). Other dual process theories (e.g., Yonelinas & Jacoby, 1996; Gardiner & Java, 1991) posit that a particular piece of information encoded at study will contribute to a particular process regardless of the requirements of the task: Over time, information may be lost, but according to these models information should not "migrate" from R to F.

To monitor changes in responses to individual items across a delay, Experiment 2 included a test-retest paradigm (i.e., a particular face was shown both on an immediate test and a delayed test). However, the retest data proved to be confounded and therefore are not reported here.[11] Because the counterbalancing of items across conditions was based on inclusion of the test-retest paradigm, reference to these item types remains in the *Method* section.

*Method*

*Participants and design.* Forty-four University of Victoria undergraduate students participated for optional extra credit in an introductory psychology course. Data from 4

---

11 In support of a functional model of recognition memory, of those studied items given remember responses on the first test, more became know responses (i.e., R-to-K) rather than new responses (i.e., R-to-New) on the second test ($\underline{z} = 3.76$, $p < .001$). In addition, for items given a know response on the first test, more were remember responses (i.e., K-to-R) rather than new responses (K-to-New) on the second test ($\underline{z} = 5.44$, $p < .001$). However, the overall proportion of K-to-R responses was greater than the proportion of R-to-K responses ($\underline{z} = 8.51$, $p < .001$). Thus, in analyzing the proportion of migrating responses it became apparent that re-testing introduced a confound. Because the first test served as an additional study episode for these items, the effect of delay (i.e., a potential weakening of memorial bonds among attributes) may have been attenuated because of the increased likelihood of strengthening of memory bonds after additional encoding.

participants were not included because they failed to return for the second test (3 participants) or they did not understand the test instructions (1 participant). Thus, data from 40 participants were analyzed.

All participants completed 2 sessions. In the first session participants judged the attractiveness of individuals in photographs. This phase was followed by an unrelated filler task and then a recognition memory test (Test 1). The second session took place one week later and participants were required only to complete another recognition memory test (Test 2). Half of the studied items were on the first test and all of the studied items were on the second test. Thus, some items were initially tested on the first test and then re-tested on the second test while others were only tested on the second test. As in Experiment 1, participants were initially unaware of the tests; They were simply informed that they would be making judgments on a series of faces. In addition, participants in Experiment 2 were informed that both sessions would involve similar-type judgments. Consequently, although they were not explicitly informed of the second test, it is highly likely that some participants were aware of the type of task in the second session. The materials and procedure were the same as those in Experiment 1 except as noted.

*Materials.* Critical item sets A and B were each divided into 4 sets of 12 photographs (A1, A2, A3, A4, B1, B2 ,B3, B4) to accommodate the repetition and test manipulation. A1 items were yoked with B1 items, A2 items were yoked with B2 items, and so forth. Participants saw A or B items at study, half of the A items and half of the B items on the first test, and all A and B items on the second test. For example, if A1 and A2 were non-repeated items, and A3 and A4 were repeated items in the study phase, then

A1, A3, and their corresponding lures, B1 and B3, were on the immediate test. All items were rotated through all item types to form 8 different study lists. There were 6 buffer items at the beginning, and 6 buffer items at the end, of the study list.

The first test included 48 critical items and the second test included all 96 critical items. There were 8 buffer items at the beginning, and end, of each test. Buffer items at each end of Test 1 included 4 studied and 4 non-studied items. For Test 2, each group of buffer items included 2 never-seen items, 2 items from the original study list not presented in Test 1, 2 items from the original study list and presented in Test 1, and 2 items not on the study list and presented in Test 1.

*Procedure.* Participants rated the attractiveness of the individual in the photograph during study (response categories: 1 = very unattractive; 2 = unattractive; 3 = slightly unattractive; 4 = slightly attractive; 5 = attractive; 6 = very attractive). Each stimulus item was presented for 1 second before it was replaced by the attractiveness rating scale noted above. The scale remained on the monitor until the participant entered a number response on the keyboard. Participants were informed that each rating for a repeated face could be treated independent of the others; they were not required to remember a previous rating.

The study phase was followed by an unrelated filler task that lasted about 20 minutes. After this distractor task, participants were informed that their memory for the faces would be tested. Upon completion of Test 1, participants were reminded to return one week later to make more face judgments, and then dismissed. There was no mention of another test and participants knew that the next session would only take 15 to 20 minutes to complete.

For the second test, participants were instructed to make face recognition judgments regardless of source: Participants responded YES to a face they believed to have been presented a week earlier either at study, or at Test 1, or both at study and Test 1. They were also informed that a re-presented face was never altered from its original view. If they recognized a face, then they provided a remember or know response as defined in Experiment 1. The experimenter recorded all test responses on the keyboard.

*Results*

Remember responses, know responses, estimates of familiarity, and $d'$ values were analyzed in separate 2 (delay: none, 1-week) X 2 (item type: non-repeated, repeated) repeated measures ANOVAs.

*Lures.* There was no significant difference between remember responses to lures on Test 1 (i.e., no delay, immediate test) and those only tested on Test 2 (i.e., 1-week delay) regardless of the assigned item type, all $ps \geq .10$ (see Table 5). Know responses to lures were also unaffected by the delay, all $ps \geq .43$ (see Table 6).

*Remember responses.* Analysis of remember responses (Table 5) revealed significant main effects only: $F(1, 39) = 438.23$, $MSE = .025$, $p < .001$, $\eta^2 = .92$ for the Item Type factor and $F(1, 39) = 75.79$, $MSE = .019$, $p < .001$, $\eta^2 = .66$ for the Delay factor. As expected, repeated items ($M = .72$) were more often remembered than non-repeated items ($M = .20$). Also in keeping with predictions was the finding that the delay significantly reduced the probability of a remember response (immediate $M = .56$; delay $M = .37$).

Table 5

*Mean Remember Response Probabilities in Experiment 2*

| | No Delay | | 1W Delay | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Non-repeated | | | | |
| Targets | .31 | .03 | .10 | .02 |
| Lures | .004 | .004 | .002 | .02 |
| (H - FA) | .30 | .03 | .09 | .02 |
| (H - FA)/(1 - FA) | .30 | .03 | .09 | .02 |
| Repeated | | | | |
| Targets | .81 | .03 | .64 | .04 |
| Lures | .01 | .005 | .004 | .003 |
| (H - FA) | .80 | .03 | .63 | .04 |
| (H - FA)/(1 - FA) | .81 | .03 | .64 | .04 |

*Note.* 1W = 1 week; *M* = mean; *SEM* = standard error of the mean; H = hits; FA = false alarms.

Table 6

*Mean Know Response Probabilities in Experiment 2*

| | No Delay | | 1W Delay | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Non-repeated | | | | |
| Targets | .35 | .03 | .26 | .03 |
| Lures | .05 | .01 | .05 | .01 |
| (H - FA) | .30 | .03 | .21 | .03 |
| Repeated | | | | |
| Targets | .17 | .03 | .25 | .03 |
| Lures | .04 | .01 | .05 | .01 |
| (H - FA) | .15 | .03 | .22 | .04 |

*Note.* 1W = 1 week; *M* = mean; *SEM* = standard error of the mean; (H - FA) = hits minus false alarms.

Analyses of uncorrected remember responses and those corrected with the (H - FA) method detected significant main effects of item type and delay that were consistent with those reported above (see Tables C4 and C5 in Appendix C).

*Know responses.* In addition to a reliable effect of item type, $F(1, 39) = 5.22$, $MSE = .040, p < .03, \eta^2 = .12$, there was a significant interaction effect, $F(1, 39) = 9.38$, $MSE = .029, p < .005, \eta^2 = .19$ (see Table 6). On the immediate test, repeated items were less likely to be known than non-repeated items, $t(39) = 4.41, p < .001, r^2 = .33$. However, repeated items were just as likely to be known as non-repeated items on the delayed test, $t < 1$. *t*-tests also indicated that after the delay there was a significant decrease in know responding for non-repeated items, $t(39) = 2.46, p < .02, r^2 = .13$. In contrast, and as expected, there was a reliable increase for repeated items, $t(39) = 2.73, p < .01, r^2 = .16$.

Table C6 in Appendix C lists the analysis for uncorrected know responses. The results of this analysis do not differ from those for corrected responses.

*IRK Familiarity.* The analysis of IRK estimates of familiarity for immediate and delay items revealed reliable main effects of item type, $F(1, 39) = 16.02, MSE = .121, p < .001, \eta^2 = .29$, and delay, $F(1, 39) = 7.58, MSE = .077, p < .01, \eta^2 = .16$ (see Table 7). The decrease in familiarity for non-repeated items appeared to be less than that for repeated items. However, the probability of a Type I error associated with the interaction for this analysis was greater than the accepted alpha level, $F(1, 39) = 3.60, MSE = .092, p = .065, \eta^2 = .084$. Thus, interpretations were based on main effects only.[12] As expected,

12

Further probing of the interaction revealed no significant change in familiarity after a one-week

Table 7

*Mean IRK and d' Estimates of Familiarity in Experiment 2*

| | No Delay | | 1W Delay | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Non-repeated | | | | |
| Targets | .50 | .04 | .29 | .03 |
| Lures | .05 | .01 | .05 | .01 |
| (H - FA) | .46 | .04 | .24 | .03 |
| *d'* | 1.55 | .12 | .89 | .10 |
| Repeated | | | | |
| Targets | .61 | .07 | .59 | .06 |
| Lures | .04 | .01 | .05 | .01 |
| (H - FA) | .58 | .07 | .55 | .06 |
| *d'* | 1.94 | .27 | 1.83 | .20 |

*Note.* IRK = independence remember/know; 1W = 1 week; *M* = mean; *SEM* = standard error of the mean; (H - FA) = hits minus false alarms.

---

delay for repeated items, $t < 1$. However, familiarity associated with non-repeated items decreased after the delay, $t(39) = 4.95$, $p < .001$, $r^2 = .39$. There was no effect of repetition on the immediate test, $t(39) = 1.55$, $p = .13$, $r^2 = .058$, but repeated items were more familiar than non-repeated items on the delayed test, $t(39) = 5.11$, $p < .001$, $r^2 = .40$.

the contribution of familiarity to responding to repeated items ($M = .57$) was greater than that to non-repeated items ($M = .35$). In contrast to the predicted increase in familiarity after a delay, there was an overall reduction (immediate $M = .52$; delay $M = .40$).

Analysis of uncorrected IRK estimates of familiarity provided analogous results (see Table C7 in Appendix C).

*d' Estimate of Familiarity.* The results of analysis of *d'* values were consistent with those obtained with IRK estimates of familiarity (see Table 7). There were significant main effects and no significant interaction, $p > .11$. A main effect of item type, $F(1, 39) = 11.59$, $MSE = 1.53$, $p < .005$, $\eta^2 = .23$, revealed that sensitivity to repeated items ($M = 1.89$) was greater than that for non-repeated items ($M = 1.22$). In addition, the delay in testing significantly reduced this measure of memory, $F(1, 39) = 5.74$, $MSE = 1.02$, $p < .03$, $\eta^2 = .13$ (immediate $M = 1.74$; delay $M = 1.36$).

*Sensitivity and Bias.* Derived *A'* values are presented in Table 8. As indicated in the *General Method* section, in a single-process model a measure of memory obtained using the probability of recognition should not differ from that obtained using the probability of remember responding because according to signal detection theory these measures are independent of response criterion. In Experiment 2, these two measures of memory differed in all four conditions, and the recognition measure was always greater than the remember measure: (1) $t(39) = 8.58$, $p < .001$, $r^2 = .65$, for non-repeated items in the immediate condition; (2) $t(39) = 4.09$, $p < .001$, $r^2 = .30$, for repeated items in the immediate condition; (3) $t(39) = 7.87$, $p < .001$, $r^2 = .61$, for non-repeated items in the delay condition; and, (4) $t(39) = 4.62$, $p < .001$, $r^2 = .35$, for repeated items in the delay

Table 8

*A' and $B''_D$ Estimates in Experiment 2*

| | No Delay | | 1W Delay | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| *A'* | | | | |
| Non-repeated | | | | |
| Recognition | .87 | .01 | .75 | .02 |
| Remember | .76 | .02 | .62 | .02 |
| Know | .75 | .02 | .70 | .02 |
| Know-IRK | .81 | .02 | .71 | .02 |
| Repeated | | | | |
| Recognition | .97 | .003 | .94 | .01 |
| Remember | .93 | .01 | .88 | .01 |
| Know | .61 | .03 | .66 | .03 |
| Know-IRK | .77 | .04 | .81 | .03 |
| $B''_D$ | | | | |
| Non-repeated | | | | |
| Recognition | .65 | .05 | .85 | .04 |
| Repeated | | | | |
| Recognition | -.12 | .04 | .21 | .08 |

*Note.* 1W = 1 week; *M* = mean; *SEM* = standard error of the mean; IRK = independence remember/know, K/(1 - R).

condition.

As Donaldson (1996) indicated, repeated items showed a moderate to high positive correlation between the $A'$know measure of memory (i.e., sensitivity) and the recognition response bias in the immediate ($r = .37, p < .03$) and delay ($r = .66, p < .001$) conditions. These correlations were significantly attenuated when know responding was rescaled as described in Jacoby et al.'s, (1997) independence remember/know procedure: In the immediate condition the correlation was reduced to $.13, p = .43$, and in the delay condition it was reduced to $.21, p = .19$. Thus, when know responses in the *absence* of remember responses were considered, sensitivity, or memory, was not correlated with criterion, or response bias, as signal detection theory predicts. These results do not support the notion of a single memory process where quantitative differences on a continuum of strength give rise to either know (weak) or remember (strong) responses.

In the immediate condition, the correlation between sensitivity and criterion for non-repeated items was a positive value as Donaldson (1996) predicted, however, it was not reliable, $r = .05, p = .76$. The correlation for the delay condition was significantly negative, $r = -.35, p < .03$. These values are very different from those obtained in the first experiment (immediate $r = .45, p < .005$; 3-day delay $r = .42, p < .01$). The reasons for the differences remain unclear and perhaps warrant further investigation. However, they do not support Donaldson's predictions about the behaviour of parameter estimates in the single-process model.

*Discussion*

In summary, results of analyses of remember and know probabilities were not affected by the type of correction (i.e., none, H - FA, or (H - FA)/(1 - FA)). In addition, analyses of IRK- and *d'*-derived estimates of familiarity also provided comparable results.

As in Experiment 1, the results of Experiment 2 support the view that recollection processes were more likely available or accessible for repeated items than for non-repeated items. In addition, delaying the testing session interfered with recollection processes such that the probability of a remember response was reduced after one week.

If Gardiner's assertions (e.g., Gardiner & Parkin, 1990) are applied to know responses, then in the immediate test there exists the odd result that non-repeated faces were more familiar than repeated faces. From a phenomenological point of view these results make sense: participants were more likely to remember repeated faces leaving few faces to be known. If memory processes are expected to be independent as Jacoby and colleagues suggest (e.g., Jacoby, 1998; Yonelinas, 1994), then simply using the number of know responses to estimate the contribution of familiarity would underestimate this value for repeated faces in the immediate test because this method would fail to account for the contribution of familiarity processes to remember responses (and hence familiarity processes appear to contribute more to responses to non-repeated faces than to repeated faces).

For repeated items, as expected, the decrease in recollection was accompanied by an increase in know responding after the delay. In Experiment 1, the increase associated with these items was not significant. Although the decrease in IRK familiarity for repeated items in Experiment 2 was only slightly negative, the overall analysis indicated a significant decrease after the delay. Estimates of *d'* revealed an identical pattern of results. Thus, any R-to-F changes appeared to be overshadowed by F-to-New changes for repeated items. Non-repeated items showed a significant decrease in remember responding, know responding, and IRK familiarity suggesting that perhaps these items were more likely forgotten after a week.

One of the goals of this experiment was to improve encoding conditions so as to enhance the formation of bonds between and among memory attributes. In addition, there was an increase from 3 to 7 days to allow sufficient time to weaken but not eliminate memory processes. This would permit recall of memorial information that could not necessarily be attributed to a particular face in the study session. Consequently, information that on an immediate test would contribute to estimates of recollection would contribute to estimates of familiarity on a delayed test. Although know responding did increase for repeated items, the results of this experiment showed a significant reduction in both recollection and IRK familiarity after the delay. An examination of individual participant data for the delayed test revealed that six participants responded remember to all 12 critical repeated items and one participant responded remember to 11 and forgot 1. The removal of these seven participants from the data set produced significant decreases in R across time, significant increases in K across time, and a non-reliable increase in IRK familiarity (immediate $M = .59$; delay $M = .67$; $t(32) = 1.03, p = .31$).[13] Although the increase for familiarity was not significant, it was clearly in the predicted direction whereas the full data set indicated a slight decrease in familiarity across time (immediate $M = .58$;

13

Reported values are corrected for false alarms.

delay $M = .55$).[14] Thus, these floor and ceiling effects for know and remember responses may indicate that the delay may not have sufficiently interfered with memory processes for these seven participants.

The question of how much of a delay is sufficient to weaken but not eliminate memory processes is empirical. To curtail a prolonged search for the optimal number of item repetitions and delay period, a new paradigm was implemented that did not include either manipulation.

## Experiment 3

As mentioned in the introduction, the classic example of knowing in the absence of remembering is the nagging sense of familiarity that often arises when a familiar individual is encountered in an unusual context (e.g., the butcher on the bus). Surprisingly, however, no published study has explored the effects of context on remember/know responses to faces. Experiments 3 and 4 were designed to fill this gap in the literature, while simultaneously exploring the same basic theoretical issues that motivated Experiments 1 and 2.

In Experiments 3 and 4, participants studied a set of colour photographs of faces, each presented one at a time with a unique context photograph (e.g., an ornate fountain, a Japanese pagoda, etc.). In Test 1, studied faces (intermixed with non-studied faces) were

---

14

The pattern of results was not altered for non-repeated items when these same participants were omitted: remember and know responding decreased, and F decreased when testing was delayed for one week. $d'$ values decreased for non-repeated items (and this was also true for the analysis with 40 participants). For repeated items, the estimate of $d'$ (for N = 33) increased after the delay, but it did not approach significance, $t(32) = 1.09$, $p = .28$ (immediate $M = 1.97$; delay $M = 2.28$).

presented in either the studied or new context for yes/no recognition judgments and, for faces recognized as studied, remember/know judgments.

Consistent with the butcher-on-the-bus example, all current theories of the R/K distinction imply that changing context from study to test should reduce remember responses. In contrast, different theories make divergent predictions regarding the effects of the context manipulation on know responses. If know responses directly reflect the operation of a distinct memory system that stores general knowledge without self-reference (i.e., semantic memory), then changing context from study to test should have no effect on know responses. But if it is assumed that the memory processes that contribute to remember and know responses are independent, and that when both occur participants respond "remember, " then changing context from study to test should increase know responses (to the extent that it lowers remember responses, thereby increasing opportunities for know responses).

The effect of context manipulation on estimates of familiarity is more difficult to predict. If familiarity is assumed to reflect automatic influences of memory for the faces per se, one might expect that familiarity would be invariant across the studied/new context manipulation (because each face is unchanged and should therefore give rise to the same automatic influences of memory regardless of test context). If, however, familiarity for faces can be "contaminated" by parallel automatic influences of memory for context, then familiarity should be greater when context is the same at study and test. Finally, it may be that some memory information cued by a test face would contribute to recollection if the face was presented in its studied context, yet contribute to familiarity if that face was presented in a different context. For example, seeing a face at test might lead a participant to generate a very vague and ambiguous thought or image having to do with the studied context; if the studied context photograph is available at test, participants may use it to interpret the ambiguous information that came to mind as recollection of the study episode, but without it they may experience that information as undifferentiated evidence of familiarity. If so, then estimates of familiarity may increase with changes in context.

Experiments 3 and 4 also included a second memory test, immediately following Test 1, in which studied faces presented in the nonstudied-context condition of Test 1 were presented in their studied contexts. The lures on Test 2 were the same faces as the lures in Test 1, but now presented in a different context. Participants were told that all of the Test-2 faces had been seen in Test 1, and that their task was to discriminate between those that were versus were not seen during the study phase. Restoring context on Test 2 was expected to restore remember and know responding to levels comparable to those observed for studied-context items in Test 1.

*Method*

*Participants and design.* Twenty-five University of Victoria undergraduate students participated for optional extra credit in an introductory psychology course. Data from one participant were removed because of an uncorrected visual impairment. Thus, data from 24 participants were analyzed.

Participants completed the three phases of the experiment in one session. During study, participants encoded each face stimulus item with an accompanying unique context photograph. They used a Likert-type scale to rate the likelihood that the person depicted in the photo was associated with the context (e.g., how likely would this person be associated with this context illustrating pyramids in the desert). Test 1 immediately followed study, and participants were asked to make face recognition judgments followed by remember/know judgments (if applicable) on studied and non-studied face and context pairs. They were informed that some of the studied faces would be presented with a context different than the one accompanying it during study. In addition, they were incorrectly informed that some of the studied faces could be presented with a context not originally studied with that face (i.e., switched, studied context). In fact, in this experiment all faces in the new condition in Test 1 were presented with non-studied contexts. In Test 2, new-context items from Test 1 were re-paired with their original study contexts. Included in Test 2 were lures from Test 1 re-paired with new contexts. Participants were asked to decide whether or not they recognized faces from the study list, followed by remember/know judgments when necessary. The materials and procedure were the same as those in previous experiments except as noted.

*Materials.* The 96 critical items described previously were divided into 4 sets of 24 faces (i.e., A, B, C, D). The sets were reconfigured from previous experiments. As in previous experiments, each set included all age groups and equal numbers of male and female faces for each age group.

Internet searches using keywords such as "photography" and "travel" generated a pool of 196 unique colour photographs that served as contexts for the faces. Context photograph themes included animals, sports, building interiors and exteriors, travel scenery, and vehicles. If a context photograph included people, the faces were always impossible to identify. The types of context themes across item sets were similar. In addition, the number of contexts of a particular theme (e.g., sports) was similar across item sets.

Each critical face was paired with two different contexts to create two context-face files (i.e., *a* and *b* files; 28.35 pixels/cm). The context themes of *a* and *b* versions for each face were distinct: For example, the *a* version of one of the critical faces included a picture of a rollercoaster while the *b* version included a picture of kittens under a chair. The face photographs from previous experiments were not altered. Context photographs were placed to the left of the face. The width of the face was 160 pixels (5.64 cm) and the height was 230 pixels (8.11 cm). The context-face content of the file measured 405 pixels (14.29 cm) in length and 230 pixels in height (and hence the context portion was 245 pixels [8.64 cm] in length). The face and context were centered on a black background that filled the screen (see Figure 7). The *a* version was always used when items were to be presented in the studied context condition at test (and restored studied context condition in Test 2), and *b* version items were always used for the new context condition in Test 1. Because buffer items were the same for all participants, only 4 of the 24 non-critical items required two versions.

There were four item types: studied face/studied context (i.e., *a* files); studied face/ new context (i.e., *b* files); non-studied face/"studied" context (i.e., *a* files); non-studied face/"new" context (i.e., *b* files) (Note that the terms "studied" and "new" context for non-studied faces are arbitrary, and, in fact, all contexts for non-studied faces are non-studied.). Rotation of critical sets (i.e., A, B, C, D) through the 4 item types created 4

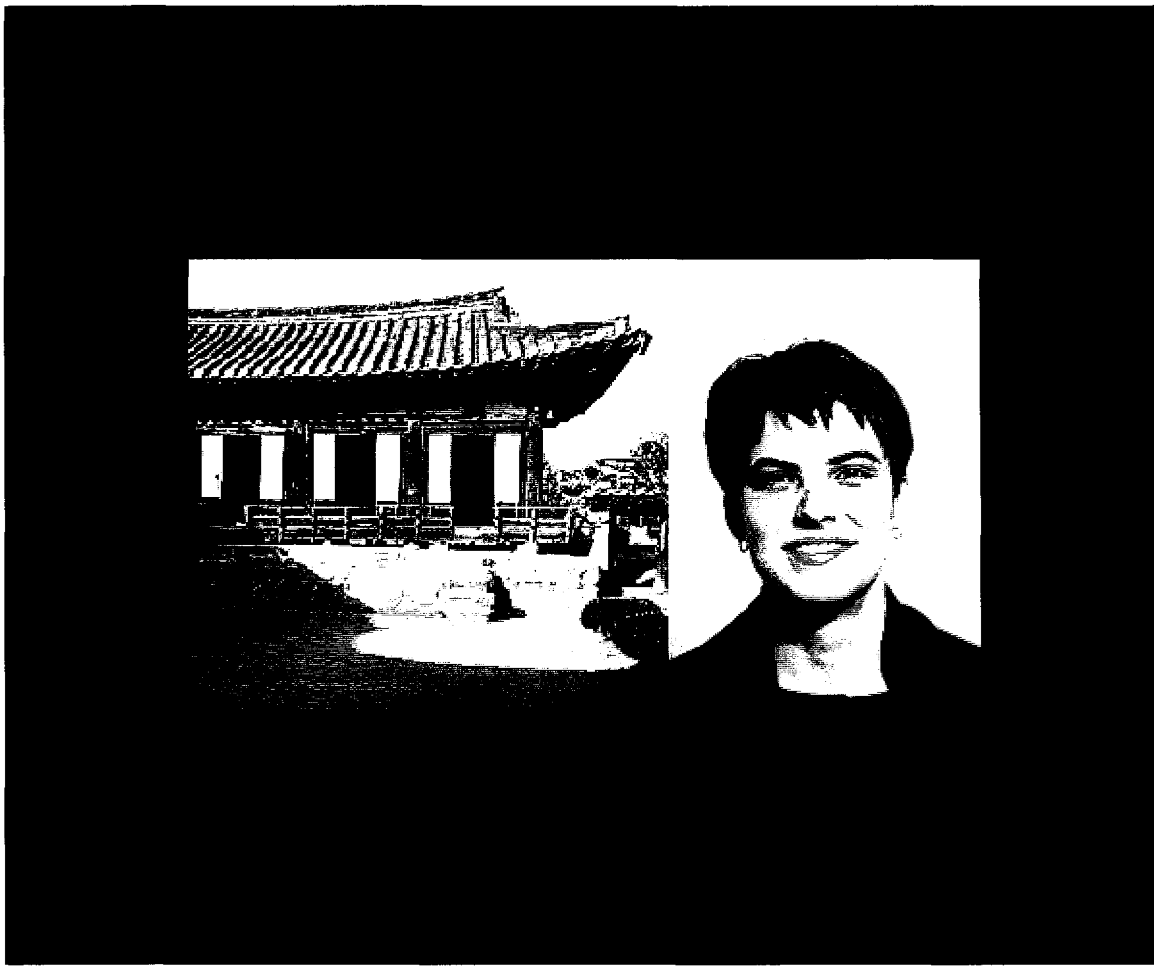


*Figure 7.* Example of stimulus item and presentation used in Experiments 3 and 4 (at test).

different study/test lists. Study lists included 48 critical items (all *a* files, e.g., A*a*, B*a*) of which half were tested in Test 1 with their studied contexts and the other half tested with new contexts (e.g., A*a*, B*b*). Eight buffer items were presented at the beginning and end of each study list. All critical faces (i.e., sets A, B, C, D) were included in Test 1. Half of these faces were old and the remaining were new. Similar to old faces, half of the new faces were presented in *a*-file format representing "studied" contexts and the remaining were presented in *b*-file format representing "new" contexts (e.g., C*a*, D*b*). There were 8 buffer items at the beginning and end of all versions of Test 1 (4 old buffer faces of which two were paired with studied contexts and two were paired with new contexts; 4 new face/context analog pairs).

Test 2 included only the Test 1 faces (i.e., old and new) that were paired with new contexts (i.e., *b* files, e.g., B*b*, D*b*). However, in Test 2, these faces were now re-paired with studied contexts (i.e., *a* files, e.g., B*a*, D*a*). Thus, there were no truly new faces on this test. This test included 48 critical items (24 old face/studied context; 24 new face/"studied" context) and 4 buffer items at the beginning and end of the list (i.e., 2 old face/studied context; 2 new face/"studied" context).

*Procedure.* At study, participants used a 6-point scale to rate the likelihood that the person on the right side of the screen would be associated with the context on the left side of the screen (1 = very unlikely; 2 = unlikely; 3 = slightly unlikely; 4 = slightly likely; 5 = likely; 6 = very likely). The face-context photo remained on the monitor for 2250 ms. The stimulus item was erased from the screen and replaced with a blank screen for 500 ms. After this pause, the encoding question was displayed along with the rating scale and interpretations. Participants recorded their ratings on a sheet of paper. The sheet included 64 number scales corresponding to each studied (critical and non-critical) photograph (e.g., 1. 1---2---3---4---5---6). Participants circled a number on the scale corresponding to their rating of the face-context photo studied. This orienting task was expected to promote deep and integrative encoding of information about each face and context pair. Participants pressed the space-bar to begin the next trial.

There were no filler tasks in this experiment. After the study phase, participants were informed that their memory for the faces they saw earlier would be tested. As in the study phase, each test face was paired with a context photo, but participants were informed that they were to make recognition and remember/know judgments about the faces only. However, they could use context information to support a recognition decision about a face (e.g., "Oh yes, I remember thinking that this person looks like a friend of mine who went to Japan to teach english."). Participants were informed that some face photos were re-presented with the context seen with it at study while others were re-presented with a different context. They were led to believe that "different" meant either a new, non-studied context or a switched, studied context (e.g., Face X was re-paired with a context originally seen with Face Y at study). This deception was intended to discourage participants from assuming that they could base their face-recognition judgments on recognition of the context photo alone. However, all faces presented with a different context were presented with a new, non-studied context. The experimenter recorded test responses (i.e., yes/no recognition followed by remember/know judgments for yes responses only) on the keyboard.

Completion of Test 1 was followed by the third, and final, phase of the experiment: Test 2. Studied faces with new contexts and non-studied faces with "new" contexts from Test 1 (i.e., *b* files for studied and non-studied faces) were presented on Test 2. The studied faces were now re-paired with the original studied context photo and the non-studied faces were re-paired with "studied" contexts (i.e., *a* files for studied and non-studied faces). Participants were informed that they were to make judgments on faces only, regardless of the testing context, although memory for a studied context could be used to make a decision about the face. They were correctly informed that all of the faces on Test 2 had been presented on Test 1. Thus, their task was not to indicate whether or not the face was presented in Test 1; rather, they were to decide whether or not they recognized the face from the study phase of the experiment. They were reminded that during that phase they had made likelihood ratings and that they recorded their ratings on a sheet of paper. If they recognized a face from the study session, then they provided a remember or know response.

*Results*

*Lures.* Remember and know response probabilities to lures in Test 1 are presented in Tables 9 and 10, respectively. As expected, false alarm responding was not affected by the arbitrarily-assigned context type (i.e., "studied" or "new"; lures on the first test consisted of non-studied faces with never-seen contexts) on the first test for either type of responding: $t < 1$ for remember responses, and $t(23) = 1.67, p > .10, r^2 = .11$, for know responses. Thus, false alarm probabilities were collapsed over item type for each response type (i.e., remember false alarm $M = .03$; know false alarm $M = .12$). Although new faces

Table 9

*Mean Remember Response Probabilities in Experiment 3*

| | Test 1 | | Test 2 | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Studied Context | | | | |
| Targets | .62 | .04 | | |
| Lures | .03 | .01 | | |
| (H - FA) | .59 | .04 | | |
| (H - FA)/(1 - FA) | .61 | .04 | | |
| New Context | | | | |
| Targets | .28 | .04 | .66 | .04 |
| Lures | .03 | .01 | .09 | .01 |
| (H - FA) | .25 | .03 | | |
| (H - FA)/(1 - FA) | .26 | .03 | | |

*Note.* *M* = mean; *SEM* = standard error of the mean.

Table 10

*Mean Know Response Probabilities in Experiment 3*

| | Test 1 | | Test 2 | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Studied Context | | | | |
| Targets | .22 | .02 | | |
| Lures | .11 | .02 | | |
| (H - FA) | .12 | .02 | | |
| New Context | | | | |
| Targets | .36 | .03 | .17 | .03 |
| Lures | .14 | .02 | .17 | .02 |
| (H - FA) | .24 | .03 | | |

*Note.* *M* = mean; *SEM* = standard error of the mean; (H - FA) = hits minus false alarms.

on Test 2 were not studied in the first phase of the experiment, they were presented on Test 1. Hence incorrect responding was expected to be greater on the second test rather than the first for these items. This was true for remember responses, $t(23) = 4.30$, $p < .001$, $r^2 = .46$. Re-testing lures from Test 1 also directionally increased know responding, but the change was not significant, $t(23) = 1.29$, $p = .21$, $r^2 = .067$. Comparisons between Test 1 and Test 2 items were not the primary purpose of this experiment. However, it was of some interest to evaluate the effect of restoring context. Because of the lack of true control items on Test 2, any change in response criterion across tests could not be measured. Thus, comparisons between the two tests were made with uncorrected probabilities. Consequently, I acknowledge that any results across tests may be biased.

*Remember responses.* In Test 1, uncorrected probabilities showed that faces tested with the studied context were better remembered than faces tested with new contexts, $t(23) = 9.59$, $p < .001$, $r^2 = .80$ (see Table 9). As expected, the hits-minus-false alarm method and conditionalized responses produced identical results, $t(23) = 9.59$, $p < .001$, $r^2 = .80$, $t(23) = 9.69$, $p < .001$, $r^2 = .80$, respectively. When the new context in Test 1 was replaced with the original studied context on the second test, remember responses were found to increase to a level previously obtained for studied context items on Test 1, $t(23) = 1.11$ $p = .28$, $r^2 = .051$.

*Know Responses.* Know responses varied for the two context types in a direction opposite to that for remember responses (see Table 10). In the first test, participants were more likely to provide know responses to new context items than to studied context items, $t(23) = 4.30$, $p < .001$, $r^2 = .45$. Correcting responses with false alarms also revealed an increase in know probabilities in Test 1, $t(23) = 3.74$, $p < .005$, $r^2 = .38$. In Test 2, replacing a new context with a studied context reduced know responding, $t(23) = 2.19$, $p < .04$, $r^2 = .17$.

*IRK Familiarity.* Altering context between study and test non-significantly reduced familiarity-based recognition, $t(23) = 1.62$, $p = .12$, $r^2 = .10$ (see Table 11). Familiarity estimates corrected for new items also revealed a non-reliable reduction between studied and new context items, $t(23) = 1.73$, $p = .097$, $r^2 = .12$. Items from Test 2 (i.e., Test 1 faces with new contexts now re-paired with studied contexts) were significantly less familiar than same context items of Test 1, $t(23) = 2.21$, $p < .04$, $r^2 = .17$.

*d' Estimates of Familiarity.* Because false alarm responding was not available for Test 2 items, *d'* values were estimated for Test 1 items only (see Table 11). Similar to corrected IRK familiarity estimates, the result of examination of *d'* estimates also indicated a non-significant decrease in this memory process, $t(23) = 1.53$, $p = .14$, $r^2 = .092$.

*Sensitivity and Bias.* As noted above for *d'* estimates, *A'* values were estimated for Test 1 items only (see Table 12). In Test 1, memory measured with overall recognition responding was greater than that obtained with remember responding for studied, $t(23) = 2.15$, $p < .05$, $r^2 = .17$, and new contexts, $t(23) = 4.70$, $p < .001$, $r^2 = .49$. Correlations between sensitivity and bias were positive and significant for all context types (studied $r = .63$, $p < .001$; new $r = .41$, $p < .05$). More important was the finding that rescaling know responses for each context type substantially reduced and statistically eliminated these correlations (studied $r = -.12$, $p = .58$; new $r = -.10$, $p = .63$). Thus, Donaldson's (1996) first prediction about measures of sensitivity for the single-process

Table 11

*Mean IRK and d' Estimates of Familiarity in Experiment 3*

| | Test 1 | | Test 2 | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Studied Context | | | | |
| Targets | .61 | .05 | | |
| Lures | .13 | .02 | | |
| (H - FA) | .48 | .05 | | |
| *d'* | 1.59 | .18 | | |
| New Context | | | | |
| Targets | .52 | .04 | .47 | .04 |
| Lures | .13 | .02 | | |
| (H - FA) | .39 | .04 | | |
| *d'* | 1.34 | .13 | | |

*Note.* *M* = mean; *SEM* = standard error of the mean; IRK = independence remember/know; (H - FA) = hits minus false alarms.

Table 12

*A' and $B''_D$ Estimates for Test 1 in Experiment 3*

| | *A'* | | *$B''_D$* | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Studied Context | | | | |
| Recognition | .91 | .01 | .01 | .12 |
| Remember | .89 | .01 | | |
| Know | .62 | .04 | | |
| Know-IRK | .82 | .03 | | |
| New Context | | | | |
| Recognition | .84 | .01 | .39 | .11 |
| Remember | .77 | .01 | | |
| Know | .72 | .02 | | |
| Know-IRK | .80 | .02 | | |

*Note.* *M* = Mean; *SEM* = standard error of the mean; IRK = independence remember/know, K/(1 - R).

model (i.e., $A$'recognition = $A$'remember) was not supported by these results. The second prediction about a positive correlation between $A$'recognition and $B''_D$ know was supported. However, the significant attenuation of the correlation when remember responding was removed (i.e., know responses were adjusted with the IRK equation) also supports the view that familiarity without recollection can be successfully modeled with signal detection theory.

*Discussion*

As predicted, changing context from study to test dramatically lowered the rate of remember responses. Changing context also substantially increased the rate of know responses. If know responding is taken as a direct estimate of the operation of a distinct memory system as Gardiner and others have proposed (e.g., Gardiner & Parkin, 1990), then this result suggests that this system paradoxically benefits from altered contextual cues. A more plausible explanation is that participants only respond "know" when they do not experience recollection (as argued by Jacoby and others, e.g., Jacoby et al., 1997, Yonelinas et al., 1995).

The effects of the context manipulation on estimates of familiarity were somewhat ambiguous. The pattern of statistical significance are consistent with Jacoby's (1991, 1998) and Yonelinas' (1994) notion of independent dual processes: Despite the large effect on remembering (i.e., recollection), the context manipulation did not reliably alter familiarity (measured as IRK or $d'$). There was a non-reliable tendency for familiarity to decrease when context changed, suggesting some "contamination" of familiarity for faces by automatic influences of memory for context (alternatively, they may "automatically" reject faces as new when they are accompanied by a new context). The findings did not support the idea that memory information that would contribute to remembering in the studied-context condition might instead contribute to familiarity in the new-context condition. This does not mean that the proposed phenomenon did not occur, but rather that if it did occur its effects were more than offset by the opposing tendency for the studied context to contribute to estimates of familiarity.

The data from Test 2 must be interpreted with caution, because the lures on that test had also been presented (albeit with different context pictures) in Test 1, obscuring interpretation of false alarms and making it impossible to use false-alarm data to calculate various measures of familiarity. Nonetheless, the central finding is strikingly clear: Restoring study context in Test 2 dramatically increased the rate of remembering responses. Indeed, remember responses were as frequent for Test-2 restored items as they were for Test-1 studied-context items. Contrary to the earlier prediction, know responses to re-paired Test-2 items were lower than that for Test-1 studied-context items.[15] It should be noted that overall recognition of Test-2 items ($M = .83$) was restored to Test-1 same-context levels ($M = .84$), $t < 1$. Together with the results for remember and know responses, these results suggest that restoring study context led participants to either fully recognize faces and say "remember" or to most likely reject the face rather than recognize it and say "know." It is likely that the experience of viewing Test-2 studied-context items

---

15

Accurate knowledge of false alarm rate would only reduce this value further. However, it is also possible that inclusion of appropriate lures would have altered the phenomenal experience of recognition for all items on the second test.

was very different than for Test-1 studied-context items because of the intervening inconsistent pair for the former. As suggested by Whittlesea (2002), the new-context pairs may have evoked perceptions of incongruity (i.e., "wrongness: some aspect does not fit with others in a identifiable way", p. 326) or discrepancy (i.e., "strangeness: some aspects of processing mismatch in an indefinite way", p. 326) when information about the face was recalled. For most restored pairs in Test 2 the presence of the appropriate studied context may have evoked an overwhelming feeling of congruity (i.e., "well-formedness: all salient aspects seem to fit well together," p. 326) leading participants to say "remember." For other items the intervening phenomenal experience with the new context may have interfered with memory for the face so much so that when re-presented on Test 2 (albeit with the studied context) there may not have been sufficient memory associated with the face to recognize it. In keeping with this notion, the estimate of familiarity for restored pairs on Test 2 was also reduced. This suggests that participants' decisions about faces on Test 2 were primarily based on recollection processes. It appears that the new context may have interfered with memory processing of some faces and upon re-presentation there was not enough memory available or accessible to support even a nagging feeling of familiarity.

## Experiment 4

Participants in Experiment 3 were incorrectly informed that a context from a studied face/context pair could be recycled and re-paired with another face on the test list. Although the observed difference between studied- and new-context familiarity estimates was not statistically significant, the decrease suggested that participants may have become aware of the misleading information and used the new context as a basis for rejecting the face. In addition to the two context conditions of Experiment 3, the final experiment included switched context items.

The loss of familiarity for new-context items in Experiment 3 may have been due to the non-familiarity of the context (and subsequent rejection of the studied face, i.e., F-to-New migration) which more than offset any tendency for R-to-F migration. Given this possibility, I predicted that the inclusion of truly switched contexts in Experiment 4 would make it more difficult to reject a face solely because the context was different. Consequently, studied faces with switched contexts should have more familiarity associated with them than studied faces with studied contexts. In addition, there would be no change in familiarity between studied-context items and new-context items.

*Method*

*Participants and design.* Twenty-seven University of Victoria undergraduate students participated for optional extra credit in an introductory psychology course. Data from 3 participants were removed: two participants never responded Remember and one participant never responded Know. Thus, data from 24 participants were analyzed.

Similar to Experiment 3, participants in this experiment completed three phases in one session. During study, participants provided ratings of the likelihood of association between the person and the context shown on the screen. Test 1 included studied and non-studied faces each with either a studied or new context. In addition, there were truly switched contexts: Some studied and non-studied faces were paired with contexts originally presented in phase 1 with a different face. Participants were aware of the different item types. Their task was to indicate whether or not they recognized the face from the study phase followed by a remember or know judgment if they did recognize it. For Test 2, studied faces that had been presented with a new or switched context in Test 1 were re-paired with the original studied context; intermixed with the same lures as in Test 1 but with rearranged contexts. Again, participants made recognition judgments and remember or know judgments when applicable. The materials and procedure were the same as those in Experiment 3 except as noted.

*Materials.* Critical item sets were reconfigured to form 6 sets of 16 faces (i.e., F1, F2, F3, F4, F5, F6). All sets included each age group and equal numbers of males and females. Each face was paired with a unique context photo obtained from Experiment 3. Context photos formed the sets C1, C2, C3, C4, C5, and C6. Context themes were the same for all sets and balanced across sets. Buffer items were identical to those in the previous experiment.

Files for the study pairs were similar to those of Experiment 3 with the following modifications. A small black border was added to the outer perimeter of the face-context pair with the final measurement for the pair and border equal to 415 pixels (14.64 cm) in length and 240 pixels (8.47 cm) in height. The pair-with-border figure was centered on a white background. These changes were expected to provide additional distinctive information that participants could use, during Test 2, to differentiate memories of the study phase from memories of Test 1. As in Experiment 3, test pairs did not have a border and were centered on a black background.

Appendix B provides a description of the different conditions for Experiment 4.

Original face-context pairs were designated as F1C1, F2C2, and so forth. These pairs included studied face/studied context item types and non-studied face/new context item types. The remaining item types were formed by rearranging original pairs (e.g., F1C2, F1C3, etc.). However, not every possible combination was formed. In addition to the original context, each face was paired with four other contexts (e.g., for F1 the following pairs were formed, F1C1, F1C2, F1C3, F1C5, F1C6). This was sufficient to form the 6 study/test lists that allowed for each face to be tested as each item type once. In all, there were 6 item types: studied face/studied context; studied face/switched context; studied face/new context; non-studied face/studied context; non-studied face/"switched" context; non-studied face/new context. The "switched" designation for non-studied faces provided information about the counterbalancing of context item sets only. In essence, "switched"contexts and new contexts (for non-studied faces) were both never-seen contexts as far as the participant was concerned.

Study lists included 48 critical face-context pairs and 8 buffer items at the beginning and end of each study list. All 96 critical faces were presented on Test 1. This test was also buffered by 8 non-critical items at the beginning and end of the list (i.e., two from the first 8 study-list buffers; two from the last 8 study-list buffers; and, 4 new non-critical items). For Test 2, re-arranged Test 1 pairs were returned to their original form. There were 64 critical items and 4 buffer items at the beginning and end of this test (i.e., one old and one new non-critical item from the first 8 Test 1-list buffers; and, one old and one new non-critical item from the last 8 Test 1-list buffers).

Lures on the first test were accompanied by a context that was either previously studied or new. On the second test, the Test 1-studied-context lure was now accompanied by a new context. However, this new context was previously presented on the first test with a non-studied face. In addition, the remaining new context of Test 2 had, on Test 1, been placed beside a studied face (i.e., new-context condition). Consequently, all new contexts presented with non-studied faces in Test 2 were viewed earlier on Test 1. As in Experiment 3, there were no true lures on the second test.

*Procedure.* The procedure for Experiment 4 was similar to that for Experiment 3 except that participants were also reminded at each test that faces at study were presented with contexts on a white background that filled the screen. In addition, on the first test, they were correctly informed that some of the faces would be presented with switched contexts.

For the second test, switched context items (i.e., really switched or arbitrarily assigned for new faces) and new context items from Test 1 were re-paired with original contexts. Participants were told that all faces on Test 2 had been presented earlier either at study, or on Test 1, or both. They were asked to identify only those faces from the earlier study session. Upon recognition of a face, they provided either a remember or know judgment.

*Results*

The Context factor in the summary tables for Experiment 4 (Tables 13-16) refers to the context type on the first test. Although switched and new contexts were replaced with studied contexts on the second test, response rates or parameter estimates for the second test are described by the context types viewed on the first test.

To counterbalance items in Experiment 4, two sets of lures with different contexts were created for the first test: One set consisted of the original pairings of faces and contexts and the other was created by rearranging original pairs (non-studied face, "switched" context; non-studied face, new context; see Appendix B). However, from the perspective of the participant these sets were identical and should provide the same false alarm rate. Analysis of remember and know responses supported this prediction, all $t$s < 1. For this reason, one average rate for each response type was derived from the two sets and used in the following analyses (remember $M = .03$; know $M = .13$). Response probabilities for lures with new contexts were used to correct responding to critical items with studied, switched, or new contexts.

The effect of context type on remember and know responding to target faces in Test 1 was analyzed with separate one-way repeated measures ANOVAs with 3 levels (i.e., studied, switched, new). Mean uncorrected probabilities or estimates between Test 2-switched items and Test 1-studied items (as well as mean uncorrected probabilities or estimates between Test 2-new items and Test 1-studied items) were compared to examine the effect of restoring a studied context.

*Lures.* To measure the effects of context on new faces separate 2 (context: studied, new[16]) X 2 (test: Test 1, Test 2) repeated measures ANOVA analyzed remember and know responses. There were no appreciable changes in know responses to lures, all $p$s > .13 (see Table 14). In contrast, remember responses to lures showed significant

---

16

As indicated earlier, the rate for new context lures was an average of non-studied face/new context items and non-studied face/"switched" context items.

variability in the form of an interaction, $F(1, 23) = 4.81$, $MSE = .003$, $p < .04$, $\eta^2 = .17$ (see Table 13). In Test 1, remember responding to lures with studied contexts was greater than that for lures with new contexts, $t(23) = 3.27$, $p < .005$, $r^2 = .32$.[17] In addition, re-presenting a Test 1-new-context lure led to an increased probability of false alarms in Test 2, $t(23) = 2.22$, $p < .04$, $r^2 = .18$. All other contrasts of interest were not significant, all $ps > .38$.

*Remember Responses.* A significant effect of context type for Test 1 critical items (i.e., (H - FA)/(1 - FA), see Table 13), $F(2, 46) = 93.86$, $MSE = .010$, $p < .001$, $\eta^2 = .80$, was probed and revealed significant decreases in remember responding between studied and switched contexts, $t(23) = 8.11$, $p < .001$, $r^2 = .74$, and between switched and new contexts, $t(23) = 5.88$, $p < .001$, $r^2 = .60$.

Results of analyses for uncorrected and (H - FA) response probabilities were identical. In addition, they were consistent with results for conditionalized responses presented above (see Table C8 in Appendix C).

Replacement of a switched context on Test 2 with a studied context improved remember responding such that no significant difference was detected between this value and that obtained for studied context items on the first test, $t < 1$. In addition, responding to restored new context items on Test 2 was not reliably different than responding to Test

---

17 Despite never having seen a face, simply placing a studied context beside it led to more remember responding than when a "switched" (i.e., new) context was present. To adjust for this influence, one may use corrections made with response probabilities for studied context/new faces. Analyses with such corrected probabilities revealed the same trends and results.

Table 13

*Mean Remember Response Probabilities in Experiment 4*

| | Test 1 | | Test 2 | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Studied Context | | | | |
| Targets | .61 | .03 | | |
| Lures | .08 | .02 | .07 | .02 |
| (H - FA) | .58 | .03 | | |
| (H - FA)/(1 - FA) | .60 | .03 | | |
| Switched Context | | | | |
| Targets | .37 | .04 | .58 | .04 |
| Lures | .03 | .01 | .07 | .02 |
| (H - FA) | .34 | .03 | | |
| (H - FA)/(1 - FA) | .35 | .03 | | |
| New Context | | | | |
| Targets | .23 | .03 | .56 | .04 |
| Lures | .03 | .01 | | |
| (H - FA) | .20 | .03 | | |
| (H - FA)/(1 - FA) | .21 | .03 | | |

*Note.* *M* = mean; *SEM* = standard error of the mean; H = hits; FA = false alarms.

Table 14

*Mean Know Response Probabilities in Experiment 4*

| | Test 1 | | Test 2 | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Studied Context | | | | |
| Targets | .22 | .02 | | |
| Lures | .18 | .03 | .16 | .02 |
| (H - FA) | .12 | .03 | | |
| Switched Context | | | | |
| Targets | .38 | .03 | .26 | .03 |
| Lures | .14 | .03 | .15 | .02 |
| (H - FA) | .25 | .03 | | |
| New Context | | | | |
| Targets | .38 | .03 | .20 | .03 |
| Lures | .12 | .02 | | |
| (H - FA) | .25 | .03 | | |

*Note.* *M* = mean; *SEM* = standard error of the mean.

1 studied context items, $t(23) = 1.51, p = .15, r^2 = .090$.

*Know responses.* Context type significantly affected know responding to items (i.e., (H - FA), see Table 14) on the first test, $F(2, 46) = 10.83, MSE = .013, p < .001, \eta^2 = .32$. There was a reliable increase in response probability between studied items and switched items, $t(23) = 4.40, p < .001, r^2 = .46$. However, there was no change in responding between switched and new items, $t < 1$. Results for uncorrected know responses did not vary from those corrected for false alarms (see Table C9 in Appendix C).

On Test 2, participants were just as likely to respond to restored switched items and to restored new items as they were to studied context items of Test 1, $t(23) = 1.48, p = .15, r^2 = .087$ and $t < 1$, respectively.

*IRK Familiarity.* A planned comparison tested the prediction that familiarity would increase from the studied-context condition to the switched-context condition. As shown in Table 15, the difference was in the predicted direction, but it was not statistically reliable, $t < 1$ ($t < 1$ for uncorrected values). Consistent with the prediction that familiarity would not differ between studied-context faces and new-context faces, a planned comparison between those two conditions yielded a null effect, $t(23) = 1.23, p = .23, r^2 = .062$ ($t(23) = 1.14, p = .27, r^2 = .053$, for uncorrected values).

Familiarity estimates were also unaffected by changing a switched context for a studied context on Test 2, $t(23) = 1.41, p = .17, r^2 = .080$. Similarly, the change to a studied context from a new context did not reliably decrease the estimate of familiarity, $t(23) = 1.77, p = .09, r^2 = .12$.

Table 15

*Mean IRK and d' Estimates of Familiarity in Experiment 4*

| | Test 1 | | Test 2 | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Studied Context | | | | |
| Targets | .56 | .04 | | |
| (H - FA) | .42 | .04 | | |
| *d'* | 1.35 | .15 | | |
| Switched Context | | | | |
| Targets | .60 | .04 | .60 | .05 |
| (H - FA) | .46 | .05 | | |
| *d'* | 1.49 | .17 | | |
| New Context | | | | |
| Targets | .50 | .03 | .46 | .05 |
| Lures | .14 | .02 | | |
| (H - FA) | .36 | .04 | | |
| *d'* | 1.19 | .14 | | |

*Note.* IRK = independence remember/know; *M* = mean; *SEM* = standard error of the mean; (H - FA) = hits minus false alarms.

*d' Estimates of Familiarity.* As with IRK estimates of familiarity, *d'* values showed an increase from studied contexts to switched contexts (see Table 15). A planned comparison between these two values also proved the increase to be unreliable, $t(23) = 1.03, p = .32, r^2 = .044$. The predicted null effect between studied- and new-context items was supported, $t(23) = 1.13, p = .27, r^2 = .052$.

*Sensitivity and Bias.* *A'* and $B''_D$ values for Experiment 4 are presented in Table 16. In the first test, recognition and remember *A'* values were similar for studied-context items, $t(23) = 1.39, p = .18, r^2 = .078$, but not for switched-context items, $t(23) = 4.43, p < .001, r^2 = .46$, nor for new-context items, $t(23) = 2.43, p < .03, r^2 = .20$. As in previous experiments *A'* values for overall recognition and remembering were not always equivalent, suggesting the presence of more than one recognition process. The studied-context condition included the least amount of know responding (and the most remember responding) implying that contributions from processes associated with know responding would be minimal in this condition. Thus, it makes sense that when additional processing is minimized overall recognition more likely reflects a single process rather than dual processing. Consequently, *A'*remember and *A'*recognition values would be similar.

Correlations between $B''_D$ recognition and *A'*know values for studied ($r = .71, p < .001$) and new ($r = .52, p < .01$) contexts of Test 1 were positive and significant. There was also a tendency toward a positive correlation between sensitivity and bias for switched-context items but it was not reliable ($r = .31, p = .14$). When the measure of sensitivity was based on know responses in the absence of remember response (i.e., K / (1 - R)), the correlation associated with each context type was reduced and non-significant

Table 16

*A' and B"$_D$ Estimates for Test 1 in Experiment 4*

| | *A'* | | *B"$_D$* | |
|---|---|---|---|---|
| | *M* | *SEM* | *M* | *SEM* |
| Studied Context | | | | |
| Recognition | .90 | .01 | .11 | .10 |
| Remember | .88 | .01 | | |
| Know | .60 | .04 | | |
| Know-IRK | .80 | .02 | | |
| Switched Context | | | | |
| Recognition | .87 | .02 | .27 | .10 |
| Remember | .81 | .01 | | |
| Know | .71 | .03 | | |
| Know-IRK | .81 | .03 | | |
| New Context | | | | |
| Recognition | .81 | .02 | .49 | .08 |
| Remember | .75 | .01 | | |
| Know | .72 | .02 | | |
| Know-IRK | .78 | .02 | | |

*Note.* *M* = Mean; *SEM* = standard error of the mean; IRK = independence remember/know, K/(1 - R).

(studied $r = .15, p = .48$; switched $r = -.08, p = .71$; new $r = .18, p = .40$). Thus, the data pattern that would support the single-process model of recognition memory was not always present. In contrast, when the effects of one of the processes in a dual-process model of recognition were removed then measures appeared to be based on a remaining single process and the predictions of a signal detection model were confirmed (i.e., no correlation between sensitivity and response criterion).

*Discussion*

In summary, as in the previous experiment, altering context at test reduced remember responding and hence the contribution of recollection processes to recognition. Furthermore, there was a graded decrease in recollection with the greatest reduction associated with new-context items. Know responses revealed a different pattern of results: Participants were more likely to say know to changed context items regardless of the type of change. The results for remember and know responding were not affected when corrections for any changes in response criterion were applied. As predicted, switched-context items had greater IRK, or $d'$, estimates of familiarity than studied-context items. However, the change was not reliably significant. Consistent with earlier findings there was no change in familiarity (either IRK or $d'$) between studied-context items and new-context items.

Experiment 4 revealed a process dissociation with context affecting recollection but not familiarity. The results of this experiment provide support for a dual-process model of recognition memory rather than a single-process model. The increase in know responding for changed context items does not pose a problem for the functional model of memory that proposes a single process with independence among memory attributes: Information that in a studied-context condition supported clear remembering or recollection of the face would in a different context condition only contribute to an unspecified sense of oldness or knowing. However, if a dual-process model is considered with an exclusive relationship between processes, then it would seem odd that a studied face in the presence of a different context would have more familiarity associated with it than when in the presence of the studied context. As the independence model predicted, know responding increased for changed context items, but when these probabilities were translated into IRK familiarity estimates or *d'* values there was no parallel increase in familiarity. Such a result is consistent with a dual-process model with independent processes. Furthermore, the best fit among a number of different multinomial models tested also provided support for independent processes in recognition memory. The models are presented in Appendix D.

Re-pairing faces with studied contexts on Test 2 fully restored all responding and memory processes to levels obtained with faces with studied contexts of Test 1. This was only true for remember responding in Experiment 3: Know responding and familiarity were reduced to levels below those obtained for same context items. In Experiment 4, any interference due to changed contexts appears to have been overcome when contexts were restored. It appears that inclusion of items paired with switched-studied contexts altered the subjective experience of each item type. To better assess the effects of restoring context, further studies should include studied-context items and truly new items in addition to restored context items on a second test.

Bodner and Lindsay (in press) studied the effects of context on remembering and knowing of words. They found that the phenomenological experience of stimulus items was dependent upon test context. Specifically, they found that responding to words at test that were studied with a medium level-of-processing (LOP) task differed depending upon the encoding history of the other words. Thus, when tested in the presence of shallow LOP words (i.e., medium-with-shallow condition), participants were more likely to remember than know medium LOP words. In contrast, in the presence of deep LOP words (i.e., medium-with-deep condition), they were more likely to say know rather than remember to recognized medium LOP words. These changes in the subjective experience of items occurred despite the lack of significant differences in overall recognition for medium-LOP items in the two conditions. In addition, IRK estimates of familiarity for medium LOP items did not differ in the two context conditions. Because overall recognition and familiarity contributions to responding was the same for both conditions, it would be incorrect to say that the decrease in remember responding truly reflects disruption of recollection processing. Rather, Bodner and Lindsay suggested that response differences were due to changes in the subjective experience of recognition associated with awareness of memorial attributes. Thus, in the presence of shallow items retrieved information about a medium item is sufficient to identify its origins in the study session and promote "clear" remembering. However, in the presence of deep items the same retrieved information may only inspire a feeling of oldness and know responding. The recall of highly-specifying information for deep items may compel a phenomenological experience very different from that associated with the recall of information for medium items that may not be as detailed but is equally source-specifying. Consequently, this may have prompted participants to say know rather than remember.

In this experiment, the type of context presented at test affected overall recognition. Furthermore, there was a significant decrease in recognition between studied-context items and switched-context items, and between switched-context items and new-context items (means and results of analyses are available in Table C10 in Appendix C). As in Experiment 3, recollection processes in this experiment were negatively affected by a change in context at test. The reduction in overall recognition may be due solely to changes in recollection with familiarity processes unaffected by the context manipulation. Alternatively, it is also possible that the decrease in overall recognition is a result of a greater reduction in recollection than an increase in familiarity for a context change. The current experiment supports the notion of independent processes as described by proponents of dual-process theories of recognition (e.g., Jacoby, 1998; Yonelinas, 1994) rather than independent attributes associated with functional definitions of phenomenological experiences that vary with the purpose of the task (Bodner & Lindsay, in press; Gruppuso et al., 1997).

Perhaps differences between Experiment 4 (and Experiment 3) and the Bodner and Lindsay experiments are due to the type of stimulus items. Words typically have some pre-experimental familiarity associated with them and this may promote the availability of memorial information to be recalled. In contrast, participants had no pre-experimental familiarity for the faces and contexts used in this experiment. True, they may have experienced similar context scenarios but it would be unlikely that they had prior knowledge of the specific context photographs. The study task emphasized the association between a face and a context, and there may not have been sufficient encouragement to encode face-without-context information and context-without-face information (i.e., the processing of individual sub-components of the stimulus item). Thus, when pairs were disrupted, it would follow that this would interfere with much of the memory attributes available to be recalled. Consequently, it would be difficult to observe changes in responding that would support the notion of functional definitions of recollection and familiarity.

Some additional insight into the effect of context manipulation on items with no pre-experimental familiarity can be provided by the research of Macken (2002). The purpose of his research was to assess the effect of changing environmental context (i.e., colour of item on screen, location of item on screen, and background screen colour) on estimates of recollection and familiarity processes in recognition. Macken found that changing context reduced the contribution of recollection processes but that the contribution of familiarity processes to word recognition was left unchanged. He also showed that stimulus items without any pre-experimental familiarity (i.e., nonwords: consonant-vowel-consonant trigrams) produced identical effects. Regardless of the familiarity of stimulus items in the Macken experiments the pattern of results for know response probabilities was similar to that found in Experiments 3 and 4: Know responding increased for changed context items.

Future experiments can perhaps look at unfamiliar faces in familiar contexts. This situation may be a more comparable laboratory model of an eyewitness event. Typically, an individual witnessing an event would be in a location that has some familiarity to her. Mandler's (1980) example suggests that when you see the butcher on the bus you are likely to say that you know him rather than remember him. While you don't regularly see the butcher on the bus, you may regularly be on the bus. Thus, you have some familiarity with this location. In the laboratory, participants can encode faces with accompanying context photographs of unique campus locations. Pre-experimental familiarity with particular contexts may provide a foundation upon which to encode associative information and specific sub-component or item information. Disruption of pairs may scramble associative information, but may not be extremely detrimental to the availability of face-only information required to complete the goal of the recognition memory task. Alternatively, it is possible to re-introduce the repetition manipulation to increase the encoding of information about the pairs and their parts. This situation would not be an analogue to the eyewitness event because these events are typically unique.

## Summary

In Experiment 1, participants studied faces with a shallow encoding task, with some faces presented once and others presented three times, and subsequently made recognition and remember/know judgments on an immediate or 3-day delayed test. The delay reduced remember responses for both repeated and non-repeated faces. For non-repeated faces, the delay also decreased know responses and estimates of familiarity. For repeated faces, in contrast, delay had no reliable effect on know responses or estimates of familiarity. On the one hand, these results did not provide direct evidence of the predicted tendency for memory influences that contributed to remember responses on the immediate test to contribute to know responses on the delayed test (that is, familiarity did not increase with delay). On the other hand, the pattern of results is consistent with the ideas that (a) there was indeed some "migration" from recollection to familiarity across delay, but (b) that effect was offset (for repeated items) or more than offset (for non-repeated items) by the concurrent loss of memory information across delay (i.e., familiarity to new).

Experiment 2 was similar to Experiment 1, but used a deeper encoding task and a 1-week delay. As expected, reports of remembering declined with delay. For repeated items, know responses increased with delay. Also, on the immediate test, know responses were greater for non-repeated than for repeated faces. Thus, if an exclusive relationship between the processes that support know and remember responses is assumed, such that know responses are viewed as a direct index of familiarity (as implied by Gardiner and colleagues), one would conclude that familiarity decreased with repetition and increased with delay.

When an independent relationship is assumed, with remember responses dominating over know responses, the rate of know responses is not taken as a direct index of familiarity. In Experiment 2, estimates of familiarity obtained with the IRK equation or $d'$ produced more theoretically sensible results than did the raw rate of know responses. Specifically, the estimate of familiarity for repeated items was greater than that for non-repeated items. Delay reduced the estimate of familiarity for both repeated and non-repeated items. The increase in know responding after the delay provides some support for "migration" from R to F, but the overall decrease in familiarity after the delay suggested that the "migration" from F to New was greater.

The results of Experiments 1 and 2 do not refute the idea that memory information can contribute to either recollection or familiarity processes depending upon the demands of the task (i.e., the "migration" of R to F), but nor do they offer compelling support for it. There may be a constellation of study and delay conditions that would yield more convincing evidence of the hypothesized migration of memory influences from R to F, but I was unable to arrive at a principled way of selecting those conditions. Experiments 3 and 4 therefore took a different tack to exploring these issues.

In Experiment 3, participants studied faces paired with distinctive contexts, and later made recognition and remember/know judgments to test faces that were presented with their studied context or with a new context. Experiment 4 was similar, except that at test some studied faces were presented in a context seen in study with a different face (i.e., switched context). Similar to the effect of delay in Experiments 1 and 2, changing context at test in Experiments 3 and 4 reduced remember responses. In addition, Experiment 4 revealed a graded decrease in remember responses, such that target faces accompanied by a new (non-studied) context were least often remembered, those accompanied by a switched context were somewhat more often remembered, and those accompanied by their studied contexts were most often remembered.

In Experiments 3 and 4, Know responses were less frequent for faces presented in their studied contexts than for those presented in new or switched contexts. As indicated for the first two experiments, these results pose problems for dual-process models that posit an exclusive relationship between familiarity and recollection, with know responses taken as a direct index of familiarity. In Experiment 3, there was a non-reliable tendency for IRK or *d'* familiarity to be decreased with the context manipulation at test suggesting that memory for context may have "contaminated" estimates for familiarity. Thus, any R-to-F migration was offset by F-to-New migration. In Experiment 4, the inclusion of different but familiar (i.e., switched) contexts was expected to lessen such contamination. There was a predicted increase in familiarity for switched-context items as compared to studied-context items, but the increase was non-significant. In addition, and also as predicted, there was no change in the estimate of familiarity between studied- and new-context items.

Experiments 3 and 4 also included a second test phase, in which faces presented in Test 1 with new or switched contexts were presented with their studied contexts. Restoring the studied context led to rates of remembering and knowing that were equivalent to those observed for studied faces presented in their studied contexts in Test 1.

Across 10 conditions in Experiments 1 and 2, all *A'* recognition estimates were directionally greater than *A'* remember estimates, with 8 of those differences being statistically significant. Similarly, across five comparisons in Experiments 3 and 4 *A'* values for recognition were directionally greater than *A'* values for remember responses, with four of those comparisons being statistically significant. According to Donaldson (1996) a single-process strength model of recognition memory predicts that measures of discriminability derived from overall hits and false alarms will not differ from measures obtained from remember hits and false alarms. The greater value for the recognition estimate suggests that recognition relies on additional processing not observed with remember responding (Gardiner & Conway, 1999; Gardiner & Gregg, 1997; Gardiner et al., 2001). Hence these results support a dual-process model of recognition memory.

In summary, the results of these four experiments provide strong support for the independence model of the remember/know distinction, and challenge both the exclusive-processes view of Gardiner and colleagues and the single-process view of Donaldson and others. The findings provide only scant support for the hypothesis that as a consequence of delay or change in context, memory attributes that contribute to recollection can come to contribute to familiarity.

## References


Aaronson, D., & Watts, B. (1987). Extensions of Grier's computational formulas for *A'* and *B"* to below-chance performance. *Psychological Bulletin, 102,* 439-442.

Bodner, G. E., & Lindsay, D. S. (in press). Remembering and knowing in context. *Journal of Memory and Language.*

Bodner, G. E., & Masson, M. E. J., & Caldwell, J. I. (2000). Evidence for a generate-recognize model of episodic influences on word-stem completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 267-293.

Buchner, A., Erdfelder, E., & Vaterrodt-Plunnecke, B. (1995). Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General, 124,* 137-160.

Curran, T., & Hintzman, D. L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 531-547.

Curran, T., & Hintzman, D. L. (1997). Consequences and causes of correlations in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 496-504.

Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience, 13,* 89-109.

Dodson, C. S., & Johnson, M. K. (1996). Some problems with the process-dissociation approach to memory. *Journal of Experimental Psychology: General, 125,* 181-

194.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General, 121,* 275-277.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition, 24,* 523-533.

Dunn, J. C., & Kirsner, K. (1989). Implict memory: Task or process? In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit Memory: Theoretical Issues* (pp. 17-31). Hillsdale, NJ: Lawrence Erlbaum Associates.

Duzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H. J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Science, 94,* 5973-5978.

Elliott, P. (1964). Appendix 1-Tables of *d'*. In J. A. Swets (Ed.), *Signal Detection and Recognition by Human Observers: Contemporary Readings* (pp. 651-684). New York: John Wiley & Sons, Inc.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28,* 1-11.

Fantz, R. L. (1961). The origin of form perception. *Scientific American, 204,* 66-72.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition, 16,* 309-313.

Gardiner, J. M., & Conway, M. A. (1999). Levels of awareness and varieties of experience. In B. H. Challis, & B. M. Velichkovsky (Eds.), *Stratification in Cognition and Consciousness* (pp. 237-254). Amsterdam, Netherlands: John

Benjamins Publishing Co.

Gardiner, J. M., & Gregg, V. H. (1997). Recognition with little or no remembering: Implications for a detection model. *Psychonomic Bulletin & Review, 4,* 474-479.

Gardiner, J. M., Gregg, V. H., Mashru, R., & Thaman, M. (2001). Impact of encoding depth on awareness of perceptual effects in recognition memory. *Memory & Cognition, 29,* 433-440.

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition, 18,* 23-30.

Gardiner, J. M., & Java, R. I. (1991). Forgetting in recognition memory with and without recollective experience. *Memory & Cognition, 19,* 617-623.

Gardiner, J. M., & Java, R. I. (1993). Recognising and remembering. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of Memory* (pp. 163-188). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gardiner, J. M., Java, R. I., & Richardson-Klavehn, A. (1996). How level of processing really influences awareness in recognition memory. *Canadian Journal of Experimental Psychology, 50,* 114-122.

Gardiner, J. M., Kaminska, Z., Dixon, M., & Java, R. I. (1996). Repetition of previously novel melodies sometimes increases both remember and know responses in recognition memory. *Psychonomic Bulletin & Review, 3,* 366-371.

Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition, 18,* 579-583.

Gardiner, J. M., Richardson-Klavehn, A., & Ramponi, C. (1998). Limitations of the signal detection model of the remember-know paradigm: A reply to Hirshman. *Consciousness and Cognition, 7,* 285-288.

Graf, P., & Komatsu, S. (1994). Process dissociation procedure: Handle with caution! *European Journal of Cognitive Psychology, 6,* 113-129.

Green, D. M., & Swets, J. A. (1974). *Signal Detection Theory and Psychophysics.* New York: Robert E. Krieger Pub. Co., Inc.

Gregg, V. H., & Gardiner, J. M. (1994). Recognition memory and awareness: A large effect of study-test modalities on "know" responses following a highly perceptual orienting task. *European Journal of Cognitive Psychology, 6,* 131-147.

Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 259-278.

Hirshman, E., & Master, S. (1997). Modelling the conscious correlates of recognition memory: Reflections on the remember-know paradigm. *Memory & Cognition, 25,* 345-351.

Hu, X., & Batchelder, W. H. (1994). The statistical analysis of general processing tree models with the EM algorithm. *Psychometrika, 59,* 21-47.

Hu, X., & Phillips, G. A. (1999). GPT.EXE: A powerful tool for the visualization and analysis of general processing tree models. *Behavior Research Methods, Instruments, & Computers, 31,* 220-234.

Inoue, C., & Bellezza, F. S. (1998). The detection model of recognition using know and remember judgments. *Memory & Cognition, 26,* 299-308.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513-541.

Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 3-26.

Jacoby, L. L., Begg, I. M., & Toth, J. P. (1997). In defense of functional independence: Violations of assumptions underlying the process-dissociation procedure? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 484-495.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306-340.

Jacoby, L. L., Jones, T. C., & Dolan, P. O. (1998). Two effects of repetition: Support for a dual-process model of know judgments and exclusion errors. *Psychonomic Bulletin & Review, 5,* 705-709.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122,* 139-154.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific Approaches to Consciousness* (pp. 13-47). Mahwah, NJ: Lawrence Erlbaum Associates.

Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention and control. *Psychology and Aging, 8,* 283-293.

Johnson, M. K., & Chalfonte, B. L. (1994). Binding complex memories: The role of reactivation and the hippocampus. In D. L. Schacter & E. Tulving (Eds.), *Memory Systems 1994,* (pp. 311-350). Cambridge, MA: The MIT Press.

Kadlec, H. (1999). Statistical properties of *d'* and β estimates of signal detection theory. *Psychological Methods, 4,* 22-43.

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

Kelley, C. M., & Lindsay, D. S. (1996). Conscious and unconscious forms of memory. In R. Bjork & E. Bjork (Eds.), *Handbook of Perception and Cognition: Memory* (Volume 10, pp. 31-63). New York: Academic Press.

Lindsay, D. S., & Kelley, C. M. (1996). Creating illusions of familiarity in a cued recall remember/know paradigm. *Journal of Memory and Language, 35,* 197-211.

Macken, W. J. (2002). Environmental context and recognition: The role of recollection and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 153-161.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge: Cambridge University Press.

Macmillan, N. A., & Creelman, C. D. (1996). Triangles in ROC space: History and theory of "nonparametric" measures of sensitivity and bias. *Psychonomic Bulletin & Review, 3,* 164-170.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252-271.

Mäntylä, T. (1997). Recollections of faces: Remembering differences and knowing similarities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1203-1216.

Mäntylä, T., & Cornoldi, C. (2002). Remembering changes: Repetition effects in face recollection. *Acta Psychologica, 109,* 95-105.

Masson, M. E. J. (1989). Fluent reprocessing as an implicit expression of memory for experience. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit Memory: Theoretical Issues* (pp. 123-138). Hillsdale, NJ: Erlbaum.

Masson, M. E. J., & Freedman, L. (1990). Fluent identification of repeated words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 355-373.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Memory, 16,* 519-533.

Mulligan, N. W., & Hirshman, E. (1997). Measuring the bases of recognition memory: An investigation of the process-dissociation framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 280-304.

Parkin, A. J., Gardiner, J. M., & Rosser, R. (1995). Functional aspects of recollective experience in face recognition. *Consciousness and Cognition, 4,* 387-398.

Parkin, A. J., & Walter, B. M. (1992). Recollective experience, normal aging, and frontal dysfunction. *Psychology and Aging, 7,* 290-298.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21,* 89-102.

Rajaram, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 365-377.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39,* 475-543.

Richardson-Klavehn, A., Gardiner, J. M., & Java, R. I. (1996). Memory: Task dissociations, process dissociations and dissociations of consciousness. In G. Underwood (Ed.), *Implicit Cognition* (pp. 85-158). New York: Oxford University Press.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95,* 318-339.

Roediger, H. L., III & McDermott, K. B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Volume 8, pp.62-131). Amsterdam: Elsevier.

Roediger, H. L., III & McDermott, K. B. (1994). The problem of differing false-alarm rates for the process dissociation procedure: Comment on Verfaellie and Treadwell (1993). *Neuropsychology, 8,* 284-288.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501-518.

Schneider, W. (1995). *Micro-Experimental Laboratory Professional* [Computer program]. Pittsburgh, PA: Psychology Software Tools, Inc.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34-50.

Toth, J. P. (1996). Conceptual automaticity in recognition memory: Levels-of-processing effects on familiarity. *Canadian Journal of Experimental Psychology, 50,* 123-138.

Toth, J. P. (2000). Nonconscious forms of human memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford Handbook of Memory* (pp. 245-261). New York: Oxford University Press.

Toth, J. P., Reingold, E. M., & Jacoby, L. L. (1994). Toward a redefinition of implicit memory: Process dissociations following elaborative processing and self-generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 290-303.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26,* 1-12.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 359-380.

Whittlesea, B. W. A. (2002). Two routes to remembering (and another to remembering not). *Journal of Experimental Psychology: General, 131,* 325-348.

Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language, 29,* 716-732.

Xu, M., & Bellezza, F. S. (2001). A comparison of the multimemory and detection theories of know and remember recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1197-1210.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1341-1354.

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition, 25,* 747-763.

Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and response speed. *Canadian Journal of Experimental Psychology, 48,* 516-534.

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language, 34,* 622-643.

Yonelinas, A. P., & Jacoby, L. L. (1996). Noncriterial recollection: Familiarity as automatic, irrelevant recollection. *Consciousness and Cognition, 5,* 131-141.

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology, 12,* 323-339.

Yonelinas, A. P., Regehr, G., & Jacoby, L. L. (1995). Incorporating response bias in a dual-process theory of memory. *Journal of Memory and Language, 34,* 821-835.

## Appendix A

### Test Instructions

In this phase of the experiment a series of faces will be presented one at a time. Some of the faces were presented earlier in the first phase of the experiment (i.e., the phase where you made sex judgments) while others were not.

Your task is to decide whether or not each face was presented earlier. The faces you will see in this second phase are presented only once; therefore, if you believe you've seen a face before then it must have been in the first phase of the experiment. The exact same photo used in the first phase of the experiment is used in this phase of the experiment.

In addition, after those trials where you respond Yes indicate whether you remember (R) or know (K) that the face was presented earlier.

If you REMEMBER the face then you are sure that it was presented earlier and you are consciously aware of some aspect or aspects of its prior occurrence. A face that is remembered is one that is accompanied by a particular association (e.g., you remember making a judgment) or information about its position among different faces (e.g., it was the first face, or it appeared after a particular face, or it appeared before a particular face, etc.) or something more personal (e.g., you thought this person reminds you of a friend).

If you KNOW the face then you are sure that it was presented earlier but you cannot specify an aspect or aspects of your prior experience with the face--you just know that this face was one presented in the first phase of this experiment.

An example will help clarify the difference between these responses. Imagine you see a man on the bus and you immediately say to yourself that you've seen him before. If you consciously recollect seeing him in a white apron and cap and that he was standing behind the meat counter then you recognize him in the REMEMBER sense: He's the butcher at the local grocery store. If you cannot consciously recollect any specific information indicating where or how you know him but you are certain that you know him, then you recognize him in the KNOW sense.

# Appendix B

## Example of Counterbalancing scheme for Experiment 4

F = Face sets; C = Context sets

Condition 1:

| Study | Test 1 | Test 2 |
|---|---|---|
| F1C1 | F1C1: studied face, studied context | |
| F2C2 | F2C4: studied face, new context | F2C2 |
| F3C3 | F3C2: studied face, switched context | F3C3 |
| | F4C5: non-studied face, "switched" context | F4C4 |
| | F5C3: non-studied face, studied context | F5C5 |
| | F6C6: non-studied face, new context | |

# Appendix C

Table C1

*Significant Results for Uncorrected and (H - FA) Remember Response Probabilities in Experiment 1*

| | Source | *F* | df | *MSE* | *p* | $\eta^2$ |
|---|---|---|---|---|---|---|
| Uncorrected | | | | | | |
| | Item Type | 197.31 | 1, 78 | .013 | < .001 | .72 |
| | Delay | 6.80 | 1, 78 | .041 | < .02 | .039 |
| (H - FA) | | | | | | |
| | Item Type | 197.16 | 1, 78 | .013 | < .001 | .72 |
| | Delay | 4.22 | 1, 78 | .029 | < .05 | .021 |

*Note.* (H - FA) = hits minus false alarms.

Table C2

*Significant Results and Post Hoc Comparisons for Uncorrected Know Response Probabilities in Experiment 1*

| Source | *F* | df | *MSE* | *p* | $\eta^2$ |
|---|---|---|---|---|---|
| Item Type | 5.38 | 1, 78 | .010 | < .03 | .065 |
| Item Type x Delay | 5.69 | 1, 78 | .010 | < .02 | .068 |

Post Hoc Comparisons:

Non-repeated: $t(78) = 2.28, p < .03, r^2 = .062$

Repeated: $t < 1$

No delay: $t < 1$

3D delay: $t(39) = 3.76, p < .001, r^2 = .27$

*Note.* 3D = 3 days.

Table C3

*Significant Results and Post Hoc Comparisons for Uncorrected IRK Familiarity Estimates in Experiment 1*

| Source | *F* | df | *MSE* | *p* | $\eta^2$ |
|---|---|---|---|---|---|
| Item Type | 68.05 | 1, 78 | .016 | < .001 | .47 |
| Interaction | 4.16 | 1, 78 | .016 | < .05 | .051 |

Post Hoc Comparisons:

Non-repeated: $t(78) = 2.81, p < .01, r^2 = .17$

Repeated: $t < 1$

No delay: $t(39) = 4.00, p < .001, r^2 = .29$

3D delay: $t(39) = 8.16, p < .001, r^2 = .63$

*Note.* Interaction = Item Type x Delay; 3D = 3 days.

Table C4

*Significant Results for Uncorrected Remember Response Probabilities in Experiment 2*

| Source | *F* | df | *MSE* | *p* | $\eta^2$ |
|---|---|---|---|---|---|
| Item Type | 439.11 | 1, 39 | .025 | < .001 | .92 |
| Delay | 75.95 | 1, 39 | .020 | < .001 | .66 |

Table C5

*Significant Results for (H - FA) Remember Response Probabilities in Experiment 2*

| Source | *F* | df | *MSE* | *p* | $\eta^2$ |
|---|---|---|---|---|---|
| Item Type | 428.95 | 1, 39 | .025 | < .001 | .92 |
| Delay | 71.95 | 1, 39 | .020 | < .001 | .65 |

*Note.* (H - FA) = hits minus false alarms.

Table C6

*Significant Results and Post Hoc Comparisons for Uncorrected Know Response Probabilities in Experiment 2*

| Source | *F* | df | *MSE* | *p* | $\eta^2$ |
|---|---|---|---|---|---|
| Item Type | 9.55 | 1, 39 | .041 | < .005 | .20 |
| Interaction | 8.96 | 1, 39 | .030 | < .005 | .19 |

Post Hoc Comparisons:

Non-repeated: $t(39) = 2.31, p < .03, r^2 = .12$

Repeated: $t(39) = 2.80, p < .01, r^2 = .17$

No delay: $t(39) = 5.14, p < .001, r^2 = .40$

1W delay: $t < 1$

*Note.* 1W = 1 week; Interaction = Item Type x Delay.

Table C7

*Tests of Significance and Post Hoc Comparisons for Uncorrected IRK Familiarity Estimates in Experiment 2*

| Source | *F* | df | *MSE* | *p* | $\eta^2$ |
|---|---|---|---|---|---|
| Item Type | 13.38 | 1, 39 | .119 | < .001 | .26 |
| Delay | 5.40 | 1, 39 | .093 | < .03 | .12 |
| Interaction | 3.69 | 1, 39 | .103 | .062 | .086 |

Post Hoc Comparisons:

Non-repeated: $t(39) = 4.72, p < .001, r^2 = .36$

Repeated: $t < 1$

No delay: $t(39) = 1.21, p = .23, r^2 = .036$

1W delay: $t(39) = 4.72, p < .001, r^2 = .36$

*Note.* IRK = independence remember/know; 1W = 1 week; Interaction = Item Type x Delay.

Table C8

*Significant Results and Post Hoc Comparisons for Uncorrected and (H - FA) Remember Response Probabilities in Experiment 4*

| Source | $F$ | df | $MSE$ | $p$ | $\eta^2$ |
|---|---|---|---|---|---|
| Context | 91.38 | 2, 46 | .009 | $< .001$ | .80 |

Post Hoc Comparisons:

Studied Context vs. Switched Context: $t(23) = 8.09$, $p < .001$, $r^2 = .74$

Switched Context vs. New Context: $t(23) = 5.75$, $p < .001$, $r^2 = .59$

*Note.* Results of analyses for uncorrected and (H - FA) response probabilities were identical.

Table C9

*Significant Results and Post Hoc Comparisons for Uncorrected Know Response Probabilities in Experiment 4*

| Source | *F* | df | *MSE* | *p* | $\eta^2$ |
|---|---|---|---|---|---|
| Context | 14.78 | 2, 46 | .013 | < .001 | .39 |

Post Hoc Comparisons:

Studied Context vs. Switched Context: $t(23) = 4.83, p < .001, r^2 = .50$

Switched Context vs. New Context: $t < 1$

*Note.* (H - FA) = hits minus false alarms.

Table C10

*Significant Results and Post Hoc Comparisons for Mean Recognition Probabilities in Experiment 4*

| Source | *F* | df | *MSE* | *p* | $\eta^2$ |
|---|---|---|---|---|---|
| Context | 22.18 | 2, 46 | .013 | < .001 | .49 |

Post Hoc Comparisons:

Studied Context ($M$ = .83, $SEM$ = .02) vs. Switched Context ($M$ = .75, $SEM$ = .03): $t(23) = 2.85$, $p < .01$, $r^2 = .26$

Switched Context vs. New Context ($M$ = .61, $SEM$ = .03): $t(23) = 3.82$, $p < .001$, $r^2 = .63$

*Note.* $M$ = mean; $SEM$ = standard error of the mean.

## Appendix D

Multinomial models for the remember/know paradigm for recognition memory described in Experiment 4 were tested using the General Processing Tree (GPT) program (Hu & Phillips, 1999). These are statistically based models that provide a technique that attempts to estimate "unobservable" cognitive processes hypothesized to buttress observable behavioural events (e.g., recognition memory test responses such as Remember and Know) (Riefer & Batchelder, 1988). Simple equations are used to represent processing sequences that are expected to lead to observable data (and hence the relationship among equations is constrained by the theorized relationship between cognitive processes and behavioural data). In addition, there may be diverse sequences of events that lead to a particular behaviour. Each set of events represents a branch of the model much like a branch of a tree (Hu & Batchelder, 1994). In addition to their ability to estimate parameters using maximum likelihood methods, they are also valuable tools in assessing hypotheses about parameter values across different experimental conditions. For the latter the fit between the observed data and the expected values (given a particular hypothesis) is tested. The GPT program provided a value for the chi-square ($\chi^2$) goodness-of-fit statistic for the models tested in Experiment 4.

The models tested for the data of Experiment 4 were based on hypothesized independence, exclusivity, or redundancy between recollection and familiarity processes. Equations were derived to account for the relationship between processes and the three judgment alternatives in the task (i.e., representing different phenomenal states of memory): remember, know, and not in memory (i.e., "new"). In addition to parameters for each process (i.e., recollection and familiarity), each model included parameters to account for those instances where a participant may base a response upon a guessing mechanism. The A-type guessing parameters were decision-making processes assigned to instances where recollection (R) failed or was not available. B-type guessing parameters were those decision-making processes assigned to familiarity (F) processes. However, these guessing mechanisms would primarily account for remember and know responses to non-studied faces (i.e., false alarms)(Bodner, Masson, & Caldwell, 2000; Jacoby, 1998; Xu & Bellezza, 2001).

In Experiment 4, there were four parameters for each of the three context conditions (i.e., studied [s], switched [sw], and new [n]) for a total of 12 parameters: Rs, Fs, As, Bs, Rsw, Fsw, Asw, Bsw, Rn, Fn, An, Bn. Tables D1 to D6 provided at the end of this appendix list the parameters, equations, and expected responses for studied and non-studied faces in each of the context conditions for each model. For each face type (i.e., studied and non-studied) there were three response types (i.e., remember, know, and new) for each context condition for a total of 18 observable categories. Each set of remember, know, and new response frequencies for each face type/context type combination provided a sum of unity (i.e., 1 - (remember + know + new) = 0). Consequently, the degrees of freedom for each set was 2 rather than 3 (i.e., knowing the response frequencies for any two response types was sufficient to derive the response frequency for the remaining response type). This would reduce the expected 18 degrees of freedom (re: different observable frequencies) to 12 (Xu & Bellezza, 2001, p. 1207). Because there were also 12 parameters to be estimated, constraints were required for model testing.

Based on hypotheses derived from the results of Experiment 4, one of the most interesting models included the constraints that guessing parameters (i.e., A, B) and the familiarity parameters in the switched- and new-context conditions should be equivalent (i.e., Asw = An, Bsw = Bn, and Fsw = Fn). For these constraints the independent model fit the data, $\chi^2(3) = 6.71$, $p > .08$, whereas the exclusive and redundant models did not, $\chi^2(4) = 24.84$, $p < .001$[18], and $\chi^2(3) = 17.55$, $p < .001$, respectively. Table D7 lists the observed and expected frequencies for each response type for each model. In addition, Table D8 lists the parameter estimates as derived with each model.

These results provide additional support for the assumption that memory processes that underlie remember and know responding are independent, as opposed to models in which recollection depends upon familiarity, as in the redundancy model, and as opposed to models in which processes are mutually exclusive.

---

18

For the exclusive model, the familiarity parameters could not be constrained without constraining recollection parameters.

Table D1

*Multinomial Model of Recognition Memory for Studied Faces: Exclusive Processes*

| Context Condition | Parameters | Response |
|---|---|---|
| Studied | $R_s$ | Remember |
| | $F_s$ | Know |
| | $(1 - R_s - F_s)\ A_s$ | Remember |
| | $(1 - R_s - F_s)\ B_s$ | Know |
| | $(1 - R_s - F_s)\ (1 - A_s - B_s)$ | New |
| Switched | $R_{sw}$ | Remember |
| | $F_{sw}$ | Know |
| | $(1 - R_{sw} - F_{sw})\ A_{sw}$ | Remember |
| | $(1 - R_{sw} - F_{sw})\ B_{sw}$ | Know |
| | $(1 - R_{sw} - F_{sw})\ (1 - A_{sw} - B_{sw})$ | New |
| New | $R_n$ | Remember |
| | $F_n$ | Know |
| | $(1 - R_n - F_n)\ A_n$ | Remember |
| | $(1 - R_n - F_n)\ B_n$ | Know |
| | $(1 - R_n - F_n)\ (1 - A_n - B_n)$ | New |

*Note.* R = recollection; F = familiarity; A = recollection-type guess process; B = familiarity-type guess process; s = studied context; sw = switched context; n = new context.

Table D2

*Multinomial Model of Recognition Memory for Non-studied Faces: Exclusive Processes*

| Context Condition | Parameters | Response |
|---|---|---|
| Studied | $A_s$ | Remember |
| | $B_s$ | Know |
| | $(1 - A_s - B_s)$ | New |
| Switched | $A_{sw}$ | Remember |
| | $B_{sw}$ | Know |
| | $(1 - A_{sw} - B_{sw})$ | New |
| New | $A_n$ | Remember |
| | $B_n$ | Know |
| | $(1 - A_n - B_n)$ | New |

*Note.* R = recollection; F = familiarity; A = recollection-type guess process; B = familiarity-type guess process; s = studied context; sw = switched context; n = new context.

Table D3

*Multinomial Model of Recognition Memory for Studied Faces: Independent Processes*

| Context Condition | Parameters | Response |
|---|---|---|
| Studied | $R_s F_s$ | Remember |
| | $R_s (1 - F_s)$ | Remember |
| | $(1 - R_s) F_s A_s$ | Remember |
| | $(1 - R_s) F_s (1 - A_s)$ | Know |
| | $(1 - R_s) (1 - F_s) A_s B_s$ | Remember |
| | $(1 - R_s) (1 - F_s) A_s (1 - B_s)$ | Remember |
| | $(1 - R_s) (1 - F_s) (1 - A_s) B_s$ | Know |
| | $(1 - R_s) (1 - F_s) (1 - A_s) (1 - B_s)$ | New |
| Switched | $R_{sw} F_{sw}$ | Remember |
| | $R_{sw} (1 - F_{sw})$ | Remember |
| | $(1 - R_{sw}) F_{sw} A_{sw}$ | Remember |
| | $(1 - R_{sw}) F_{sw} (1 - A_{sw})$ | Know |
| | $(1 - R_{sw}) (1 - F_{sw}) A_{sw} B_{sw}$ | Remember |
| | $(1 - R_{sw}) (1 - F_{sw}) A_{sw} (1 - B_{sw})$ | Remember |
| | $(1 - R_{sw}) (1 - F_{sw}) (1 - A_{sw}) B_{sw}$ | Know |
| | $(1 - R_{sw}) (1 - F_{sw}) (1 - A_{sw}) (1 - B_{sw})$ | New |
| New | $R_n F_n$ | Remember |
| | $R_n (1 - F_n)$ | Remember |
| | $(1 - R_n) F_n A_n$ | Remember |
| | $(1 - R_n) F_n (1 - A_n)$ | Know |
| | $(1 - R_n) (1 - F_n) A_n B_n$ | Remember |
| | $(1 - R_n) (1 - F_n) A_n (1 - B_n)$ | Remember |
| | $(1 - R_n) (1 - F_n) (1 - A_n) B_n$ | Know |
| | $(1 - R_n) (1 - F_n) (1 - A_n) (1 - B_n)$ | New |

*Note.* R = recollection; F = familiarity; A = recollection-type guess process; B = familiarity-type guess process; s = studied context; sw = switched context; n = new context.

Table D4

*Multinomial Model of Recognition Memory for Non-studied Faces: Independent Processes*

| Context Condition | Parameters | Response |
|---|---|---|
| Studied | $A_s\ B_s$ | Remember |
| | $A_s\ (1 - B_s)$ | Remember |
| | $(1 - A_s)\ B_s$ | Know |
| | $(1 - A_s)\ (1 - B_s)$ | New |
| Switched | $A_{sw}\ B_{sw}$ | Remember |
| | $A_{sw}\ (1 - B_{sw})$ | Remember |
| | $(1 - A_{sw})\ B_{sw}$ | Know |
| | $(1 - A_{sw})\ (1 - B_{sw})$ | New |
| New | $A_n\ B_n$ | Remember |
| | $A_n\ (1 - B_n)$ | Remember |
| | $(1 - A_n)\ B_n$ | Know |
| | $(1 - A_n)\ (1 - B_n)$ | New |

*Note.* R = recollection; F = familiarity; A = recollection-type guess process; B = familiarity-type guess process; s = studied context; sw = switched context; n = new context.

Table D5

*Multinomial Model of Recognition Memory for Studied Faces: Redundant Processes*

| Context Condition | Parameters | Response |
|---|---|---|
| Studied | $F_s R_s$ | Remember |
| | $F_s (1 - R_s) A_s$ | Remember |
| | $F_s (1 - R_s) (1 - A_s)$ | Know |
| | $(1 - F_s) B_s R_s$ | Remember |
| | $(1 - F_s) B_s (1 - R_s) A_s$ | Remember |
| | $(1 - F_s) B_s (1 - R_s) (1 - A_s)$ | Know |
| | $(1 - F_s) (1 - B_s)$ | New |
| Switched | $F_{sw} R_{sw}$ | Remember |
| | $F_{sw} (1 - R_{sw}) A_{sw}$ | Remember |
| | $F_{sw} (1 - R_{sw}) (1 - A_{sw})$ | Know |
| | $(1 - F_{sw}) B_{sw} R_{sw}$ | Remember |
| | $(1 - F_{sw}) B_{sw} (1 - R_{sw}) A_{sw}$ | Remember |
| | $(1 - F_{sw}) B_{sw} (1 - R_{sw}) (1 - A_{sw})$ | Know |
| | $(1 - F_{sw}) (1 - B_{sw})$ | New |
| New | $F_n R_n$ | Remember |
| | $F_n (1 - R_n) A_n$ | Remember |
| | $F_n (1 - R_n) (1 - A_n)$ | Know |
| | $(1 - F_n) B_n R_n$ | Remember |
| | $(1 - F_n) B_n (1 - R_n) A_n$ | Remember |
| | $(1 - F_n) B_n (1 - R_n) (1 - A_n)$ | Know |
| | $(1 - F_n) (1 - B_n)$ | New |

*Note.* R = recollection; F = familiarity; A = recollection-type guess process; B = familiarity-type guess process; s = studied context; sw = switched context; n = new context.

Table D6

*Multinomial Model of Recognition Memory for Non-studied Faces: Redundant Processes*

| Context Condition | Parameters | Response |
|---|---|---|
| Studied | $B_s A_s$ | Remember |
| | $B_s (1 - A_s)$ | Know |
| | $(1 - B_s)$ | New |
| Switched | $B_{sw} A_{sw}$ | Remember |
| | $B_{sw} (1 - A_{sw})$ | Know |
| | $(1 - B_{sw})$ | New |
| New | $B_n A_n$ | Remember |
| | $B_n (1 - A_n)$ | Know |
| | $(1 - B_n)$ | New |

*Note.* R = recollection; F = familiarity; A = recollection-type guess process; B = familiarity-type guess process; s = studied context; sw = switched context; n = new context.

Table D7

*Expected (Exp) and Observed (Obs) Frequencies for the Exclusive (E), Independence (I), and Redundant (Rd) Models when Fsw = Fn, Asw = An, and Bsw = Bn*

| | Response | Obs | Exp E | Exp I | Exp Rd |
|---|---|---|---|---|---|
| S Face, S Context | | | | | |
| | Remember | .607 (.030) | .607 | .607 | .607 |
| | Know | .219 (.023) | .219 | .219 | .219 |
| | New | .174 (.021) | .174 | .174 | .174 |
| S Face, Sw Context | | | | | |
| | Remember | .372 (.035) | .302 | .372 | .338 |
| | Know | .375 (.032) | .376 | .338 | .340 |
| | New | .253 (.030) | .322 | .289 | .322 |
| S Face, N Context | | | | | |
| | Remember | .232 (.030) | .302 | .232 | .258 |
| | Know | .378 (.026) | .376 | .414 | .420 |
| | New | .391 (.031) | .322 | .354 | .322 |
| NS Face, S Context | | | | | |
| | Remember | .076 (.018) | .076 | .076 | .076 |
| | Know | .182 (.027) | .182 | .182 | .182 |
| | New | .742 (.034) | .742 | .742 | .742 |
| NS Face, Sw Context | | | | | |
| | Remember | .029 (.014) | .031 | .031 | .031 |
| | Know | .141 (.028) | .132 | .132 | .132 |
| | New | .831 (.030) | .837 | .837 | .837 |
| NS Face, N Context | | | | | |
| | Remember | .034 (.011) | .031 | .031 | .031 |
| | Know | .122 (.023) | .132 | .132 | .132 |
| | New | .844 (.029) | .837 | .837 | .837 |

*Note.* Numbers in parentheses represent the standard error of the mean; S = Studied, NS = Non-studied; Sw = Switched, N = New.

Table D8

*Parameter Estimates and 95% Confidence Intervals for the Independent and Redundant Models when Fsw = Fn, Asw =An, and Bsw = Bn, and for the Exclusive Model when Fsw = Fn, Rsw = Rn, Asw = An, Bsw = Bn*

| Parameter | Exclusive | Independent | Redundant |
|---|---|---|---|
| Rs | .589 (.537-.641) | .575 (.520-.629) | .625 (.542-.709) |
| Rsw | .290 (.257-.323) | .352 (.302-.403) | .379 (.290-.468) |
| Rn | .290 (.257-.323) | .207 (.162-.252) | .233 (.132-.334) |
| Fs | .176 (.130-.222) | .447 (.345-.550) | .765 (.712-.818) |
| Fsw | .326 (.287-.365) | .467 (.416-.518) | .616 (.575-.657) |
| Fn | .326 (.287-.365) | .467 (.416-.518) | .616 (.575-.657) |
| As | .076 (.049-.102) | .076 (.049-.102) | .293 (.203-.383) |
| Asw | .031 (.019-.044) | .031 (.019-.044) | .192 (.123-.261) |
| An | .031 (.019-.044) | .031 (.019-.044) | .192 (.123-.261) |
| Bs | .182 (.144-.221) | .197 (.156-.239) | .258 (.214-.302) |
| Bsw | .132 (.108-.155) | .136 (.111-.160) | .163 (.137-.189) |
| Bn | .132 (.108-.155) | .136 (.111-.160) | .163 (.137-.189) |
| $\chi^2$ | 24.841* | 6.707 | 17.554* |
| df | 4 | 3 | 3 |

*Note.* F = familiarity; A = recollection-type guess process; B = familiarity-type guess process; R = recollection; sw = switched context, n = new context; s = studied context; df = degrees of freedom; * $p < .001$.