Evidence for Discrete-State Processing in Perceptual Word Identification

April R. Swagman, Jordan M. Province, Jeffrey N. Rouder

University of Missouri

April Swagman

AprilSwagman@mail.missouri.edu

Abstract

We contrast predictions from a discrete-state model of all-or none information loss with a signal-detection model of graded strength for the identification of briefly flashed English words. Previous assessments have focused on whether ROC curves are straight or not, which is a test of a discrete-state model where detection leads to the highest confidence response with certainty. We along with many others argue this certainty assumption is too constraining, and, consequently, the straight-line ROC test is too stringent. Instead, we assess a core property of discrete-state models, conditional independence, where the pattern of responses depends only on which state is entered. The conditional independence property implies that confidence ratings are a mixture of detect and guess state responses, and that stimulus strength factors, the duration of the flashed word in this report, affect only the probability of entering a state but do not affect responses conditional on a state. To assess this mixture property, fifty participants saw words presented briefly on a computer screen at three variable flash durations followed by either a two-alternative confidence ratings task or a yes-no confidence ratings task. Comparable discrete-state and signal-detection models were fit to the data for each participant and for each confidence-ratings task. The discrete-state model outperformed the signal detection model for 90% of participants in the two-alternative confidence ratings task and for 70% of participants in the yes-no confidence ratings task. We conclude discrete-state models are viable for predicting performance across stimulus conditions in a perceptual word identification task.

Evidence for Discrete-State Processing in Perceptual Word Identification

One of the enduring debates in cognition is whether there are discrete mental states with one state corresponding to guessing or complete information loss. Blackwell (1953), for example, proposed the *high-threshold model* where an accuracy of 75% may arise because on half the trials the participant detects the stimulus (and responds with perfect accuracy) and on the other half she or he guesses (say with 50% accuracy). This notion that on some trials participants have no access to information and must guess without stimulus information is at the core of a few models in perception (e.g. Townsend, 1971), working memory (e.g. Zhang & Luck, 2009), and recognition memory (e.g. Broder & Schutz, 2009), and has been expanded upon in the multinomial processing tree approach of Riefer & Batchelder (1988).

Although models with complete information loss are presented in the literature, they stand as a minority viewpoint. Most models do not have recourse to guessing states, and information is lost more gradually or gracefully. Perhaps the most seminal is the theory of signal detection where the representation of stimuli are represented in a graded fashion without the need for guessing states (Swets, 1961). Other graceful degradation theories include Ratcliff's diffusion model of perception (Ratcliff, 1978; Ratcliff & Rouder, 1998), multidimensional representation models (Ashby, 1992) and Bayesian ideal-observer models (Ma & Huang, 2009; Shams, Ma, & Beierholm, 2005; Sims, Jacobs, & Knill, 2012).

The critical question of this report is whether a perceptual task, the identification of briefly flashed and subsequently masked words, is better described by a discrete-state model of all-or-none loss or a latent-strength signal-detection model of graceful degradation. Although this paper is focused on perception, these issues have inspired a robust and vigorous debate in the recognition-memory literature. Consequently, we borrow from both literatures.

One popular method of testing between latent strength and discrete-state models is assessment through *receiver operating characteristics* (ROCs, Macmillan & Creelman, 2005). The prevailing wisdom is that discrete-state models and latent-strength models predict straight-line and curved ROC isosensitivity curves, respectively (Kinchla, 1994). Almost all empirical ROCs are curved (see Glanzer, Kim, Hilford, & Adams, 1999, for a review in recognition memory), and Figure 1A-B shows this curvature. Consequently, many memory researchers consider discrete-state models untenable (e.g., Dube & Rotello, 2012; Kinchla, 1994; Slotnick & Dodson, 2005; Yonelinas & Parks, 2007). The same types of results are obtained in perception as well, where ROC analysis has been a staple since Tanner & Birdsall (1958) and Egan (1958). In the visual domain, ROCs have been used in conjunction with latent-strength signal-detection models in luminance detection (Carterette & Cole, 1962) and the identification of briefly displayed words (Thierman, 1968; Haase & Fisk, 2001; Jacobs, Graf, & Kinder, 2003), and more complex perception tasks such as medical imaging procedures (Goodenough, Rossman, & Lusted, 1972). Discrete-state models remain unpopular in these applications because ROCs are known to be curved.

This conventional wisdom that discrete-state models are incompatible with curved ROCs has been challenged by two different critiques. One critique, an aggregation critique, is that ROC curves are indeed straight lines but that previous researchers have failed to disaggregate the data from individuals, leading to distorted views of ROC curves (Kellen, Klauer, & Broder, 2013; Pratte & Rouder, 2011; Rouder & Lu, 2005). Results from analyses at the level of individuals have proved contentious and discordant—Klauer & Kellen (2011) and Broder & Schutz (2009) find considerable support for discrete-state models over latent-strength ones while Dube & Rotello (2012) find support for the opposing position. The debate centers on disagreement over methodological details such as what constitutes an appropriate penalty for model complexity and how to manipulate base rates in an ecologically valid manner.

We think, however, that there is a fundamental critique of the conventional wisdom—realistic discrete-state models are not constrained to predict straight-line ROCs. The conventional discrete-state model for confidence ratings is shown in Figure 1C. In this model the detection state leads always to the endorsement of the highest-confidence response option, and we term this specification the *certainty assumption*. Models with the certainty assumption do indeed predict straight-line ROCs. Luce (1963) argued that participants may enter detect states without volition or control, and may distribute their responses rather than act with certainty to meet real or perceived demand characteristics of the experiment such as utilizing all response options. Similar arguments are made by Erdfelder & Buchner (1998), Malmberg (2002), and Province & Rouder (2012) for confidence ratings and by Rouder, Province, Swagman, & Thiele (submitted) for binary choice. Figure 1D shows an appropriate generalization of the discrete-state model without certainty, and, as shown by Rouder et al. (submitted), it can predict curved ROCs.

In fact, the discrete-state model without the certainty assumption can predict any single ROC curve; that is, it is not testable with any one ROC curve from any one condition (Rouder et al., submitted). Fortunately, discrete-state models make strong and testable predictions, but these hold across conditions rather than in any one (Province & Rouder, 2012) as discussed next.

Testable Predictions of Discrete-State Models

The predictions of discrete-state models without certainty are most easily discussed with reference to the experimental paradigm. In the present study, participants saw briefly flashed words. For some blocks, they performed a two-alternative response task where after a mask, two test words were displayed to the left and right of fixation (see

Figure 2). Participants indicated which test word was briefly flashed and their confidence by moving a slider with a computer mouse. We refer to this response task as a two-alternative confidence ratings task (2A-CR). The discrete-state constraints apply across different durations of the briefly flashed word. On some trials, the word was flashed for a relatively long duration, say 50 ms, and performance was relatively high. In other conditions, the word was flashed quickly, say 30 ms, and in one condition, the word was not presented at all (or was flashed for 0 ms). This last condition is useful because the responses conditional on guessing are observable should guessing indeed occur.

The key psychological property in discrete-state models is *conditional independence*. Stimulus factors like flash duration affect whether a trial is mediated by detection or guessing. They do not, however, affect the response distributions conditional on being in a state. Consider for example a trial where a participant has entered a guessing state even though the word was flashed for a modestly long duration. The distribution of responses in this guessing state must be the same as when the participant enters the guessing state for much shorter durations. And the same holds for detection; the distribution of responses is the same conditional on detection regardless of flash duration.

Figure 3 shows the predictions resulting from conditional independence. Figures 3A-B show hypothetical distributions of confidence ratings for the guess and detect states (for when the flashed word is tested on the right), respectively. The response distribution conditional on guessing is centered; the response distribution conditional on detection is shifted toward the correct target. Figures 3C-E show how these distributions are mixed to provide predictions for different flash duration conditions. Panel C is the case where the target is flashed for 0 ms, and responses come from the guessing state. Panel D shows the case for a brief duration, and responses are a 50-50 mixture of guessing and detection. Panel E shows the case for a modestly long duration, and responses are a 90-10 mixture. For each of these panels, the effect of duration is on the mixture probabilities and not on the guessing and detection components themselves.

The predictions of Figures 3C-E may be combined as shown in Figure 3F. The response for the intermediate brief-duration condition is projected downward to avoid clutter. The overall geometrical patterns of the predictions are evidence. The distribution for the brief-duration condition mirrors the composite of the other two in so much as the modes line up. The pattern is much like looking at two mountains reflected in a lake. Figure 3G shows the predictions of a comparable latent-strength model. Here, the effect of duration is simply to shift the distributions toward a higher confidence correct answer. The distributions no longer have a mirror-reflection geometry in the 30 ms condition.

Province and Rouder (2012) used the above logic to assess whether discrete states or latent strength better accounted for recognition memory and found overwhelming support for the discrete states. Of 89 participants, 76 (85%) had patterns better accounted for by a discrete-states model than by a comparable latent-strength model. Here we extend the logic to perceptual word identification. We also include a yes/no response task (see Figure 2) to assess the role of response task.

Method

Participants

Fifty University of Missouri students served as participants for course credit. There were 25 participants each in Experiments 1A and 1B.

Stimuli

A list of 600 nouns, each six to seven letters long, from the MRC Psycholinguistic Database (Coltheart, 1981) served as test items. A second distinct list of 80 nouns was reserved for the calibration phase, which is discussed subsequently.

Design

A session consisted of two tasks: a two-alternative confidence ratings (2A-CR) task and a yes/no confidence-ratings (YN-CR) task. For the 2A-CR task, there were two factors, flash duration and side of target at test, which were crossed in a within-subject manner. For each participant, flash duration was manipulated between four levels. Three of these levels were set for each participant individually in a calibration phase. The remaining level was a 0 ms flash. For the YN-CR task, each trial was either a yes trial, that is the test item was the target, or a no trial, and the test item was a lure. Test item and flash duration were crossed in a within-subject manner.

Procedure

For experimental trials, participants were presented briefly flashed targets that were preceded and followed by pattern masks. In the YN-CR trials, a test stimulus was presented, and participants were asked to rate their confidence whether this test item was indeed the target by moving a mouse on a scale anchored by "Sure Not X" and "Sure X", where X is the test item. In the 2A-CR trials, two items, the target and a lure, were presented at test, and participants rated their confidence about which of the two test items was the target by moving a mouse on a scale between "Sure X_1" and "Sure X_2", where X_1 and X_2 are the tested items.

The session started with the *calibration phase* to find individualized flash durations so that a participant's performance was neither at chance nor at baseline. This phase was a 40-trial adaptive 2-up/1-down staircase (Treutwein, 1995) with two-alternative forced choice response (without confidence ratings). The flash duration for the first trial was set to 100 ms. The adaptive 2-up/1-down procedure then followed. For every two successive correct responses, the flash duration decreased by 10 ms. For every incorrect response, the flash duration increased by 10 ms. The average duration of the last 25 trials of this phase was computed and served as the middle-level flash duration. A longer and a shorter duration served as the remaining two levels, and these were set by adding and subtracting 10 ms from the middle level, respectively. Participants' middle flash settings ranged from 20 ms to 60 ms with a mean of 40.8 ms and a standard deviation of 10.3 ms. The average accuracy across the 0 ms, short, middle, and long flash durations were 0.50, 0.63, 0.76, and 0.88, respectively, for the 2A-CR task and 0.48, 0.60, 0.73, and 0.86, respectively, for the YN-CR task. Standard deviations in accuracy across participants averaged 0.13 across all conditions. The calibration phase, while brief, was successful at achieving performance that was at neither floor nor ceiling across all conditions.

Immediately following the calibration phase was the *2A-CR phase*, which consisted of 150 trials divided into 3 blocks. Of these 150 trials, there were 44 at each flash duration except for 0 ms, for which there were 18 trials. Target items were equally likely to appear to the left or right. The *YN-CR phase* began thereafter, and it was similar to the 2A-CR phase with the exception that it was comprised of yes/no trials rather than two-alternative trials.

Feedback was provided on each trial, but it varied across Experiments 1A and 1B. In Experiment 1A, participants gained points after a correct response and lost points after an incorrect response. The points system followed a polynomial loss rate and a linear gain rate across the confidence scale such that point losses for high-confidence responses were disproportionately larger than the gains (see Province & Rouder, 2012, for details). In Experiment 1B, participants did not receive points, but instead received audio feedback for correct answers regardless of confidence. Because data is analyzed at the individual level, we refer to Experiments 1A and 1B jointly going forward.

Results

Figure 4 shows the histograms of confidence ratings across flash duration conditions for a few selected participants. The patterns in panels A, B, and C display the characteristic mirror-reflection property in Figure 3F, and provide support for discrete-state processing for these participants. The data in panel D, however, do not show this or any clear pattern.

To quantify the support for discrete states and latent strength, we constructed simple discrete-state and latent strength models. The discrete model was a generalized double-threshold model and the latent strength model was an equal-variance signal detection model (details are provided in the supplement). In the 2A-CR task, confidence ratings display a fair degree of symmetry in that responses to targets on the left show a symmetric pattern to responses to targets on the right. Therefore, we restricted detection to be the same whether the target was on the left or right in the discrete-state model, and we restricted the shift to be the same in magnitude whether the target was on the left or right in the latent-strength model. Both the latent strength and discrete-state models had 10 parameters, and both of these models were fit to individual data by maximizing likelihood. Because both models had an equal number of parameters, model comparison was performed through deviance—the model that had the lower deviance was preferred. This deviance approach is equivalent to selection by AIC and BIC because of the equality of the number of parameters in the competing models. In the YN-CR task, the symmetry is not evident in the confidence ratings data, and it was necessary to expand both models. In the discrete-state model, we relaxed the restriction that detection was equal when the target was flashed or was a lure. Instead, we required for each flash duration condition that $d_t = \rho d_l$, where d_t and d_l are the detection of target and lure items, respectively, and where ρ is an additional free parameter. In the latent-strength model, we relaxed the restriction by allowing the shift for lures to be a constant fraction of that for target items, e.g., $\mu_l = -\rho\mu_t$, where μ_t and μ_l are the shifts relative to midpoint for target and lure items, respectively, and ρ is a free parameter. Consequently, the discrete-state and latent strength models had 11 parameters each, and model comparison could still be done by comparison of deviance.

Figure 5 shows the deviance results for both tasks. For 45 of the 50 participants (90%), the discrete-state model was preferred for the two-alternative confidence ratings task. For 35 of the 50 participants (70%), the discrete-state model was preferred in the yes-no confidence-ratings task. This evidence may be further quantified by computing Bayes factors for these results. The Bayes factor is the probability of the data under one model relative to that under an alternative model (Jeffreys, 1961). We analyzed the proportion of participants who had lower deviance for the discrete-state than the latent-strength model. If the discrete-state model is a better description overall, then the underlying true probability should be high while if the latent-strength model is a better description overall, then the underlying true probability should be low. The Bayes factor values were 860 million and 430, respectively, for the 2A-CR and YN-CR tasks, both in favor of the discrete-state model.[1] The value for the 2A-CR task indicates decisive support, and that for the YN-CR task, while much less, is still strong. Readers who insist on frequentist interpretations may perform sign tests on these frequencies.

Discussion

In this report we provide a more general and less assumptive discrete-state model as well as a crucial test of its core property of conditional independence. Conditional independence—that responses are only determined by mental states—may be assessed in a straightforward manner with confidence ratings. We find support for a discrete-state model relative to a comparable latent strength competitor, and this result bolsters the possibility that impoverished words are mediated by stable detect and guessing states.

Our findings are similar to our previous work with recognition memory, indicating that this mediation may be quite general in the processing of words.

We find appealing the subtle Fechnarian view of discrete states where sub-threshold information may be represented, that is, it is not lost entirely, but is unavailable consciously and does not influence direct behavioral report. Stimulus strength factors such as flash duration may still affect perception in a graded manner, but they affect the probability of detection rather than the response on any trial (see Rouder and Morey, 2009). We note that these results stand in contrast to more accepted graceful degradation views of information loss in the literature on perception and memory. Traditionally, researchers have been reluctant to consider discrete-state models, but perhaps these results will entice researchers to reconsider this reluctance. This research raises future questions of whether behavioral reports in many or most domains are mediated by discrete states, and how stimulus factors are represented and affect the probability of entering one state or another.

References

Ashby, G. F. (1992). *Multidimensional models of perception and cognition.* Hillsdale, NJ: Erlbaum.

Blackwell, H. R. (1953). Psychological thresholds: Experimental studies of methods of measurement. *Bulletin of the Engineering Research Institute of the University of Michigan, No. 36.*

Broder, A., & Schutz, J. (2009). Recognition rocs are curvilinear - or are they? On premature arguments against the two-high-threshold model of recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 587-606.

Carterette, E., & Cole, M. (1962). Comparison of the receiver-operating characteristics for messages received by ear and by eye. *Journal of the Acoustical Society of America*, *34*, 172-178.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33A*, 497-505.

Dube, C., & Rotello, C. M. (2012). Binary rocs in perception and recognition memory are curved. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 130-151.

Egan, J. P. (1958). *Recognition memory and the operating characteristic* (Tech. Rep. No. AFCRC-TN-58-51). Bloomington, Indiana: Indiana University Hearing and Communication Laboratory.

Erdfelder, E., & Buchner, A. (1998). Comment: Process-dissociation measurement models: Threshold theory or detection theory. *Journal of Experimental Psychology: General*, *127*(1), 83–97.

Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 500-513.

Goodenough, D., Rossman, K., & Lusted, L. (1972). Radiographic applications of signal detection theory. *Radiology*, *105*(1), 199-200.

Haase, S., & Fisk, G. (2001). Confidence in word detection predicts word identification: Implications for an unconscious perception paradigm. *American Journal of Psychology*, *114*(3), 439-468.

Jacobs, A., Graf, R., & Kinder, A. (2003). Receiver operating characteristics in the lexical decision task: Evidence for a simple signal-detection process simulated by the multiple read-out model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 481-488.

Jeffreys, H. (1961). *Theory of probability (3rd edition).* New York: Oxford University Press.

Kellen, D., Klauer, K., & Broder, A. (2013). Recognition memory models and binary-response rocs: A comparison by minimum description length. *Psychonomic Bulletin & Review*, 1-27.

Kinchla, R. A. (1994). Comments on Batchelder and Riefer's multinomial model for source monitoring. *Psychological Review*, *101*, 166-171.

Klauer, K., & Kellen, D. (2011). The flexibility of models of recognition memory: An analysis by the minimum-description length principle. *Journal of Mathematical Psychology*, *55*(6), 430-450.

Ma, W., & Huang, W. (2009). No capacity limit in attentional tracking: Evidence for probabilistic inference under a resource constraint. *Journal of Vision*, *9*(11), doi:10.1167/9.11.3.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, N.J.: Lawrence Erlbaum Associates.

Malmberg, K. J. (2002). On the form of rocs constructed from confidence ratings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 380-387.

Pratte, M. S., & Rouder, J. N. (2011). Hierarchical single- and dual-process models of recognition memory. *Journal of Mathematical Psychology*, *55*, 36–46.

Province, J. M., & Rouder, J. N. (2012). Evidence for discrete-state processing in recognition memory. *Proceedings of the National Academy of Sciences*.

Qin, J., Raye, C. L., Johnson, M. K., & Mitchell, K. J. (2001). Source ROCs are (typically) curvilinear: Comment on Yonelinas (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1110-1115.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for decisions between two choices. *Psychological Science*, *9*, 347-356.

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

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin and Review*, *12*, 573-604.

Rouder, J. N., & Morey, R. D. (2009). The nature of psychological thresholds. *Psychological Review*, *116*, 655-660.

Rouder, J. N., Province, J. M., Swagman, A. R., & Thiele, J. E. (submitted). From ROC curves to psychological theory.

Shams, L., Ma, W., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *NeuroReport: For Rapid Communication of Neuroscience Research*, *16*(17), 1923-1927.

Sims, C., Jacobs, R., & Knill, D. (2012). An ideal observer analysis of visual working memory. *Psychological Review*, *119*(4), 807-830.

Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. *Memory & Cognition*, *33*, 1499-1517.

Swets, J. A. (1961). Is there a sensory threshold? *Science*, *134*, 168-177.

Tanner, J., W. P, & Birdsall, T. G. (1958). Definition of d' and η as psychophysical measures. *Journal of the Acoustical Society of America*, *30*, 922-928.

Thierman, T. (1968). A signal detection approach to the study of set in tachistoscopic recognition. *Perceptual and Motor Skills*, *27*(1), 96-98.

Townsend, J. T. (1971). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, *9*, 40-50.

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, *35*, 2503-2522.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin*, *133*, 800-832.

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science*, *20*, 423-428.

Footnotes

[1] The Bayes factor modeling went as follows. The number of participants who had lower deviance scores for the discrete-state model, denoted y, was modeled as $y \sim \text{Binomial}(p, n)$, where p is the true probability and n is the number of participants. The discrete-state-favored model was instantiated by the prior $p \sim \text{Uniform}(.5, 1)$; the latent-strength favored-model was instantiated by the prior $p \sim \text{Uniform}(0, .5)$. The Bayes factor is

$$B = \frac{\int_{.5}^{1} p^y (1-p)^{n-y} dp}{\int_{0}^{.5} p^y (1-p)^{n-y} dp}.$$

The integrals are evaluated numerically.

Figure Captions

Figure 1. ROC curves, discrete-state models, and ROC predictions. **A.** An example from Slotnick and Dotson (2005). The points are Slotnick and Dotson's observed hit and false alarm rates, the dashed straight line is the best fit of the high-threshold model with certainty (see Panel C), and the solid curve is the best fit of the high-threshold model without certainty (see Panel D). **B.** An example from Qin, Raye, Johnson, & Mitchell (2001). Points are the observed data and the lines are the best model fits. **C.** A double-high threshold model with the certainty assumption that predicts straight-line ROCs. **D.** A double-high threshold model without the certainty assumption that yields piecewise linear ROCs that mimic curvature. *HC*: high-confidence. *LC*: Low-confidence. Parameters d_s and d_n denote the probability of detection for studied and new items, respectively. Parameters $\gamma_1, \ldots, \gamma_K$ denote the probability of endorsing a confidence-ratings option conditional on guessing; parameters $\alpha_1, \ldots, \alpha_K$ and $\beta_1, \ldots, \beta_K$ denote the same conditional on respective detect states.

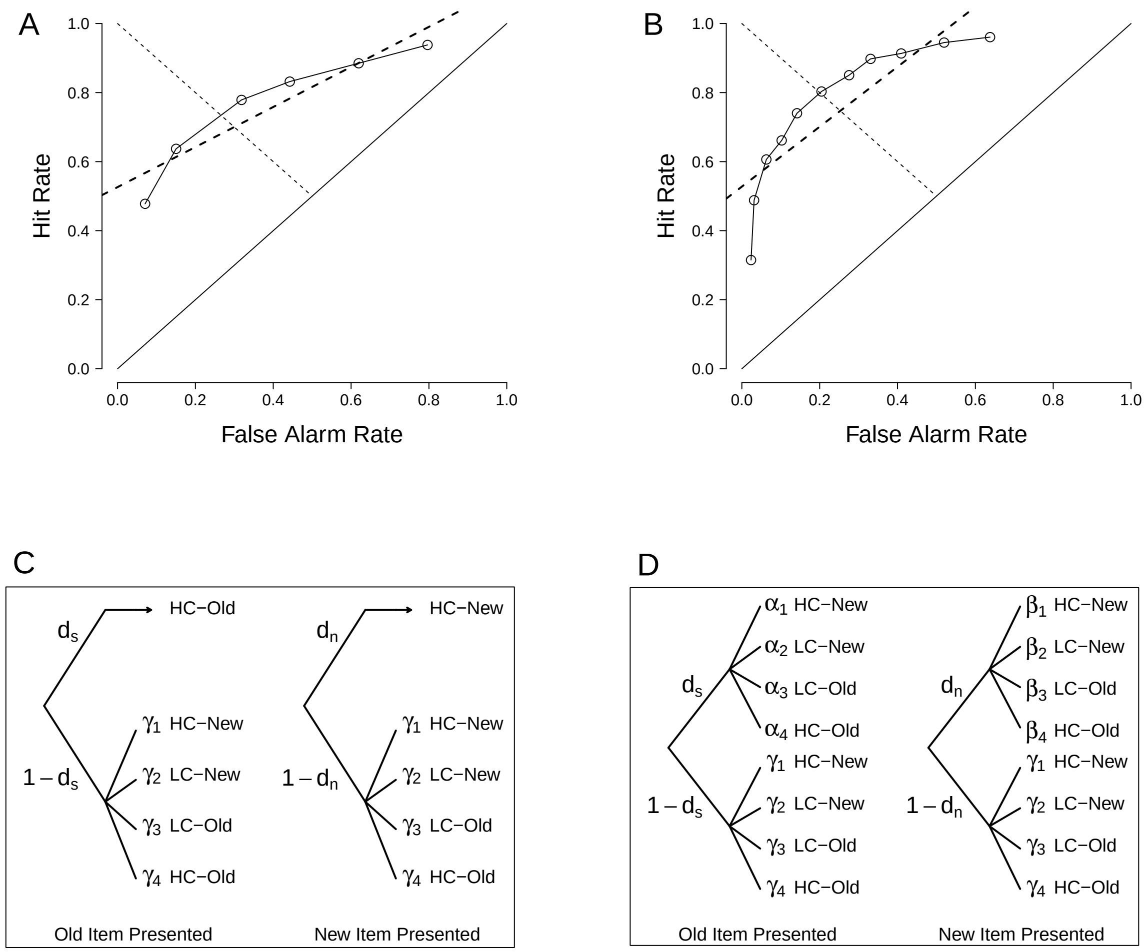
Figure 2. Word identification task and both test types. The study sequence is shown on the left, and at the right are the two-alternative confidence-ratings and yes-no confidence-ratings test displays. The test item in the yes-no confidence ratings task could be either the target word (STATUE), as pictured, or some lure (e.g. MIRROR).

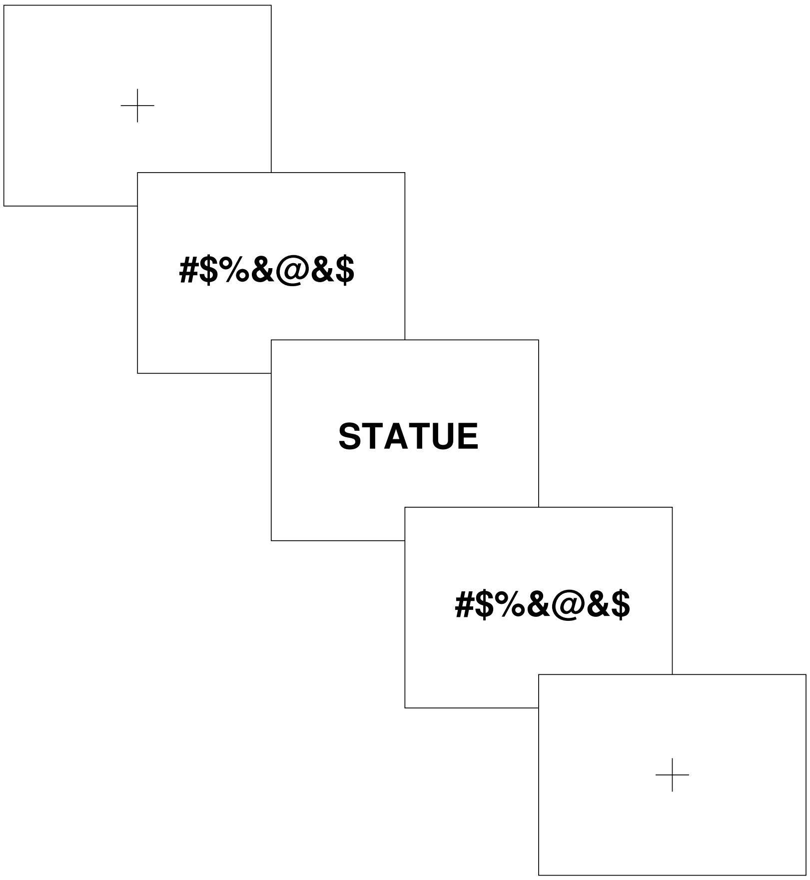
Figure 3. Predictions for distributions of confidence ratings from a discrete-state model where STATUE was the flashed word and MIRROR was the lure. **A.** Distribution of confidence ratings from a hypothetical guess state. **B.** Distribution of confidence ratings from a hypothetical detect state. **C.** Distribution of responses from the 0 ms condition which reflects only the guess state. **D.** Distribution of responses from a 30 ms condition which is an equal mixture of responses from the guess and detect states. **E.** Responses from a 50 ms condition, which are a 90-10 mixture with a predominance of

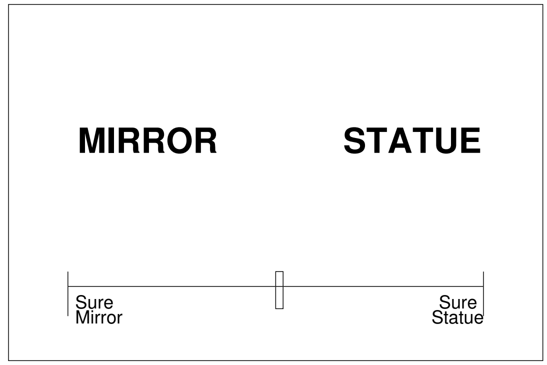
detection-mediated responses. **F.** Predictions from a discrete-state model for all three conditions. Note the brief-duration (30 ms) condition is projected downward. **G.** Predictions of a comparable latent-strength model.

Figure 4. Histograms of selected participants' confidence ratings for the 2A-CR and YN-CR tasks. Deviance scores from these four participants are highlighted in Figure 5. **A-C.** Patterns showing evidence of discrete-state processing. **D.** A data set better fit by the latent-strength model.

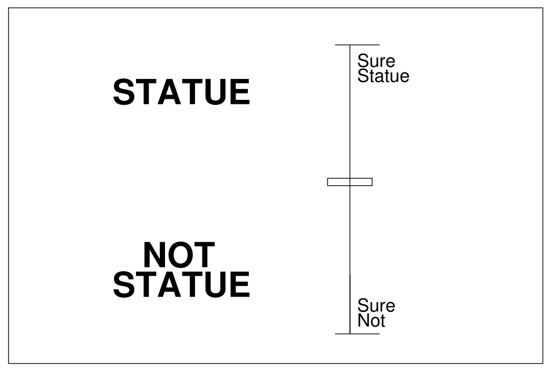
Figure 5. Deviance differences (discrete-state deviance subtracted from latent strength deviance) for each participant for 2A-CR (plotted as circles) and YN-CR (plotted as triangles) tasks. Positive values show a better fit for the discrete-state model than the latent-strength model. Participants whose confidence ratings were displayed in Figure 4 are labeled A-D.


A
B
Hit Rate
False Alarm Rate
0.0
0.2
0.4
0.6
0.8
1.0
C
D
d_s
d_n
$1 - d_s$
$1 - d_n$
HC−Old
HC−New
γ_1 HC−New
γ_2 LC−New
γ_3 LC−Old
γ_4 HC−Old
α_1 HC−New
α_2 LC−New
α_3 LC−Old
α_4 HC−Old
β_1 HC−New
β_2 LC−New
β_3 LC−Old
β_4 HC−Old
Old Item Presented
New Item Presented

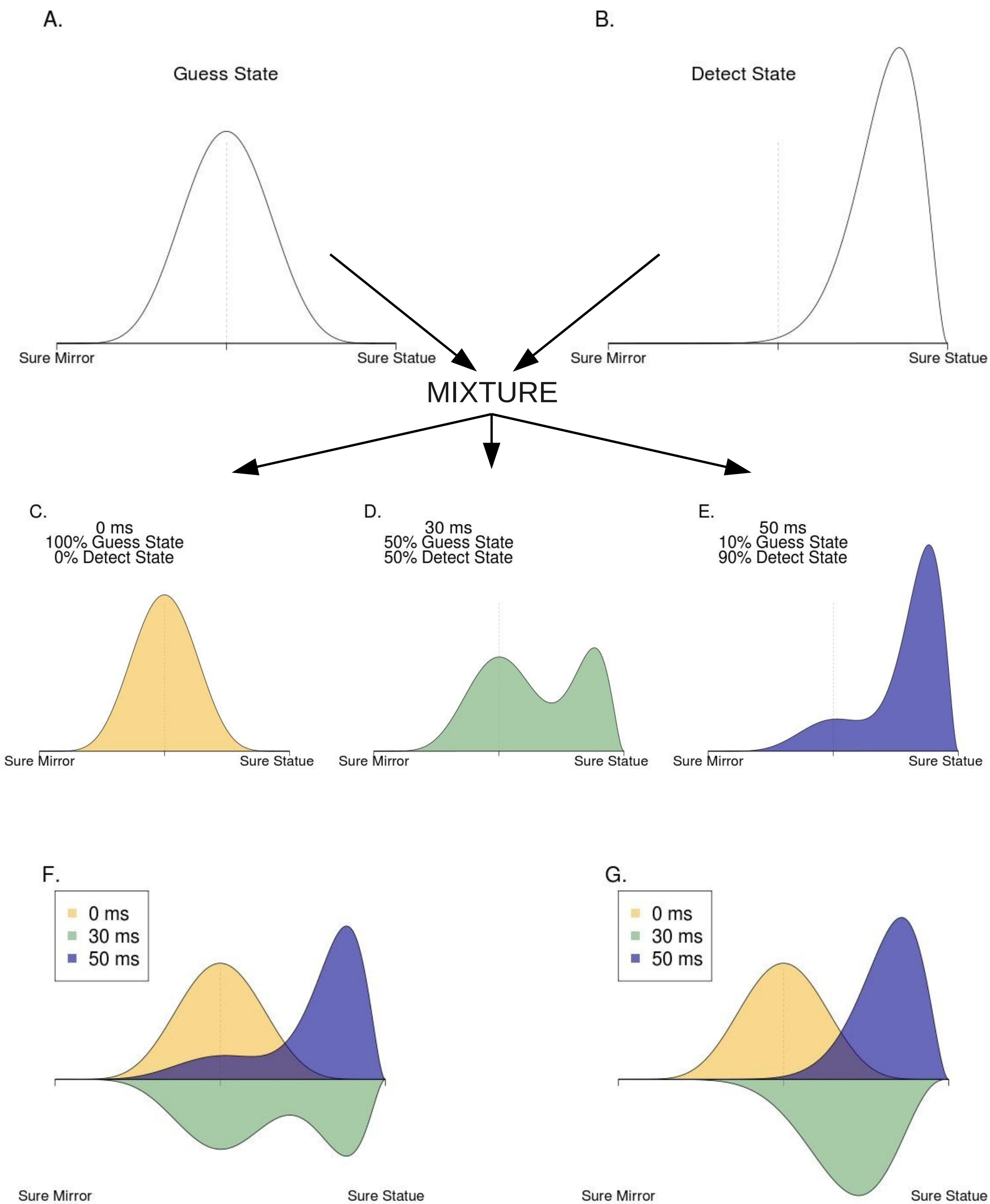


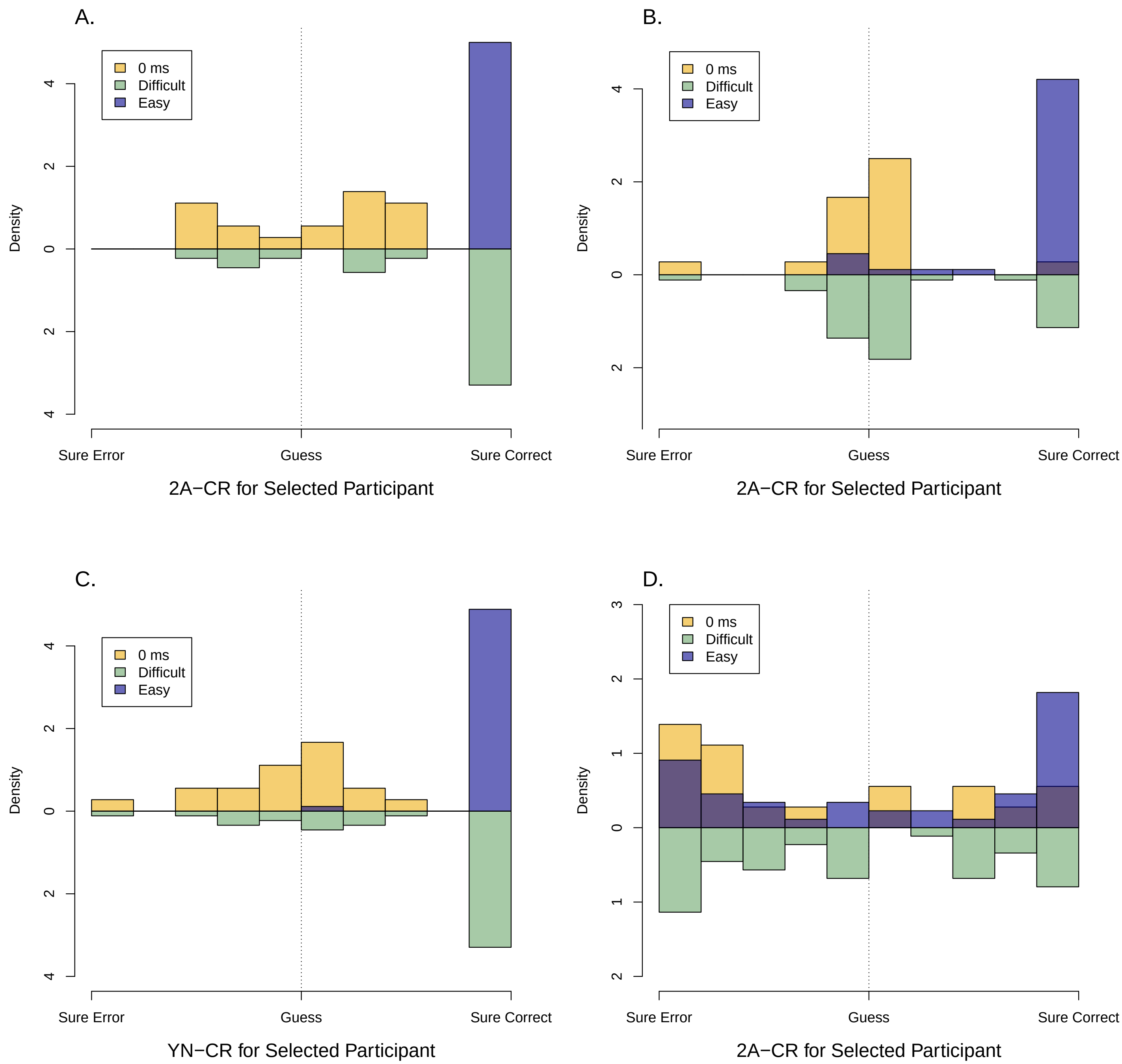


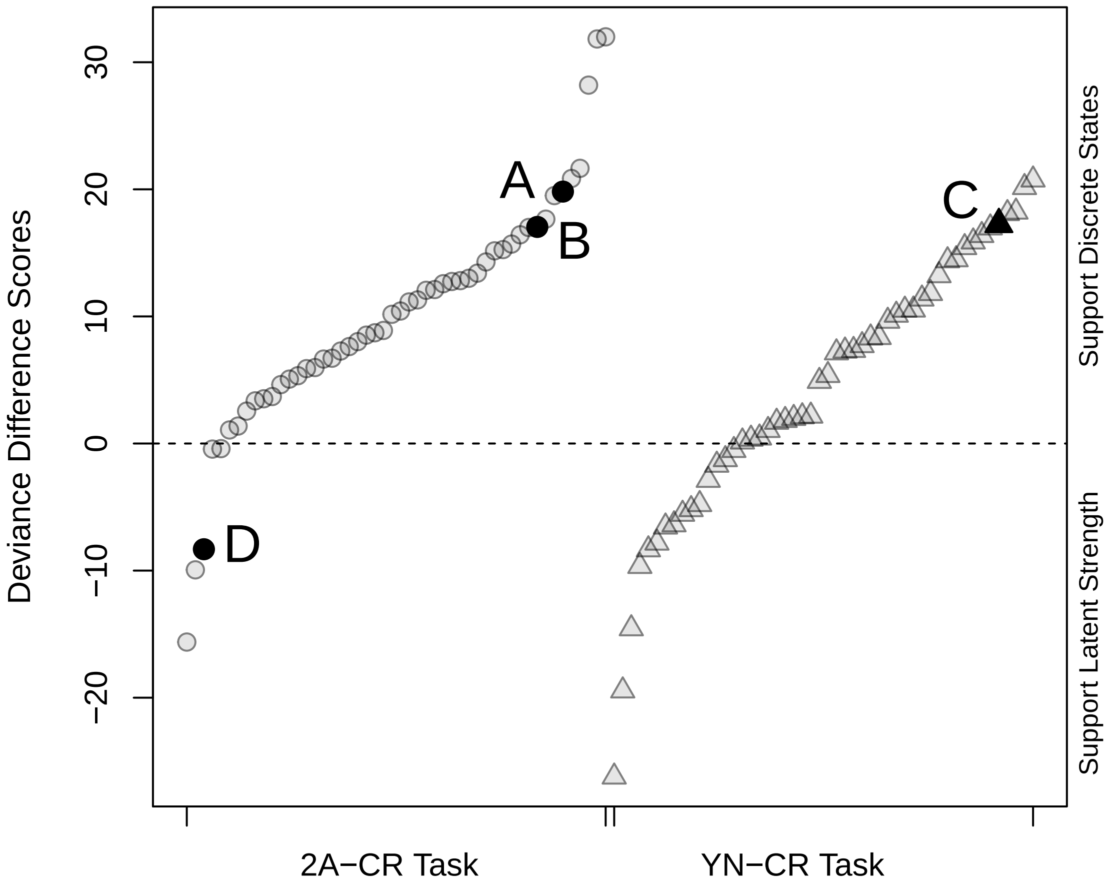
2 Alternative – Confidence Ratings



Yes/No – Confidence Ratings


A.
Guess State
Sure Mirror
Sure Statue
B.
Detect State
Sure Mirror
Sure Statue
MIXTURE
C.
0 ms
100% Guess State
0% Detect State
Sure Mirror
Sure Statue
D.
30 ms
50% Guess State
50% Detect State
Sure Mirror
Sure Statue
E.
50 ms
10% Guess State
90% Detect State
Sure Mirror
Sure Statue
F.
0 ms
30 ms
50 ms
Sure Mirror
Sure Statue
G.
0 ms
30 ms
50 ms
Sure Mirror
Sure Statue


A.
0 ms
Difficult
Easy
Density
Sure Error
Guess
Sure Correct
2A−CR for Selected Participant
B.
0 ms
Difficult
Easy
Density
Sure Error
Guess
Sure Correct
2A−CR for Selected Participant
C.
0 ms
Difficult
Easy
Density
Sure Error
Guess
Sure Correct
YN−CR for Selected Participant
D.
0 ms
Difficult
Easy
Density
Sure Error
Guess
Sure Correct
2A−CR for Selected Participant



Deviance Difference Scores
30
20
10
0
−10
−20
A
B
C
D
Support Discrete States
Support Latent Strength
2A−CR Task
YN−CR Task