Title
Broadening the Scope of Recognition Memory

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 39(0)

Authors
Olds, Justin M.
Marewski, Julian N.

Publication Date
2017

Peer reviewed

Broadening the Scope of Recognition Memory

Justin M. Olds (justin.olds@unil.ch) & Julian N. Marewski (julian.marewski@unil.ch)
Department of Organizational Behavior, University of Lausanne
1015 Lausanne, Switzerland

Abstract

Within the literature of psychological and decision sciences, there is a critical difference in the way recognition is defined and studied experimentally. To address this difference, the current experiment examines and attempts to disentangle the influence of two recognition judgment sources (from within an experiment and from an individual's prior life experiences) upon two different recognition judgments. By presenting participants with a set of related stimuli that vary naturally in environmental occurrence and by manipulating exposure within an experimental context, this experiment allows for a broader and more ecologically valid assessment of recognition memory. Contrasting with the typical word-frequency effect, the results reveal an overall bias to judge high-frequency items as studied on an episodic recognition test. Additionally, the results underscore the role of context by showing that a single study exposure increases the probability that individuals will judge stimuli as presented outside the laboratory.

Keywords: Recognition memory; decision-making; ecological validity

Recognizing the Difference between Recognition and Recognition

In general, *recognition* refers to the experience wherein upon encountering a stimulus an individual has a sense that she has encountered that stimulus before. A *recognition judgement*, in turn, is when an individual explicitly claims that a stimulus was previously encountered. Within the literatures of psychological and decision sciences, there is a critical difference in the way recognition is studied experimentally. While one set of research focuses on an individual's capacity to recognize stimuli presented previously within an experimental episode (*episodic recognition memory*), the other set focuses on an individual's capacity to recognize stimuli as previously encountered during the individual's prior life experiences before beginning an experiment (*pre-experimental recognition memory*). Although these types of recognition are typically studied independently, the sources of experience related to both are inherently interconnected. Indeed, the experience of recognition is influenced by an individual's prior life experiences as well as by the experiences she has during an experiment. The current work provides a framework for studying recognition memory in a way that more readily relates to these two intertwined factors.

In what follows, we first broadly describe the lines of research related to both types of recognition judgements, including prior work that has examined their interconnected nature. Within this review, we note criticisms of each line of research. Following this, we (1) describe a research methodology that draws upon both lines of work to address these concerns, (2) present the results of an experiment adopting this approach, and (3) discuss implications of these results and considerations for future work.

Episodic Recognition in Memory Research

Recognition memory has been studied extensively with list-learning experiments. Here, stimuli, such as words or pictures, are presented individually in the form of a *study* list. After a delay ranging from a few seconds to multiple days, participants are presented with a recognition memory *test* list, and are asked to discriminate *targets* (stimuli from the study list) from *foils* (new items). Episodic recognition memory has been the focus of decades of extensive research and has been noted as an increasingly prevalent research paradigm (e.g., Hintzman, 2011).

A major strength of the episodic recognition memory line of research is experimental control. This is, in part, achieved by minimizing the role of individual stimuli, such as by presenting mixed lists of unrelated and uncommon concrete nouns. This approach follows in the footsteps of pioneering memory researcher Ebbinghaus (1885), who used nonsense stimuli to avoid the influence of everyday exposure.

The advantage of striving for experimental control in this way is also a disadvantage when it comes to understanding how memory judgments operate within everyday decision-making. For example, Neisser (1976) argued that memory research should strive toward *ecological validity*. Drawing upon work in perception by Brunswik (1957) and Gibson (1979), ecological validity refers to applicability outside the laboratory. The importance of ecological validity is underscored by research on eyewitness testimony. For instance, when participants study mixed lists of unrelated words, a positive relationship between accuracy and confidence is typically found (e.g., Dallenbach, 1913; Dunlosky & Metcalfe, 2009). This intuitive finding dovetails with the 1972 and 1976 U.S. Supreme Court rulings suggesting that highly confident eyewitness identification is likely accurate. This pattern, however, does not hold for lists of similar material (i.e., categorized lists, see e.g., DeSoto & Roediger, 2014), such as description details of suspects presented to an eyewitness (e.g., Smith, Kassin, & Ellsworth, 1989). It is disconcerting to consider how other memory research findings might also be misleading due to a similar lack of ecological validity.

Departing from the convention of presenting participants with mixed lists of uncommon, unrelated stimuli would inherently introduce an additional factor, which might interact with or overshadow other experimental manipulations. Specifically, a by-product of presenting participants with a more ecologically valid set of related stimuli is that the individual items will vary based on how often each occurs outside the laboratory. The influence of environmental frequency upon episodic recognition judgements has been the focus of extensive research (e.g., Dennis & Humphreys, 2001; Estes, 1994; Glanzer & Adams, 1985; Lohnas & Kahana, 2013; McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997). In these experiments, study and test lists are composed of unrelated words that are sampled randomly from a range of low linguistic frequency words and high linguistic frequency words. The typical word-frequency finding on an episodic recognition memory test is that low-frequency target words are more accurately judged to be "old" than high-frequency target words and low-frequency foil words are more accurately judged to be "new" than high-frequency foils. Although this line of research does begin to reintroduce the influence of pre-experimental exposure into a laboratory setting, these experiments still tend to favor experimental control over ecological validity in numerous ways. First, the study lists in these experiments are typically composed of stimuli that are sampled from opposite poles along a continuum of environmental frequency—either extremely low-frequency or extremely high-frequency (but see Lohnas & Kahana, 2013, for an exception), essentially transforming the naturally continuous variable of environmental frequency into a dichotomized factor. Second, the study and test lists used in these experiments are often composed of unrelated words. That is, the stimuli belong to many disparate categories and, thus, do not related to one another regarding any real-world inferences, such as person details in relation to culpability.

It may be the case that, similar to experiments investigating the word-frequency effect (e.g., Glanzer & Adams, 1985), a mnemonic advantage for low-frequency items emerges if participants are tested with a related set of stimuli sampled with varying degrees of environmental frequency. There is some evidence, however, suggesting that this pattern of superior recognition accuracy for low-frequency items might not persist. Specifically, in an experiment by Jacoby, Woloshyn, and Kelley (1989), which investigated the influence of environmental frequency by composing study and test lists of famous and nonfamous names, the results of an episodic recognition test revealed that famous names (i.e., high-frequency items) were more likely to be judged as presented on the study list than nonfamous ones (i.e., low-frequency items), regardless of whether or not they were actually studied. Thus, it may turn out that the low-frequency item advantage, borne out of memory research favoring experimental control over ecological validity, may not hold when participants are tested with sets of related stimuli.

Pre-Experimental Recognition in Decision-Making Research

Numerous researchers have investigated how individuals use pre-experimental recognition, a sense of prior exposure to a stimulus outside the laboratory, during decision-making (e.g., Dougherty, Gettys, & Ogden, 1999; Erdfelder, Küpper-Tetzel, & Mattern, 2011; Goldstein & Gigerenzer, 2002; Hertwig, Herzog, Schooler, & Reimer, 2008; Marewski & Schooler, 2011). A sense of pre-experimental recognition has been shown to influence a wide array of judgments, including about population size (e.g., Marewski & Schooler, 2011), fame (Jacoby, Woloshyn, & Kelley, 1989), and company revenue (Hertwig, et al., 2008), to name but a few. This work has shown that the frequency of occurrence for a given stimulus in the environment (which can be estimated by counting how often a stimulus, say a city name, occurs on the Internet or in the print media), allows for modeling how likely and how quickly the stimulus is to be recognized (e.g., Hertwig et al., 2008). Overall, this body of research underscores the intuitive notion that a sense of recognition is paramount for making many everyday decisions.

The influence of pre-experimental recognition upon decision-making is typically studied within *probabilistic inference* experiments, in which individuals are assumed to use known attributes of a stimulus as *cues* to make inferences about an unknown or future criterion. These experiments typically include a recognition task and a paired-comparison inference task. In the recognition task, participants are shown a list of related stimuli, such as city names, and asked to judge if they recognize each item from their prior life experience. Additionally, some experiments ask individuals to report if they have additional knowledge beyond a sense of pre-experimental recognition for each object. Reponses from this task and their respective response times are later used, to predict judgments on the paired-comparison inference task. In the paired-comparison inference task, individuals are shown two items at a time from a related set of stimuli and asked to infer which alternative is higher or lower on some judgement criterion, such as which of two cities has a larger population size.

In part because of its rigid simplicity, one decision strategy in particular, the *recognition heuristic* (Goldstein & Gigerenzer, 2002), has been the focus of much research and debate. This strategy assumes that, stemming from an existing relationship in the world between environmental occurrence and a given criterion (e.g., population size), a sense of pre-experimental recognition can readily guide decisions in a straightforward way. Specifically, the recognition heuristic assumes that on a paired-comparison inference task, if one decision alternative is recognized and the other is not, individuals will judge the recognized alternative to have a higher value on the criterion.

In general, this line of research is aptly commended for showcasing and exploring how memory is employed for everyday decisions. One major criticism of this research, however, is that the assumptions about recognition memory fail to appropriately draw upon theory from the abundance of related recognition memory research (e.g., Dougherty, Franco-Watkins, & Thomas, 2008; Newell & Fernandez, 2006). It appears that much of the work on recognition-based decision-making assumes that pre-experimental recognition is a fixed commodity, whereas research concerning episodic recognition memory has revealed many ways in which a sense of recognition is influenced by contextual conditions. For example, Jacoby et al. (1989) found that presenting nonfamous names within an experiment study phase increased the likelihood that these nonfamous names would be incorrectly judged as famous later (see also Hertwig et al., 2008). Related to this, Pohl and colleagues (Pohl, Erdfelder, Michalkiewicz, Castela, & Hilbig, 2016), point out two typical experimental procedure choices that fail to consider how a sense of pre-experimental recognition might be influenced by experimental conditions. Both entail how participants, during an earlier part of the experiments, are often exposed to the stimuli that they are later asked to consider during a paired-comparison inference task. First, to obtain a large number of paired-comparison inference trials, items are often paired repeatedly with different items, such that each item appears numerous times during the task. Additionally, the order of the two tasks is often counterbalanced across participants, such that some participants have the recognition task first and others have the paired-comparison inference task first. These two typical methodology choices may influence the sense of pre-experimental recognition and respective recognition speed that individuals might use during decision-making. Although Pohl and colleagues (2016) focused on recognition speed specifically in relation to these methodology concerns, context conditions such as exposure within an experiment might also influence the probability that participants will judge stimuli as recognized from outside the laboratory. This is one of the concerns the current approach addresses.

Experiment

The purpose of the current experiment is to investigate the influence of two fundamental sources of experience, one stemming a person's prior life experiences before entering the laboratory and another stemming from the experiences within an experimental context, upon recognition judgements. Specifically, the memory source related to the experimental context here is a single study exposure and the memory source related to prior life experiences is pre-experimental exposure (estimated with web frequencies). The influence of both sources is examined for both episodic recognition (e.g., "Was this city name presented earlier during the experiment?") and pre-experimental recognition (e.g., "Have you ever heard of this city name before beginning the experiment?"). In line with ecological validity, instead of informing participants that stimuli presented during the study phase would be presented during a later memory test, incidental study exposure was adapted from Hertwig et al. (2008).

In relation to previous work, the current experiment also addresses the following two questions. First, does the typical low-frequency item advantage for episodic recognition memory (e.g., Glanzer & Adams, 1985) occur when individuals are tested with a more ecologically valid set of stimuli (i.e., stimuli from a related set that vary based on their natural occurrence outside the laboratory)? Second, to what extend does a single incidental study exposure influence judgements of pre-experimental recognition, and, if so, does this influence depend on the environmental frequency of stimuli?

Method

Participants A total of 63 individuals (mean age = 21.5 years, 56% female) recruited from the University of Lausanne were paid roughly 26 Swiss francs each (depending on performance) for participating in the experiment. They were tested individually.

Design The study was conducted as a within-subjects design with one independent variable, *study status*, one pseudo-independent variable, *environmental frequency,* and two dependent variables (measured using within-subjects blocks), *episodic recognition judgements* and *pre-experimental recognition judgements.* Study status was manipulated within-subjects by presenting half of the to-be-tested stimuli within a preceding study phase. Pre-experimental frequency was estimated for each stimulus using Wikipedia page occurrences as a proxy for environmental occurrence.

Stimuli The stimuli presented during the experiment were from a set of 200 city names from North American and Western European countries (Canada, England, France, Germany, Italy, Spain, and USA). Additionally, eight extra city names were used at the beginning and end of the study task to minimize primacy and recency study effects. The city names used in the experiment were sampled such that the entire set would include cities from each country that varied in both population size (population statistics obtained from www.citypopulation.de) and environmental occurrence (as approximated by Wikipedia page occurrences) and excluded capital cities. All stimuli were counterbalanced such that each occurred roughly equally often in all study and test conditions.

Procedure The experiment was conducted on a computer using E-Prime experimental software (Psychology Software Tools). Participants were informed that there would be three separate tasks and payment would depend of their performance in each task. First, all participants had an incidental study task, in which half of the city names (100) were presented. Specifically, participants were shown each city name individually and asked to count the number of vowels in each city name. At the beginning of each trial a fixation cross (+) appeared on the screen for 2 s along with a reminder of the task. Participants were informed that the fixation cross would occur immediately before each city name was presented to help them prepare to respond. Afterwards, a city name replaced the fixation cross and participants were prompted to press the appropriate number key on the keyboard corresponding to the number of vowels in the city name. Participants were given up to 4 s to make their response and a blank screen was presented for 2 s between trials.

The following test phase consisted of two separate tasks, an episodic recognition task ("Was this city name presented during the vowel counting task") and a pre-experimental recognition task ("Have you ever heard of this city name before beginning the experiment?"). Participants were asked to respond with one of two keyboard keys to respond "yes" or "no" for each test trial. Participants were given as long as needed to make their response in both tasks. Each task block consisted of 100 trials, which included half studied and half unstudied city names. Importantly, city names were not repeated across test tasks, but were counterbalanced between-subjects such that each city name would occur roughly equally often in each test phase. The order of test tasks was counterbalanced between-subjects. Similar to the study phase, each trial began with a 2 s fixation cross and was followed by a 2 s intertrial interval during both recognition tasks.

Results

The data from all 63 participants was analyzed, excluding trials for one city name due to a clerical error. Wikipedia page occurrence values were log transformed to approximate a linear relationship across city names (e.g., Marewski & Schooler, 2011). Given the two binary dependent variables, separate logistic regressions were planned for both episodic recognition and pre-experimental recognition judgments.

Episodic Recognition Results from the episodic recognition task are presented in Figure 1. From visual inspection of Figure 1, two patterns are apparent. First, the probability that participants judged city names as presented during the study phase was higher for studied city names than for unstudied city names. Second, the probability of judging city names as studied increased as a function of environmental frequency. Moreover, the difference in recognition probabilities between studied and unstudied city names did not seem to vary as a function of environmental frequency. To test the influence of both factors (study status and environmental frequency), a multilevel logistic regression analysis was conducted to fit the episodic recognition data. A test of the model against a constant only model indicated that the predictors as a set provided an improved fit (χ^2 = 980, *df* = 2, p < .001, Nagelkerke's R^2 = .20). The Wald criterion demonstrated both factors, study status (z = 5.38) and environmental frequency (z = 6.82), contributed to the model fit (p < .001 for both). Additionally, by comparing the model to another which included an additional interaction term of the two factors, evidence of an interaction between study status and environmental frequency was not found (χ^2 = .27, *df* = 1, p = .61). To control for the variance associated with the random factor of repeated measurements from individual participants, follow-up generalized linear mixed models were conducted. The same pattern emerged: both factors of study status (z = 5.22, p < .001) and environmental frequency (z = 6.77, p < .001) contributed to the model fit, and there was no evidence of an interaction between the two found (χ^2 = .78, *df* = 1, p = .38).

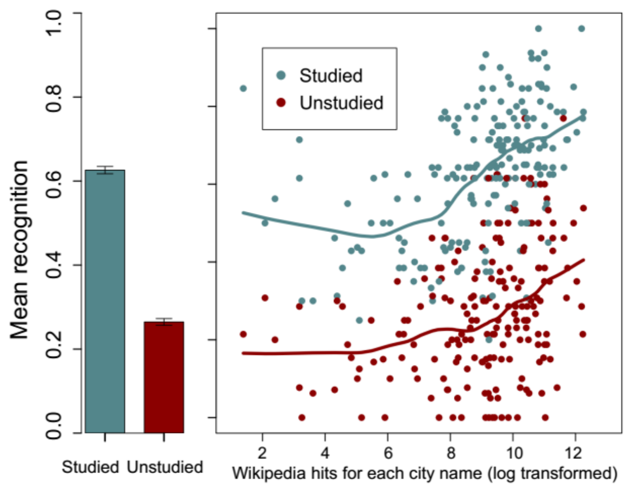

Figure 1: Mean episodic recognition rates. The left side depicts the mean influence of study exposure for all city names with standard error bars. The right side depicts the influence of both study exposure and environmental frequency for each city name with a moving average for both studied and unstudied city names across environmental frequency.

Comparison to previous results The results from the episodic recognition task contrast with the typical word-frequency effect (e.g., Glanzer & Adams, 1985), which reveal an interaction between study and environmental frequency, such that low-frequency targets are recognized more accurately (i.e., higher hit rate and lower false alarm rate) than high-frequency targets and low-frequency lures are correctly rejected more accurately than high-frequency lures (e.g., Lohnas & Kahana, 2013). Instead, the current results show that when participants are tested with a set of related stimuli, high-frequency items are more likely to be judged as studied regardless of if they were studied or unstudied (i.e., higher hit and false alarm rates for high-frequency items). Why did this pattern of results differ from the typical word-frequency effect? Although further work is required to better address this question, we can provide some speculation. The potentially stronger association with the experiment context for low-frequency items, perhaps due to item distinctiveness, may have been relatively diluted in the current experiment for a number of reasons. First, one important difference to remark upon is the overall low episodic recognition accuracy from the current experiment compared to previous word-frequency experiments (e.g., Lohnas & Kahana, 2013). We suspect this difference can be attributed to the difficulty inherent in testing sets of related stimuli and stemming from incidental study (e.g., Criss & Shiffrin, 2004). This increased task difficulty may have led participants be more influenced by the pre-existing associations for high-frequency items, since these associations are less contingent upon the study conditions than the associations between the context and studied items.

Pre-Experimental Recognition Results from the pre-experimental recognition task are presented in Figure 2. From visual inspection of Figure 2, two patterns are apparent. First, the probability of judging city names as recognized from outside the laboratory increased as a function of environmental frequency. Second, the probability that participants judged city names as recognized from outside the laboratory was slightly higher for studied city names than for unstudied city names. Additionally, it appears that the influence of a study exposure was relatively consistent across varying degrees of environmental frequency. To test the influence of both factors (study status and environmental frequency), a multilevel logistic regression analysis was conducted to fit pre-experimental recognition judgements using study status and environmental frequency as predictors. A test of the model against a constant only model indicated that both predictors as a set provided an improved fit ($\chi^2 = 1739$, $df = 2$, $p < .001$, Nagelkerke's $R^2 = .323$). The Wald criterion demonstrated both factors, study status ($z = 4$) and environmental frequency ($z = 32$), contributed to the model fit ($p < .001$ for both). Additionally, by comparing the model to another which included an additional interaction term for the two factors, evidence of an interaction between study status and environmental frequency was not found ($\chi^2 = .621$, $df = 1$, $p = .431$). Evidence for the same pattern was suggested by a generalized linear mixed model with the categorical variable of participant included as a random factor.

Comparison to previous results Similar to previous work (e.g., Marewski & Schooler, 2011), the current results support the use of web frequencies as a reasonable predictor of pre-experimental recognition. The results from the pre-experimental recognition task also converge with previous work showing that an experimental exposure increases the probability of inferring an item to be higher on a criterion related to pre-experimental exposure, such as the fame of individuals (Jacoby et al., 1989) or population size of cities (Hertwig et al., 2008). Unlike previous work, however, the current experiment examines the relationship of a single incidental study exposure across items varying in environmental frequency continuously from extremely infrequent to extremely frequent. Importantly, the current work examines pre-experimental recognition instead of inference judgments, which are assumed to be influenced by a sense of pre-experimental recognition. By focusing on this more basic memory judgment, the current approach and respective data reveal that the presentation of stimuli within an experimental context influences a sense of pre-experimental recognition that is core to much research on memory-based decision research (e.g., Goldstein & Gigerenzer, 2002). One novel finding is that, because of the lack of an interaction between study and environmental frequency, it appears a single study exposure results in a relatively constant increase in the probability of pre-experimental recognition across all items, regardless of environmental frequency.

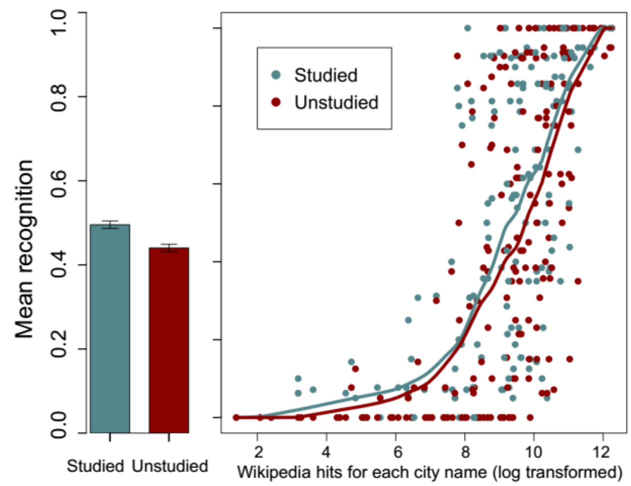

Figure 2: Mean pre-experimental recognition rates. The left side bar graph depicts the mean influence of study exposure for all city names with standard error bars. The right side depicts the influence of both study exposure and environmental frequency for each city name with a moving average for both studied and unstudied city names across environmental frequency.

Discussion

The main purpose of the current work is to showcase a broad approach to studying recognition memory—one that considers how different types of recognition emerge as a function of the interconnected factors of (1) an individual's prior life experiences and (2) an individuals' recent and current experiences within an experimental context. This approach was designed to better translate results from memory experiments into real-world situations. This was achieved by testing recognition judgements for related stimuli, which, in turn, we assume more readily relate to everyday recognition judgments and memory-based inferences. For instance, one might be asked to identify or corroborate which colleagues were present at a company meeting or holiday party. This would entail gauging exposure within a context for a related set of stimuli (e.g., co-workers) that vary based on their environmental frequency (i.e., some are more well-known than others). This kind of everyday memory task contrasts sharply with the typical kind of recognition memory task used in psychology experiments, in which stimuli are unrelated and environmental occurrence is either constrained or dichotomized into highly disparate factor levels (extremely low and high-frequency bins).

The importance of adopting a broader and more ecologically valid approach to understanding recognition memory is underscored by the current experiment results. In contrast to the typical word frequency effect (increased episodic recognition accuracy for low-frequency items) found in many previous experiments, (e.g., Glanzer & Adams, 1985; Lohnas & Kahana, 2013), the current experiment results showed an increased tendency to judge high-frequency items from a set of related stimuli as studied.

The current experiment results also provide evidence suggesting that context conditions, such as a single incidental study exposure, influence pre-experimental recognition judgements. This finding suggests that researchers examining memory-based inferences should strongly consider how often and in what manner stimuli are presented within an experiment. Related to this concern, although separate sets of stimuli were presented during both recognition test phases (episodic and pre-experimental) of the current experiment, we reanalyzed both sets of data with the inclusion of test task order as a factor to help rule out the influence of task demands upon the results. For both recognition tasks, the same main effects (study exposure and environmental frequency) and lack of interaction were supported. Importantly, these results did not interact with task order and a main effect of task order was not found.

There are numerous possible extensions of the current work. One is to incorporate the influence of context factors into models of memory-based decision-making. Additionally, the influence of list composition (e.g., Malmberg & Murnane, 2002) upon both episodic and pre-experimental recognition can be explored with sets of related stimuli, such as city names. Although the current experiment included related stimuli that varied widely on environmental frequency and the stimulus set was somewhat balanced, in that half of the city names were typically recognized, it remains largely unexplored to what degree a sense of recognition is influenced by the composition of study and test lists of related sets of stimuli. Similar to list composition effects, testing other stimulus materials, such as eyewitness-related description details, may help reveal the influence of varying environmental occurrence patterns.

Author’s Note

Correspondence: justin.olds@unil.ch. We would like to thank Ulrich Hoffrage and the four anonymous reviewers for their helpful comments on a previous version of this article. This work was supported by a grant from the Swiss National Science Foundation [159822].

References

Brunswik, E. (1957). Scope and aspects of the cognitive problem. *The essential Brunswik: Beginnings, explications, applications*, 300-312.

Criss, A. H., & Shiffrin, R. M. (2004). Interactions between study task, study time, and the low-frequency hit rate advantage in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 778.

Dallenbach, K. M. (1913). The relation of memory error to time interval. *Psychological Review*, *20*, 323.

Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review*, *108*, 452.

Roediger III, H. L., & DeSoto, K. A. (2014). Confidence and memory: Assessing positive and negative correlations. *Memory*, *22*, 76-91.

Dougherty, M. R., Franco-Watkins, A. M., & Thomas, R. (2008). Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review*, *115*, 199.

Dougherty, M. R., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, *106*, 180.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. New York, NY: Sage

Ebbinghaus, H. (1885). *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot.

Erdfelder, E., Küpper-Tetzel, C. E., & Mattern, S. D. (2011). Threshold models of recognition and the recognition heuristic. *Judgment and Decision Making*, *6*(1), 7.

Estes, W. K. (1994). *Classification and Cognition*. New York: Oxford University Press.

Gibson, J. J. (1979). The ecological approach to visual perception, New York: Houghton Mifflin.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*. 12, 8-20.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: the recognition heuristic. *Psychological Review*, *109*, 75.

Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency heuristic: a model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1191.

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the “coordinates of truth”. *Perspectives on Psychological Science*, *6*, 253-271.

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, *118*, 115.

Lohnas, L. J., & Kahana, M. J. (2013). Parametric effects of word frequency in memory for mixed frequency lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1943.

Malmberg, K. J., & Murnane, K. (2002). List composition and the word-frequency effect for recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 616.

Manson v. Brathwaite, 432 U.S. 98,97 S.Ct. 2243 (1976).

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: an ecological model of strategy selection. *Psychological Review*, *118*, 393.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effect of experience in recognition memory. *Psychological Review*, 105, 724-760.

Neil v. Biggers, 409 U.S. 188 (1972).

Neisser, U. (1976). *Cognition and Reality*. San Francisco: W.H. Freeman.

Newell, B. R., & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, *19*, 333.

Pohl, R. F., Erdfelder, E., Michalkiewicz, M., Castela, M., & Hilbig, B. E. (2016). The limited use of the fluency heuristic: Converging evidence across different procedures. *Memory & Cognition*, *44*, 1114-1126.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145-166.

Smith, V. L., Kassin, S. M., & Ellsworth, P. C. (1989). Eyewitness accuracy and confidence: within-versus between-subjects correlations. *Journal of Applied Psychology*, *74*, 356.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207-232.