CrossMark

One mirror effect: The regularities of recognition memory

Andrew Hilford [1] · Murray Glanzer [1,2] · Kisok Kim [3] · Laurence T. Maloney [1,2]

Abstract
The mirror effect is a pattern of results generally found in two-condition recognition memory experiments that is consistent with normative signal detection theory as a model of recognition. However, the claim has been made that there is a distinct mirror effect, the "strength mirror effect," that differs from the normative one. This claim is based on experiments on recognition memory in which repetition or study time is varied to produce differences in accuracy, where typically the ordinary mirror effect pattern is absent. If this claim is correct, it has major implications for theories of recognition memory. Therefore, a full examination of the data that support the claim was called for. To do that, we replicated the basic demonstration of the no-mirror-effect data and analyzed it further in a series of experiments. The analysis showed the following: (1) Whether or not the mirror effect occurs is determined by whether the experimenter furnishes effective discriminanda that distinguish the weak and strong conditions for the participant. (2) Once Finding 1 is taken into account, no adjustments of or additions to the normative signal detection theory explanations are necessary. (3) There is only one mirror effect, and no separate "strength mirror effect."

Keywords Mirror effect · Signal detection theory · Likelihood ratio · Recognition memory

In many recognition memory experiments, the participant is asked to discriminate between previously studied items (*old*) and new items (*new*). The task is a yes–no signal detection task (Green & Swets, 1966/1974), and human performance is typically modeled within the framework of signal detection theory. The presented item—old or new—generates a strength variable X that is drawn from either the old or the new distribution (Fig. 1). The observer's task is, given X, to decide whether to respond "old" or "new."

There are four possible outcomes (Green & Swets, 1966/1974):

- False alarm (FA), when the item is new and the participant responds "old";
- Hit (H), when the item is old and the participant responds "old";

- Correct rejection (CR), when the item is new and the participant responds "new";
- Miss (M), when the item is old but the participant responds "new."

These outcomes are sometimes paired with explicit payoffs—reward or punishment.

The dependent measures in an old/new recognition task are the conditional probabilities of each outcome—for example, $P[\text{FA} \mid \text{new}]$, the probability of an FA when a new item is presented. Given estimates of these measures, the experimenter can reconstruct the placement of the two distributions, as shown in Fig. 1 (Green & Swets, 1966/1974; Maloney & Zhang, 2010). (Note that the figure is plotted so that the distributions are normalized, to permit easier comparison of the means of the distributions along the x-axis.)

In many experiments there are two conditions, each with a separate signal detection task. The performance in a more discriminable, stronger condition (S) is then compared with the performance in a weaker condition (W). Now there are four distributions—SO, SN (the stronger conditions), WO, and WN (the weaker ones)—and we illustrate their placement on the strength axis in Fig. 2A. Along the strength axis, there is a separation of the studied (old) items. In this interpretation of the theory, the strong old items (e.g., low-frequency words) appear farther to the right on the strength axis than the weak

✉ Andrew Hilford
ah13@nyu.edu

1 Department of Psychology, New York University, New Yotablerk, NY, USA

2 Center for Neural Science, New York University, New York, NY, USA

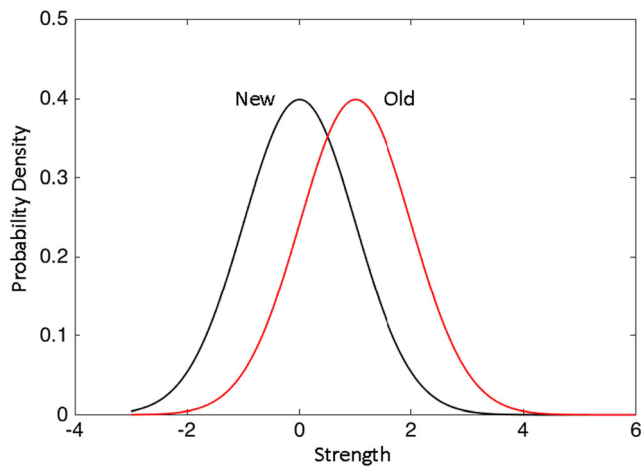3 Department of Psychology, Iona College, New Rochelle, NY, USA

Fig. 1 Signal detection theory account of an old–new recognition memory experiment in which the participant attempts to discriminate previously presented items (old) from items nor previously presented (new). In the signal detection model, each stimulus presented corresponds to a strength variable X taken from one of two distributions, old and new. On the basis of this evidence, the participant must judge whether the stimulus is old or new. See the text for details

old items. Since neither the strong or the weak new items have been studied, a single distribution represents their placement along the strength axis.

An intriguing pattern that is commonly found is the *mirror effect*: In the strong condition, it is easier both to identify new items as new and to identify old items as old than in the weak condition. This pattern corresponds to the inequalities

$$FA_{SN} < FA_{WN} < H_{WO} < H_{SO}$$

We use a transparent notation for such inequalities, in which each term refers to the proportion of "yes" responses (either FAs or hits):

$$SN < WN < WO < SO$$

This pattern is represented in Fig. 2B. An example of the mirror effect is in Hilford, Maloney, Glanzer, and Kim (2015), Experiment 3, with S denoting familiar names and W unfamiliar names:

$$SN = .20 < WN = .30 < WO = .64 < SO = .78$$

Many other examples of the mirror effect can be found in the literature (Glanzer & Adams, 1985; Glanzer, Hilford, & Maloney, 2009). It is produced by varying any of the following variables: concreteness, delay, encoding, familiarity of names, familiarity of faces, familiarity of scenes, list length, neighborhood, number of words similar to the presented word, orthographic similarity, pictures versus words, repetition, semantic similarity, similarity of targets and lures, study time, transformation, or word frequency.

For some of the variables, such as the familiarity of names, the difference is intrinsic to the stimuli themselves. The experimenter can then simply select a number of familiar names (S) and a number of unfamiliar names (W) for the study list. These appear in the subsequent test list as the old items. The experimenter also selects another set of familiar and unfamiliar names to serve as new items in the test list. These selections allow the mirror effect to appear with a single test list because the list contains old familiar names (SO), new familiar names (SN), old unfamiliar names (WO), and new unfamiliar names (WN).

Three of the variables listed above—repetition, study time, and encoding—differ from the others, in that the strong and weak items are not intrinsically different. Instead, they differ in terms of number of repetitions, or study time, or the encoding chosen by the experimenter. The stimuli in each condition are typically marked differently, usually by color, during both study and test. Of course, the new items at test have never been seen before and are not differentiated. At test,
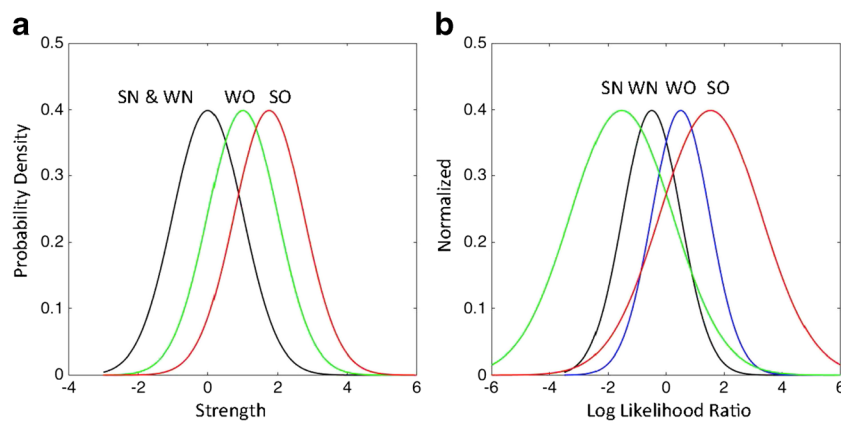

Fig. 2 The SDT-λ view of two-condition recognition memory. The information available on a single trial is represented by a random variable X, as in panel A. In SDT-λ, the SN, WN, WO, and SO distributions of X in panel A are plotted versus the log likelihood ratio of the corresponding condition, strong or weak. That is, SN is plotted

versus the log likelihood ratio for the strong condition, WN is plotted versus the log likelihood ratio for the weak condition, and so forth. We have normalized the distributions in panel B to have the same maxima, in order to facilitate comparison of the locations of the maxima

the new items are randomly assigned to two sets (SN and WN) and arbitrarily marked.

For any of these variables we can check for a mirror effect, and although the mirror effect is often found, it sometimes fails, particularly when the variable differentiating strong and weak is one of those just singled out. Stretch and Wixted (1998), for example, carried out recognition experiments in which a mirror effect was expected to appear. It did not. Since that regularity has an important role in the evaluation of recognition theories (Glanzer et al., 2009), that negative result has generated a large number of subsequent studies. The impact of these findings is summarized in the following quotation from Criss (2009):

> Stretch and Wixted (1998) suggested a fundamental difference between the underlying cause of strength based (e.g., repetition during a study list) and frequency based (e.g., normative word frequency) mirror effects. Namely, they attributed the frequency based mirror effect to differences in the underlying distributions and the strength based mirror effect to a criterion shift.

The Stretch and Wixted findings led to several other studies (Morrell, Gaitan, & Wixted, 2002; Starns, White, & Ratcliff, 2010; Verde & Rotello, 2007). They also gave rise to studies that specified procedures to counter the mixed-list-strength-no-mirror effect (Bruno, Higham, & Perfect, 2009; Franks & Hicks, 2016; Hicks & Starns, 2014; Starns & Olchowski, 2015) and to differentiation models that attribute the word frequency mirror effect and the strength-based mirror effect to different theoretical mechanisms (though not to the same mechanisms as in Stretch & Wixted, 1998).

The idea that there two different mirror effects has become an accepted part of the literature. This can be seen in the widespread use of the qualifier "strength" in the term "strength mirror effect." Whereas the original mirror effect is a simple consequence of normative signal detection theory (Glanzer et al., 2009), this new effect seems to require changes or additions to the normative theory—specifically, to the decision rule employed.

Before accepting the idea of two different mirror effects and two different decision rules, we will reexamine the original finding, the disappearance of the mirror effect. To do that, we first present background information on the mirror effect, how it is produced, and its relation to theory. Then we show how it can be made to disappear and to reappear.

The key issue will prove to be how participants set criteria: the decision rule they use to select a response. The strength mirror effect presupposes an alternative decision rule. Consequently, we begin with a brief summary of how criteria are selected in normative signal detection theory (SDT) and in an alternative SDT that differs from normative in how the criteria are chosen.

Signal detection theory—Two versions

The two different versions of SDT we will consider both apply to two-condition recognition experiments and both share the same common statistical framework. They differ only in the decision rule used to decide the participant's response—specifically, the choice of a criterion. The first version (Glanzer, Adams, Iverson, & Kim 1993; McClelland & Chappell, 1998; Osth, Dennis, & Heathcote, 2017; Shiffrin & Steyvers, 1997) assumes that decisions are made on the basis of likelihood ratios (see below) and that the likelihood ratios are compared to criteria chosen to maximize the expected value. This is a normative form of SDT that maximizes expected value (Green & Swets, 1966/1974), and we label it *SDT-λ*.

The other version, which we label *SDT-s*, assumes that recognition decisions are based on familiarity or strength. It represents an approach that is commonly assumed by researchers of recognition memory (Balota, Burgess, Cortese, & Adams, 2002; Cary & Reder, 2003; Franks & Hicks, 2016; Hicks & Starns, 2014; Higham, Perfect, & Bruno, 2009; Hockley & Niewiadomski, 2007; Morrell et al., 2002; Singer, Fazaluddin, & Andrew, 2012; Singer & Wixted, 2006; Starns & Olchowski, 2015; Stretch & Wixted, 1998). There are competing models of how a criterion is selected for any recognition memory task. See Glanzer et al. (2009) for a comparison and evaluation of SDT-λ versus SDT-s as models of recognition memory.

Figure 2B presents the SDT-λ view of two-condition recognition memory. The information available on a single trial is represented by a random variable X, as in Fig. 2A. The distribution of X is $f_O(x)$ when the item is old (O), and $f_N(x)$ when it is new (N) for each of the conditions. If X is a continuous random variable, then $f_O(x)$ and $f_N(x)$ are probability density functions.

Given X, the likelihood ratio (LR) for "old" over "new" responses is

$$L(X) = f_O(X)/f_N(X),$$

a measure of the strength of evidence in the data favoring "old" over "new." In Fig. 2B we plot the LR functions corresponding to the distributions in Fig. 2A. In SDT-λ the participant's response is determined by the rule

$$L(X) > \beta, \text{ if and only if "respond old."}$$

where β is the likelihood criterion. The rule

$$\log L(X) > \log \beta, \text{ if and only if "respond old"}$$

is equivalent, and it is sometimes convenient to work with the log likelihood ratio $\lambda = \log L(X)$.

At first glance, the decision rule just described is just another comparison of evidence, but one based on

likelihoods. However, the particular rule chosen here differs from other comparison mechanisms. In normative SDT, the criterion β (also referred to as "bias") is determined by the prior odds of old versus new and the values (reward and penalty) assigned to the four possible outcomes (Green & Swets, 1966/1974); we will illustrate its computation in Experiment 2. This choice of criterion gives the unique decision rule with the greatest expected value. The exact computation of β is less important than the facts that the choice of criterion is completely determined by the structure of the signal detection task and that any deviation from this rule reduces the participant's expected value.

Figure 3 represents the SDT-s view of two-condition recognition and the mirror effect. (The same plot may be found in Bruno et al., 2009, Fig. 2; Criss, 2009, Fig. 1; Higham et al., 2009, Fig. 4; Singer et al., 2012, Fig. 1; and Verde & Rotello, 2007, Fig. 2.) The distributions are the same as those in Fig. 2A for SDT-λ given earlier.

The distributions of SN, WN, WO, and SO are arrayed as shown on a strength decision axis. SO is placed to the right of WO, representing greater accuracy. SN and WN are not separated. When, for example, repetition is the variable that separates SO from WO, the new items, because they have not been studied, or treated differently, cannot differ in strength.

To obtain the mirror effect in SDT-s, we must add the following assumption: There are two criteria, C_W and C_S, that individually follow the means of the WO and SO distributions, so that the S criterion, C_S, is higher (farther right) than the W criterion, C_W. This arrangement is depicted in Fig. 3.
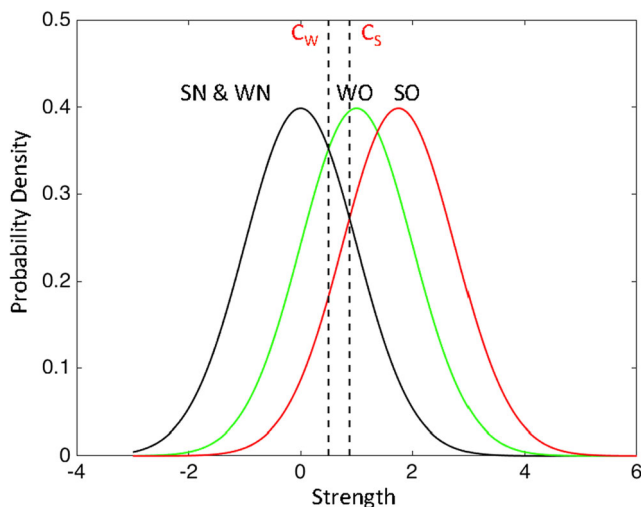

Fig. 3 The SDT-s view of two-condition recognition and the mirror effect. The distributions of SN, WN, WO, and SO are the same as those in Fig. 2A. They remain, however, arrayed, as shown, on a strength decision axis. The mirror effect is obtained by the placement of criteria C_W and C_S at the WN–WO and SN–SO intersections, respectively
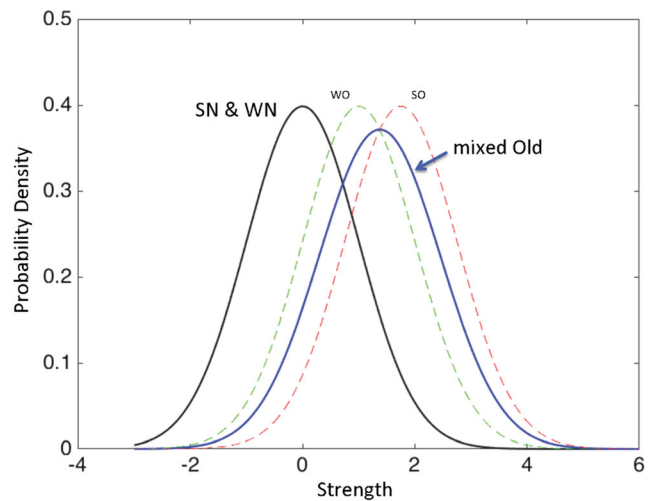

Fig. 4 The SDT-λ model for the situation in which the participant ignores the experimenter's discriminanda. The dashed curves are the WO and SO distributions that the experimenter presents. The blue curve is a mixture of these two curves. "Mixed old" is the distribution a participant experiences if he or she ignores the discriminanda

The areas to the right of the criteria in SO and WO define hit rates. The areas to the right of the criteria in SN and WN define the FA rates. This arrangement produces the full mirror pattern, but only with the ad-hoc assumptions concerning multiple criteria and their placements.

Paradigms

Three experimental paradigms can be used to produce the mirror effect. One, called *mixed-list* or *within-list*, has items from both levels of the variable (e.g., both low- and high-frequency words) in both the study and test lists. A single group of participants is involved. Examples of mixed lists are in Glanzer and Adams (1985). A second, called *pure-list* or *between-list*, has one level of the variable in one study and test list, the other level in a second study and test list. One or two groups of participants may be involved. An example of pure lists with two groups of participants, one given repeated study words, the other given unrepeated study words, can be found in Criss (2006), Experiment 1. A third paradigm has a mixed study list followed by a pure (strong or weak) test list— for instance, Hirshman (1995) and Verde and Rotello (2007). This paradigm is sometimes referred to as *mixed*. We refer to it as a *mixed–pure paradigm*.

Most work on the mirror effect has been done with mixed lists. Experimenter-imposed variables, such as repetition and study time, however, are ordinarily studied in pure lists (Ratcliff, McKoon, & Tindall, 1994, Exps. 1 and 2; Ratcliff, Sheu, & Gronlund, 1992, Exps. 1, 2, and 3; Starns, Ratcliff, & White, 2012, Exps. 1 and 2). To study them in mixed lists requires the addition of discriminanda to indicate to the participant whether a presented test item belongs to the weak or

strong condition. Stretch and Wixted (1998) did that in their Experiment 4, with different colors for the repeated versus unrepeated conditions.[1]

Other investigators—for example, Morrell, Gaitan, and Wixted (2002)—have used semantic categories to discriminate weak and strong conditions. Thus, in the study list, names of living things may be repeated and nonliving things unrepeated. The studies using semantic categories as discriminanda find the same effects as those with color discriminanda.

In recent articles (Glanzer et al., 2009; Hilford et al., 2015) we have demonstrated support for the proposition that SDT as fully presented in Green and Swets (1966/1974, pp. 26–28, 36–40, 153)—with its assumption of a likelihood ratio decision axis, namely SDT-λ—produces the mirror effect and other regularities. Note that the definition of the effect given earlier contains the phrase "two sets of items or conditions." There is no exemption for conditions such as repetition.

The mixed-list-strength-no-mirror effect

The idea of different kinds of mirror effects was introduced by Stretch and Wixted (1998). In their Experiment 4, uncorrelated condition, they used repetition as a variable in a mixed-list paradigm with color (green vs. red) discriminating the two levels of the variable. The key result was the absence of a mirror effect. The results for words in repeated (S) and unrepeated (W) study conditions were

$$SN = .195 = WN = .180 < WO = .578 < SO = .821,$$

with SN not statistically significant different from WN, $F < 1$. Mirror order would require SN < WN. We will call this a *mixed-list-strength-no-mirror* effect. Stretch and Wixted proposed two classes of mirror effect: a strength mirror effect and a frequency mirror effect.[2]

The distinction has been accepted by a large number of investigators, as indicated by their use of the term "strength mirror effect." It has been supported further by investigations that have replicated the mixed-list-strength-no-mirror effect (Bruno et al., 2009, Exp. 1; Criss, 2006, Exp. 2; Higham et al., 2009, Exp. 2; Morrell et al., 2002, Exps. 1–3; Singer

et al., 2012, Exps. 1 and 2; Starns & Olchowski, 2015, Exps. 1b and 2 [two-key]; Verde & Rotello, 2007, Exps. 1–4).

The disappearance of the repetition mirror effect presents a problem. This occurs only when strength is varied along such variables as repetition and, with those variables, only in mixed lists. Pure-list experiments with repetition as the variable show the standard mirror effect:

$$SN < WN < WO < SO,$$

as in Criss (2006), Experiment 1. Moreover, mixed-list experiments with other variables also show the standard mirror effect.

Some investigators (e.g., Morrell et al., 2002) have argued that the mixed-list-strength-no-mirror effect offers a special problem for SDT-λ models. The effect, however, presents a problem for both versions of signal detection theory, SDT-s and SDT-λ. Both have to explain why the mirror effect disappears only for strength variables such as repetition, not for stimulus condition variables such as familiarity of names, and only in mixed lists.

SDT-s theorists (Stretch & Wixted, 1998) explain the difference between the mixed-list and pure-list findings as resulting from differences in the way participants set criteria on a strength decision axis, the axis in Fig. 3. In pure lists, the criteria differ. In mixed lists, they collapse into a single criterion, like the leftmost criterion in Fig. 3. However, no rationale has been established for the assumed difference in the setting of criteria. Other variables—for example, word frequency—do not show such a collapse in mixed lists. Both versions of SDT are incomplete if they do not explain the mixed-list-strength-no-mirror effect. We therefore decided to examine the basis for this effect, starting with one of its initial demonstrations.

The key issue we addressed is the mirror effect and the implications of its absence in certain experimental designs. As we noted above, many studies have now demonstrated the lack of a mirror effect in a particular experimental design and then proposed supplementary mechanisms for setting criteria. Our conclusion—first of all—was that the results obtained so far are consistent with normative SDT models of recognition memory, and no supplementary mechanisms are needed. Still, as part of our argument, we needed to test previous claims about supplementary mechanisms.

Experiments 2–4 did just that. The first two were used to rule out one common explanation of the mixed-list-strength-no-mirror effect. Experiment 4 illustrated the importance of discriminanda in producing the mixed-list-strength-no-mirror effect. The experiments are important precisely because they challenge commonly accepted models. Experiment 1 replicated Stretch and Wixted's (1998) original result and served as a baseline for interpretation of the data of the experiments that followed. Experiments 2 and 3 tested a possible explanation of

---

[1] In three of their experiments, Stretch and Wixted (1998) used a fourth paradigm that disrupted the mirror effect: They pitted repetition against semantic category strength. During study, high-frequency words (the weak category) were repeated, and low-frequency words (the strong category) were not. This increased the hit rate of the weak condition and caused the d' difference between the weak and strong conditions to disappear or become slight. The appearance of a mirror effect requires, however, a significant accuracy difference between conditions. This paradigm has not been used by subsequent investigators.

[2] A more appropriate term for this class is *stimulus condition effect*, since it should cover other variables that do not show the mixed-list-strength-no-mirror effect: familiarity, pictures versus words, word frequency, and so forth.

the mixed-list-strength-no-mirror effect, framed within SDT-s (that criteria are set differently in pure than in mixed lists). Experiment 4 tested the alternative SDT-λ explanation of that effect. Experiments 5 and 6 furnished further support for the SDT-λ explanation.

Experiment 1: Replication of the mixed-list-strength-no-mirror effect

Method

As a first step, we replicated a paradigm that had eliminated the mirror effect. Two sets of words were intermixed in a single study list. The study words from one set were each presented one time. The study words from the other set were each presented three times. The two sets were distinguished by color (red or green) during both study and test. During test they were intermixed with two sets of lures, also discriminated by those colors.

The study and test lists were randomly created (uniquely for each participant) from the main list (of 259 words). Each study list contained 120 unique words, randomized. Of these, half (60) were randomly selected to be repeated. Each test list, also randomized, consisted of 240 words, half old and half new. Color, red or green, was also assigned at random to each participant's single versus repeated conditions.

Materials The study and test words were drawn at random from a list of 259 nouns selected from Kučera and Francis (1967). The mean frequency was 200.29 words per million, and the mean word length was 5.92 letters.

Procedure Participants were first given a practice session, a six-word study list followed by a 12-word test list. This was followed by the main list, in which each word was presented to participants in one of two colors, red or green. The words were arbitrarily placed into a "strong" (repeated) and "weak" (unrepeated) conditions. Each condition was identified by one of the two colors. In the strong condition, each individual study word was shown three times. The words in the weak condition appeared only once. The lists were uniquely randomized for each participant. Two buffer words were presented at the beginning and end of each study list and at the beginning of each test list. These buffer words were not included in the analysis of the results.

The study words each appeared one at a time for 1,250 ms in the center of the screen, with a 750-ms interval between words. The participants were instructed to study the words carefully and that some of them would be repeated. Participants were further instructed that a test list would follow the study list. The test words also appeared, participant-paced, one at a time in the center of the screen. Participants were instructed to determine whether each word was old or new. They indicated their decisions ("old" or "new") as well as the confidence of their choice. Only the old/new decisions are considered here.

Participants Twenty-four undergraduates participated in order to fulfill a class requirement. All had been speaking English since the age of 10 years or earlier.

The descriptions above of materials, method, procedure, and participants hold for the following experiments unless indicated otherwise.

Results

The results, shown in the first line of Table 1, replicate the Stretch and Wixted (1998) *mixed-list-strength-no-mirror effect*. The hit rates differ at a statistically significantly level, $F(1, 23) = 25.86$, $MSE = .350$, but the FA rates do not, $F(1, 23) = 0.105$, $MSE = 0.0003$. Here and subsequently, we adopt an alpha level of .05 for statistical significance. We also shorten the phrase "statistically significant" to "significantly."

Discussion

The paradigm for producing the mixed-list-strength-no-mirror effect was replicated. The characteristic finding of the mixed-list-strength-no-mirror effect was also replicated: a difference in hit rates and no difference in FA rates.

Corresponding to that pattern of hit and FA rates are the bias measures, β, shown in the first line of Table 2. A β of 1 indicates no bias; a β less than 1 indicates a liberal bias, or a greater tendency to say "yes" to both new and old words; and a β greater than 1 indicates a conservative bias, or a reduced tendency to say "yes" to both new and old words. The β values for Experiment 1 indicate little or no bias for the single condition and a considerable liberal bias for the repeated condition.

Table 1 Mirror effect pattern in all six experiments

Exp.	SN	WN	WO	SO
1	.3361	.3410	.6028[*]	.7736[*]
2	.3381	.3333	.5968[*]	.8111[*]
3	.3450	.3580	.6050[*]	.8050[*]
4	.3491	.3430	.6342[*]	.8289[*]
5	.2492[*]	.2935[*]	.7162[*]	.7757[*]
6	.1714[*]	.2089[*]	.7036[*]	.9375[*]

SN is the proportion of false alarms in the strong (e.g., repeated) condition. WN is the proportion of false alarms in the weak (e.g., unrepeated) condition. SO is the proportion of hits in the strong condition. WO is the proportion of hits in the weak condition. [*] Asterisks indicate a statistically significant difference in paired hits (SO vs. WO) and paired false alarms (SN vs. WN).

Table 2 Bias indices, β, for Experiments 1–4

Exp.	Payoff		Feedback	
	Single	Repeated	Single	Repeated
1	1.05	0.84		
2	1.07	0.74	1.00	2.00
3	1.03	0.76	0.20	1.00
4	1.03	0.68		

The first two columns indicate the bias measures for each experiment, based on Gardner, Dalsing, Reyes, and Brake's (1984) tables. The second two columns indicate the bias pressure in the feedback to Experiments 2 and 3.

Experiment 2: Attempt to recover the mirror effect with payoff and feedback

When participants' patterns of results show the mixed-list-strength-no-mirror effect, as in Experiment 1, they necessarily show bias: They say "yes" more often to repeated than to unrepeated words. This bias, depicted in SDT-s as the two criteria C_W and C_S collapsing into a single position, has been interpreted as the cause of disruption of the mirror effect. (See Fig. 3.) We now attempted to move the criteria and remove the bias by providing the participants with payoffs for responses to the two classes of words. We used payoffs to counter bias, thus canceling the observed bias difference. If this reasoning is correct then—once we had countered the bias—the mirror pattern should reappear. This kind of payoff arrangement (see Healy & Kubovy, 1978) had a strong effect in Hilford et al. (2015) in countering bias.

Method, materials, and procedure

These were the same as in Experiment 1, except for the addition of payoff and feedback.

Participants were told that they would receive scores for each response. The scores were given on the basis of the schedule in Table 3. The βs associated with the two payoff schemes, listed in Table 2, should pressure the participants to maintain the same response pattern in the single condition but to become less liberal in the repeated condition.

Table 3 Payoff schedule for Experiment 2

Presentation	Response			
	Hits	Miss	Correct Rej.	False Alarm
Single	30	− 10	10	− 30
Repeated	10	− 10	10	− 30

After each test response, the score for that response (30, − 30, 10, or − 10) and a running total score were shown for 1,250 ms. This was also done during the initial practice section.

The computation of the SDT-λ criterion β for any choice of payoffs is

$$\beta = [V(CR)–V(FA)]/[V(H)–V(M)]$$

(Green & Swets, 1966/1974). For example, in the repeated condition of Table 2, the payoffs are

$$V(CR) = 10, V(FA) = –30, V(H) = 10, V(M)$$
$$= –10, \text{ and criterion } \beta = 2.00.$$

Participants Twenty-one undergraduates took part in the experiment.

Results and discussion

The results, shown in the second line of Table 1, show the mixed-list-strength-no-mirror effect. The hit rates differ significantly, $F(1, 20) = 75.419$, $MSE = 0.482$, but the FA rates do not, $F(1, 20) = 0.084$, $MSE = 0.0002$. The mixed-list-strength-no-mirror effect remains. The feedback had no effect.

Experiment 3: Further attempt to recover the mirror effect by payoff and feedback

We repeated the experiment, this time drawing participants' attention to the payoff arrangement in the instructions.

Method

The materials, procedure, and characteristics of the participants were the same as in Experiment 2, except that, before the test, participants were given, in addition, a detailed description of the payoff schedule. The payoff schedule was changed in order to bring bias in the single and repeated conditions to equality by making the responses in the single condition more liberal, as in the repeated condition, and the repeated condition responses less liberal. The key scores 50 and − 50 were made larger than the values of 30 and − 30 in Experiment 2, to increase the effect of the feedback.

Of the 20 participants, 12 had the feedback scores in the table. When it became apparent that the feedback was not having any effect, the scores 50 and − 50 were increased to 100 and − 100. Those scores also did not change the bias.

The payoff arrangement is indicated in Table 4.

Table 4 Payoff schedule for Experiment 3

Presentation	Response			
	Hits	Miss	Correct Rej.	False Alarm
Single	100	− 100	10	− 10
Repeated	10	− 10	10	− 10

Participants Twenty undergraduates took part in the experiment.

Results

The results, shown in the third line of Table 1, replicated the mixed-list-strength-no-mirror effect. The hit rates differed significantly, $F(1, 19) = 57.00$, $MSE = 0.400$, but the FA rates did not, $F(1, 19) = 0.0003$, $MSE = 0.043$, $p = .839$. The βs in line 3 of Table 2 are the same as those for Experiments 1 and 2. The mixed-list-strength-no-mirror effect was unchanged.

Discussion

An alternative explanation of the results of Experiments 2 and 3 is that participants may not have been motivated to use the feedback or payoff schedule effectively. We believe that this explanation is not tenable, given past results in the literature. There is considerable evidence that feedback and payoff schedules can affect criterion in recognition memory tasks. Hilford et al. (2015), Exp. 2 found that a procedure identical to that used here produced a pattern like that in the mixed-list-strength-no-mirror effect: SN = WN < WO < SO. However, Hilford et al.'s Experiment 3 then used payoff and feedback to demonstrate that changing the bias (criterion location) for one condition resulted in reappearance of the mirror effect: SN < WN < WO < SO. That payoff procedure was identical to the one used in Experiment 3 of the present article. It is implausible that participants would have been unmotivated in one experiment but motivated in the other.

We also note that Verde and Rotello (2007) showed that the pure test blocks in their Experiments 1–4 gave the mixed-list-strength-no-mirror effect, but the addition of feedback resulted in reappearance of the mirror effect. Moreover, Experiment 2 of Hilford et al. (2015) produced a pattern like that in the mixed-list-strength-no-mirror effect—namely, SN = WN < WO < SO. Hilford et al.'s Experiment 3 then demonstrated that payoff and feedback identical to those used here recovered the mirror effect: SN < WN < WO < SO. The fact that payoff and feedback did not affect the criteria in our Experiments 2 and 3 and did not counter the mixed-list-strength-no-mirror effect implies that the placement of criteria is not the explanation of that effect.

Reconsideration of the problem SDT-s explains the mixed-list-strength-no-mirror effect as a result of the two criteria in Fig. 3 being collapsing to a single position. Differential payoff and feedback should have moved them apart. Experiments 2 and 3, however, showed the criteria immovable. The effect was not due to placement of the criteria.

Two studies in which procedures were used to recover the mirror effect in mixed lists suggest an explanation. Franks and Hicks (2016) obtained the mirror effect with mixed lists by strengthening the discriminanda between the strong and weak conditions. They used both color and position. Starns and Olchowski (2015) forced participants to attend to the discriminanda by using a three-response procedure during test (strong old, weak old, new), instead of the usual two-response procedure (old, new). The explanation of the mixed-list-strength-no-mirror effect may lie in the participants' not attending to the discriminanda (red vs. green) between the strong and weak conditions. In that case, Figs. 2 and 3 describe the experimenter's view of the experimental situation, but not the participants'.

We can set up the model for such a case in which the participants' and the experimenter's views differ.

Model for participant–experimenter difference Let us assume that SDT-λ in Fig. 2B or SDT-s in Fig. 3 represents the experimenter's view of the input to Experiment 1. Let us assume, however, that the participant is blind to the discriminanda that the experimenter has furnished. The model for this case, and its implications, are presented in Fig. 4.

The dashed curves labeled WO and SO represent the experimenter's view of the experiment: Strong and weak old stimuli are interleaved. If, however, the participant ignores the discriminanda, then he or she treats the interleaved stimuli as drawn from a single distribution, labeled "mixed old." This distribution is just the weighted mixture of WO and SO with weights .5 each. Of course, on the strength axis, the two new distributions coincide, but now they also coincide on the log likelihood ratio axis, as well, since—from the participant's viewpoint—there is only one log likelihood ratio distribution.

Using the same distributions as those used earlier for Fig. 2A, but including the participants' collapsing of the four distributions into two, we can compute the output of the collapsed model. The hits and FAs for this model show the mixed-list-strength-no-mirror effect.

$$SN = .308 = WN = .308 < WO = .691 < SO = .841$$

The hits, SO and WO, differ because the repeated words are higher within the participants' combined old distribution. The experimenter partitions the words in the participants' single new distribution into those the experimenter (but not the participant) knows were repeated and those that were not. The lures, SN and WN, cannot be

different when (for the participant) there is only a single old distribution.

Experiment 4: Producing the mixed-list-strength-no-mirror effect by eliminating discrimination of the underlying distributions

We have shown with a theoretical model that the mixed-list-strength-no-mirror effect is produced when the participant ignores the experimenter's discriminanda and works with just two distributions instead of the experimenter's four. We will now strengthen that argument with an empirical demonstration. We repeated Experiment 1 with one change: The discriminating color was eliminated. All words, both repeated and unrepeated, appeared colored black at both study and test. If the model and the implicit explanation are correct, then the mixed-list-strength-no-mirror effect should appear.

Method

The materials, procedure, and characteristics of the participants were the same as in Experiment 1, except that the colors discriminating the repeated and unrepeated words were absent. This corresponds to the situation in which the participant ignores, or is blind to, the cues that discriminate the two classes of items.

Participants Nineteen undergraduates took part in the experiment.

Results

The results, shown in the fourth line of Table 1, replicate the mixed-list-strength-no-mirror effect. The hit rates differ significantly, $F(1, 18) = 63.257$, $MSE = 0.360$, but the FA rates do not, $F(1, 18) = 0.116$, $MSE = 0.0004$, $p = .737$. The βs on line 4 of Table 2 also replicate the results of the preceding experiments: The single condition is unbiased, $\beta = 1.03$, whereas the repeated condition shows a liberal bias, $\beta = 0.68$.

These results reproduce the results of the preceding experiments and of mixed-list-strength-no-mirror-effect experiments generally. They support the proposition that the mixed-list-strength-no-mirror effect results from participants not discriminating the distributions furnished by the experimenter.

Discussion

The experimenter assumes that the participant is responding to two distinct sets of items. The participant, however, is responding to only one set. The fact that the participants have

a hit rate to the repeated items does not indicate that they are responding differentially to two separate sets of items. It indicates only that the experimenter assumes that the discriminanda are effective and partitions the responses in a way the participants have not. We will now show further evidence that the mixed-list-strength-no-mirror effect is not general. We will show that it is possible to obtain mirror effects with a strength-mixed list when there is effective discrimination of the underlying distributions.

Experiment 5: Demonstrations of the mixed-list strength mirror effect by effectively discriminating the underlying distributions

It is necessary for the experimenter to differentiate the underlying distributions effectively for participants. Arbitrarily selecting color or semantic category as discriminanda does not work. Those discriminanda are ignored by the participants.

We speculated that participants might find it more efficient to deal with decisions involving two distributions among four rather than with the four distributions in Fig. 2A. We conjectured that such discrimination could not be ignored if the discriminanda were the sensory modalities of presentation: audition versus vision. That discrimination was applied in the following experiments.

Method

The materials, procedure, and characteristics of the participants were the same as in Experiment 1, except that a sense modality (audition vs. vision), not color, discrimination was used to create the two stimulus categories. Words were presented to participants either spoken through headphones or visually on a computer screen. The strong condition was created using word repetition.

Materials The study and test words were drawn at random from a list of 365 nouns selected from Kučera and Francis (1967). The mean frequency was 62.73 words per million, and the mean word length was 6.41 letters.

Participants Thirty-seven undergraduates took part in the experiment.

Results and discussion

The results, in the fifth line of Table 1, showed a mirror effect with strength and repetition varied in a mixed list. The hit rates differed significantly, $F(1, 36) = 10.587$, $MSE = 0.0654$, and—most important, to complete the mirror effect—the FA rates also differed significantly, $F(1, 37) = 6.084$, $MSE =$

0.0363. The claim that the mixed-list-strength-no-mirror effect is a general effect was refuted.

Experiment 6: Demonstrations of the mixed-list strength mirror effect with changes in encoding tasks

Another standard way to strengthen memory of items is through encoding tasks. Requiring participants to do additional processing of words during study strengthens recall and recognition of those words. We used that strengthening method to generalize the findings of Experiment 5.

Method

The materials, procedure, and characteristics of the participants were the same as in Experiment 5, except that a strength difference was created by the encoding task instead of repetition.

As in Experiment 5, words were presented to participants either spoken through headphones or visually on a computer screen. The strong condition was created using an encoding task: evaluating the pleasantness of the word presented. In the weak condition, no encoding task was applied to the word presented. During study, in the strong condition, following each item presentation, participants were asked, on the computer screen, to rate whether the just-presented word was "positive" or "negative." Participants were instructed to indicate their choice using a 6-point rating scale, with 1 being *very positive* and 6 being *very negative*.

Materials and procedure The materials, procedure, and characteristics of participants were the same as in Experiment 5, except for the replacement of repetition by encoding.

Participants Twenty-eight undergraduates took part in the experiment.

Results and discussion

The results, in the sixth line of Table 1, again showed a mirror effect of strength differences in a mixed list. The hit rates differed significantly, $F(1, 27) = 73.473$, $MSE = 0.766$, and the FA rates also differed significantly, $F(1, 27) = 4.553$, $MSE = 0.0197$.

The view that the mixed-list-strength-no-mirror effect is a general effect was again rebutted.

Table 5 shows the d' measures[3] between the means of the four underlying distributions, SN, WN, WO, and SO

---

[3] To reduce notational clutter, we omit the prime in d' when the variable is subscripted by condition. Thus, d_S is the d' measure in the S condition.

Table 5 Distances between the distributions of Experiments 1, 5, and 6

Exp	d_S	d_W	d_{OO}	d_{NN}
1	1.15	0.66	0.48	
5	1.44	1.14	0.19	− 0.12
6	2.50	1.33	1.02	− 0.14

Comparison of the results for Experiments 1, 5, and 6: d_S is the accuracy measure (d') for the strong condition, and d_W is the accuracy measure for the weak condition. The value d_{OO} is the distance between the means of the old distributions, and d_{NN} is the distance between the means of the new distributions.

for the Experiments 1, 5, and 6—allowing us to compare across experiments. One question concerns the relation between Experiments 5 and 6. The value d_S is the accuracy measure (d') for the strong condition, and d_W is the accuracy measure for the weak condition. The value d_{OO} is the distance between the means of the old distributions, and d_{NN} is the distance between the means of the new distributions. Comparison of the values of d_S and d_W in the two experiments shows that participant accuracy was less in Experiment 5 than in Experiment 6. This leads to a second question. According to SDT-λ, the lesser accuracy of Experiment 5 should produce concentering, or a shrinkage of all the distances between the means of the underlying distributions. Did it? Examination of the paired values of d_{NN}, d_{OO}, and d_W does show concentering: The means in the less accurate Experiment 5 are indeed closer together than those in Experiment 6.

We can also use Table 5 to address a potential circularity in our reasoning. Are we assuming that the discriminanda of Experiments 5 and 6 are superior simply because they generate a mirror effect? Experiments 1 and 5 were identical in method, procedure, and undergraduate participants. They differed only in their discriminanda. Comparison of the standard accuracy measures for Experiments 1 and 5, the paired d_S and also d_W, show independently that the disciminanda of Experiment 5 (modality) are *more* effective than the discriminanda of Experiment 1 (color), not less.

In summary, we have shown that bias differences are the reason why the results of certain experiments appear to contradict SDT-λ. FAs are strongly affected by bias and can be poor indications of the positions of the underlying new distributions. This was noted by Hilford et al. (2015), page 1650:

> Glanzer et al. (2009), however, demonstrated that H/FA Mirror Index functions poorly when there are bias effects. In such cases the index may indicate the absence of a Mirror Effect even when the underlying distributions are actually in mirror order. . . . A better index of the Mirror Effect [the distance mirror index] is obtained

Table 6 Experimental summary

Experiment	Discriminanda	Strengthing Variable	Feedback	Mirror Effect
1	Color	Repetition	No	No
2	Color	Repetition	Yes	No
3	Color	Repetition	Yes	No
4	Null	Repetition	No	No
5	Modality	Repetition	No	Yes
6	Modality	Encoding	No	Yes

with more complete measures of the distance between the means of SN and WN and the distance between the means of WO and SO.

A table showing how bias difference affects the appearance or nonappearance of the mirror effect in SDT-λ has been included in an Appendix (Table 7). More appropriate statistics to compare experiments are the d' distances in weak versus strong conditions. Table 5 contains those d's. The table and the text now indicate that d_{NN}, the distance between the means of the new distributions of Experiment 5, is less than that in Experiment 6. Concentering is evident.

Final discussion

These experiments show that there is one mirror effect. That is, the mixed-list-strength-no-mirror effect that was the motivation for a separate "strength mirror effect" is caused by a deficiency in experimental design: selection of discriminanda that participants ignore. We noted earlier that participant may find it more efficient to deal with decisions involving two rather than the four distributions that the experimenter offers and assumes are effective. We have, in SDT, a theory of the way that costs and benefits will affect old/new decisions in a recognition memory task. We do not have a theory of the way that costs and benefits will affect decisions about the number of distributions involved. Table 6 summarizes the sequence of the experiments completed for this article. Discussion of the results in this table follows.

Experiments 1, 2, and 3 replicated the finding that using imposed discriminanda—for example, colors—results in a mixed-list-strength-no-mirror effect. Experiments 2 and 3 also showed that the problems posed by the mixed-list-strength-no-mirror effect cannot be solved by moving criteria, which is the solution offered by SDT-s theorists.

Experiment 4 showed that the discriminanda offered in preceding experiments that had produced the mixed-list-strength-no-mirror effect were, for the participant, irrelevant. The experimenter offered the four distributions in Fig. 2A and scored the responses on the basis of that picture. The participant, however, perceived only the two distributions in Fig. 4 (the repeated items were intermingled with unrepeated items in a single distribution) and responded on the basis of that picture. The model presented for those conflicting views generates a mixed-list-strength-no-mirror effect.

A study by Franks and Hicks (2016) supports the argument that the mixed-list-strength-no-mirror effect results from the experimenter offering participants discriminanda that the participants ignore. In that study, the mirror effect was recovered in mixed lists by introducing additional factors that further discriminated the two classes of stimuli. Franks and Hicks's Experiments 1 and 2 strengthened the discriminanda by using color, screen position, and three response keys (strong old, weak old, and new) combined.

Once it is clear that the participant is working in mixed lists with the two distributions in Fig. 4, not the experimenter's four distributions, other problems are solved. One of these is why payoffs and feedback were ineffective at recovering the mirror effect in Experiments 2 and 3. The payoffs and feedback directed the participant to behave differently to repeated and unrepeated items. However, the participant did not have two such classes of items, only one. The feedback would thus have appeared to the participant as being randomly applied to one class of items, and was therefore irrelevant.

A second problem solved is why pure lists show the mirror effect, whereas mixed lists do not. This occurs because pure lists do not permit the participant to merge strong and weak distributions.

A third problem solved is one posed by the results of Shiffrin, Huber, and Marinelli's (1995) recognition memory study. Their participants studied lists that contained many categories in order to prevent category-specific criterion shifts. Strengthening some categories on the list increased the hit rate for those categories but did not affect the FA rate. This is the mixed-list-strength-no-mirror effect in a different guise. It can thus be understood as involving a disparity between the experimenter's and the participants' view of the underlying distributions.

The mixed-list-strength-no-mirror effect does not require revision of theories of the mirror effect. It does not require the designation of two different kinds of mirror effect, for strength and frequency. Instead, it disappears when the experimenter employs discriminanda that participants respond to, as we did in Experiments 5 and 6.

Appendix

Table 7 Effect of criterion bias on the H/FA index of the mirror effect in the normal unequal-variance model

H/FA Indices					No Mirror Effect?
β	SN	WN	WO	SO	
0.25	0.56	0.26	0.62	0.94	x
0.50	0.33	0.26	0.62	0.86	x
0.75	0.23	0.26	0.62	0.79	
1.00	0.17	0.26	0.62	0.74	
1.25	0.14	0.26	0.62	0.70	
1.50	0.11	0.26	0.62	0.67	
1.75	0.10	0.26	0.62	0.64	
2.00	0.08	0.26	0.62	0.61	x
2.25	0.07	0.26	0.62	0.59	x
2.50	0.06	0.26	0.62	0.57	x
2.75	0.06	0.26	0.62	0.55	x
3.00	0.05	0.26	0.62	0.54	x

A likelihood ratio criterion $\beta = 1$ is unbiased. A likelihood ratio criterion less than 1 is biased ("liberal") and will result in a higher FA rate. A likelihood ratio criterion greater than 1 is also biased ("conservative") and will result in a lower FA rate. In the table the likelihood ratio criterion β is varied in the strong condition (S) as shown, but is held constant in the weak condition (W). An "x" in the last column indicates the absence of a mirror effect. For the mirror effect to hold, we must have SN < WN < WO < SO. With biased criteria much less than 1, SN > WN (i.e., the mirror effect fails). With biased criteria much greater than 1, WO > SO (again, the mirror effect fails).

References

Balota, D. A., Burgess, G. C., Cortese, M. J., & Adams, D. R. (2002). The word-frequency mirror effect in young, old and early-stage Alzheimer's disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language, 46*, 199–226. https://doi.org/10.1006/jmla.2001.2803

Bruno, D., Higham, P. A., & Perfect, T. J. (2009). Global subjective memorability and the strength-based mirror effect in recognition memory. *Memory & Cognition, 37*, 807–818. https://doi.org/10.3758/MC.37.6.807

Cary, M., & Reder, L. M. (2003). A dual-process account of the list length and strength-based mirror effects in recognition. *Journal of Memory and Language, 49*, 231–248.

Criss, A. H. (2006). The consequences of differentiation in episodic memory: Similarity and the strength based mirror effect. *Journal of Memory and Language, 55*, 461–478.

Criss, A. H. (2009). The distribution of subjective memory strength: List strength and response bias. *Cognitive Psychology, 59*, 297–319. https://doi.org/10.1016/j.cogpsych.2009.07.003

Franks, B. A., & Hicks, J. L. (2016). The reliability of criterion shifting in recognition memory is task dependent. *Memory & Cognition, 44*, 1215–1227.

Gardner, R., Dalsing, S., Reyes, B., & Brake, S. (1984). Table of criterion values (*b*) used in signal detection theory. *Behavior Research Methods, Instruments, & Computers, 16*, 425–436.

Glanzer, M., & Adams, J. K. (1985). The Mirror Effect in recognition memory. *Memory & Cognition, 13*, 8–20. https://doi.org/10.3758/BF03198438

Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review, 100*, 546–567. https://doi.org/10.1037/0033-295X.100.3.546

Glanzer, M., Hilford, A., & Maloney, L. T. (2009). Likelihood ratio decisions in memory: Three implied regularities. *Psychonomic Bulletin & Review, 16*, 431–455. https://doi.org/10.3758/PBR.16.3.431

Green, D. M., & Swets, J. A. (1974). Signal detection theory and psychophysics (Rev. ed.). Huntington, NY: Kriege. (Original work published 1966)

Healy, A. F., & Kubovy, M. (1978). The effects of payoffs and prior possibilities on indices of performance and cutoff location in recognition memory. *Memory & Cognition, 6*, 544–553.

Hicks, J. L., & Starns, J. J. (2014). Strength cues and blocking at test promote reliable within-list criterion shifts in recognition memory. *Memory & Cognition, 42*, 742–754.

Higham, P. A., Perfect, T. J., & Bruno, D. (2009). Investigating strength and frequency effects in recognition memory using Type-2 signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 57–80.

Hilford, A., Maloney, L. T., Glanzer, M., & Kim, K. (2015). Three regularities of recognition memory: The role of bias. *Psychonomic Bulletin & Review, 22*, 1646–1664. https://doi.org/10.3758/s13423-015-0829-0

Hirshman, E. (1995). Decision processes in recognition memory: Criterion shifts and the list-strength paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 302–313. https://doi.org/10.1037/0278-7393.21.2.302

Hockley, W. E., & Niewiadomski, M. W. (2007). Strength-based mirror effects in item and associative recognition: Evidence for within-list criterion changes. *Memory & Cognition, 35*, 679–688.

Kučera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.

Maloney, L. T., & Zhang, H. (2010). Decision-theoretic models of visual perception and action. *Vision Research, 50*, 2362–2374.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review, 105*, 724–760. https://doi.org/10.1037/0033-295X.105.4.734-760

Morrell, H. E. R., Gaitan, S., & Wixted, J. T. (2002). On the nature of the decision axis in signal-detection-based models of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 1095–1110.

Osth, A. F., Dennis, S., & Heathcote, A. (2017). Likelihood ratio sequential sampling models of recognition memory. *Cognitive Psychology, 92*, 101–126.

Ratcliff, R., McKoon, G., & Tindall, M. (1994). Empirical generality of data from recognition memory receiver-operating characteristic functions and implications for the global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 763–785. https://doi.org/10.1037/0278-7393.20.4.763

Ratcliff, R., Sheu, C.-F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99*, 518–535. https://doi.org/10.1037/0033-295X.99.3.518

Shiffrin, R. M., Huber, D. E., & Marinelli, K. (1995). Effects of category length and strength on familiarity in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 267–287.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review, 14*, 145–166. https://doi.org/10.3758/BF03209391

Singer, M., Fazaluddin, A., & Andrew, K. N. (2012). Recognition of categorised words: Repetition effects in rote study. *Memory, 21*, 467–481.

Singer, M., & Wixted, J. T. (2006). Effect of delay on recognition decisions: Evidence for a criterion shift. *Memory & Cognition, 34*, 125–137. https://doi.org/10.3758/BF03193392

Starns, J. J., & Olchowski, J. E. (2015). Shifting the criterion is not the difficult part of trial-by-trial criterion shifts in recognition memory. *Memory & Cognition, 43*, 49–59.

Starns, J. J., Ratcliff, R., & White, C. N. (2012). Diffusion model drift rates can be influenced by decision processes: An analysis of the strength-based mirror effect. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, 38, 1137–1151. https://doi.org/10.1037/a0028151

Starns, J. J., White, C. N., & Ratcliff, R. (2010). A direct test of the differentiation mechanism: REM, BCDMEM, and the strength-based mirror effect in recognition memory. *Journal of Memory and Language, 63*, 18–34. https://doi.org/10.1016/j.jml.2010.03.004

Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 1379–1396. https://doi.org/10.1037/0278-7393.24.6.1379

Verde, M. F., & Rotello, C. M. (2007). Memory strength and the decision process in recognition memory. *Memory & Cognition, 35*, 254–262. https://doi.org/10.3758/BF03193446