**A Concurrent Numerical Stroop Task Increases the Vulnerability of Procedural Skills to Unlearning**

Matthew J. Crossley1, F. Gregory Ashby2 & W. Todd Maddox3

1 Department of Psychology, University of California, Berkeley

2 Department of Psychological & Brain Sciences, University of California, Santa Barbara

3 Department of Psychology, University of Texas, Austin

August 19, 2013

Address: Matthew J. Crossley

Department of Psychology

University of California

Berkeley, CA 94720

Email: [mjcrossley@berkeley.edu](mailto:ashby@psych.ucsb.edu)

Phone: (510) 642-0135

Fax: (510) 642-0135

Running Head: Cognitive Load Potentiates Unlearning

**Abstract**

We recently reported the efficacy of various intervention protocols in unlearning procedural categorization (Crossley, Ashby & Maddox, 2012, J. of Experimental Psychology: General). This work demonstrated that an intervention protocol composed of random feedback effectively reduced accuracy during the intervention, but lead to massive savings in relearning (i.e., failed to induce true unlearning). However, an intervention protocol composed of a mixture of random and veridical feedback eliminated savings in relearning (i.e., caused true unlearning). This article shows that performing a concurrent numerical Stroop task during an intervention of random feedback reduces or else fully eliminates savings in relearning. This suggests that the mechanism that protects learning during random feedback relies on prefrontal mechanisms. **Introduction**

1. Huge amounts of money are spent on programs designed to facilitate “unlearning” of maladaptive behaviors (examples)
2. In most cases these are failures
3. In recent paper we proposed a neurobiological theory of why a large class of behaviors (PB) are so resistant to unlearning.
4. Brief outline of the theory—striatal, tans, etc
5. We showed that the model could account of the classic relearning phenomena under random feedback conditions.
6. In addition, we used model to proposed and tested an effective unlearning protocol.
7. 75/25 found to work and predicted by model
8. Goal of current study is to determine whether the mechanism that protects learning during random feedback relies on prefrontal mechanisms.
9. To test this we examine whether a concurrent numerical Stroop task during an intervention of random feedback reduces or fully eliminates savings in relearning.

PBL and II

We examined savings in relearning in the information-integration (II) category-learning task. II categorization tasks are those in which the stimuli are assigned to categories in such a way that accuracy is maximized only if information from two or more non-commensurable stimulus dimensions is integrated at some pre-decisional stage (Ashby & Gott, 1988). Typically, the optimal strategy in II tasks is difficult or impossible to describe verbally (which makes it difficult to discover via logical reasoning). An example of an II task is shown in Figure 1. In this case the four categories are each composed of single black lines that vary in length and orientation. The diagonal lines denote the category boundaries. Note that no simple verbal rule correctly separates the lines into the four categories. Nevertheless, many studies have shown that with enough practice, people reliably learn such categories, and the evidence is good that II category learning uses procedural memory and requires dopamine-dependent reinforcement learning in the striatum (e.g., Ashby & Maddox, 2005).

The II task used here included acquisition, intervention, and reacquisition phases of 300 trials each. These three phases were identical except in the nature of the feedback provided after each response. During acquisition and reacquisition, feedback indicated whether each response was correct or incorrect. During the intervention phase, the feedback was random – that is, participants were informed that their response was correct with probability ¼ and incorrect with probability ¾, regardless of what response they actually made. Every stimulus in all three phases of Experiment 1 was a line (as in Figure 1) that varied across trials in length and orientation. Identical II category structures were used in all three phases. These are represented abstractly in Figure 1. Also note that the categories overlap slightly such that the best possible accuracy with these categories is 95%. In these capacities, this task is identical to the task used in Experiment 1 of Crossley et al. (2012).

Details of Concurrent Dual Task

The present experiment diverges from Crossley et al. (2012) in that …insert description of dual task.

INSERT FIGURE 1 ABOUT HERE

**Methods**

*Participants*

There were 30 participants in Condition 1 and 30 participants in Condition 2, 28 participants in Condition 3, and 27 participants in Condition 4. All participants completed the study and received course credit for their participation. All participants had normal or corrected to normal vision. To ensure that our analyses included only participants who performed well above chance on the category learning task - while sufficiently attending to the concurrent numerical Stroop – two exclusion criteria were applied. First, a learning criterion of 40% correct (25% is chance) during acquisition trials 200-250 was applied. We additionally excluded any subject whose mean performance on the Stroop task was less than 85% correct. Using these criteria, we excluded 8 participants from Condition 1, 11 participants from Condition 2, 7 participants from Condition 3, and 9 participants from Condition 4.

*Stimuli and Procedure*

All stimuli and procedures were identical to those used in Crossley et al. (2012). Example stimuli, as well as the complete category distributions used are shown in Figure 1. The experimental design for each condition is shown in Figure 2.

Extra notes on the dual task:

1. During stimulus 1 presentation in the dual task condition, the stroop stimuli are displayed concurrently to the left and right of the categorization stimulus for 200 ms and are then replaced by a rectangular white mask (one on each side) for 200ms.
2. On 85% of trials the numerically larger number is physically smaller.
3. During the value/size query, on half of the trials the word “value” comes up and on half of the trials the word “size” comes up. This queries the participant to determine which side (left or right) had the numerically or physically larger stimulus.
4. The participant is instructed to perform the numerical task without error and to use whatever is left over to do the categorization task.
5. Participants’ current accuracy on the numerical Stroop task should be indicated at the top of the screen when they received feedback regarding their performance on the concurrent task on each trial. Their percentage correct score should be listed in green if it was above 80% and red if it was below 80%.

INSERT FIGURE 2 ABOUT HERE

**Results**

*Accuracy-based results*

The top panel of Figure 3 shows the mean accuracy for every 25-trial block of each condition. During intervention, a response was coded as correct if it agreed with the category membership shown in Figure 1. Recall that the categories and feedback were identical in all conditions throughout the entire experiment… except for the dual-task onsets and offsets. Note that participants from all conditions were able to learn the categories, and that accuracy dropped significantly during intervention, but remained well above chance for every condition. No participants in any condition showed savings in relearning, and participants in Condition 2 and Condition 3 showed an interference.

INSERT FIGURE 3 ABOUT HERE

To test these conclusions formally we … insert stats here

We then performed several post-hoc pairwise contrast tests and corrected for multiple comparisons using the Tukey test. Note that all Tukey post-hoc tests reported in this article use an experiment-wise type I error rate set to α = .05. Insert more stats here.

Next, we performed several 1-way repeated measures ANOVAs to compare performance between phases within each condition. Insert more stats here.

Finally, we computed difference scores between acquisition and reacquisition for every subject in each condition and performed a mixed-design, repeated measures ANOVA to compare the amount of savings present in each condition. These difference scores are shown in Figure 4.

INSERT FIGURE 4 ABOUT HERE

*Decision Bound Modeling*

Optimal performance on the category structures used here can only be obtained via a procedural strategy. However, explicit strategies can nevertheless yield better-than-chance performance (e.g., the accuracies observed here). Since our goal is to examine savings in procedural learning, we must carefully rule out the contribution to savings from explicit strategies. To this end, we partitioned the data from each participant into blocks of 100 trials and fit different types of decision bound models (e.g., Maddox & Ashby, 1993; Ashby, Waldron, Lee, & Berkman, 2001) to each block of data from every participant. One type assumed a rule-based decision strategy, one type assumed an II (i.e., procedural) strategy, and one type assumed random guessing. See Appendix 1 of Crossley et al. (2012) for more details.

Table 2 shows the number of participants in the four conditions best fit by a model of these three types. Create this table and discuss it.

INSERT TABLE 2 ABOUT HERE

**Discussion**

1. Previously developed model of performance in 3-phase pbl task
2. Key is TANS
3. Using model identified condition (75/25) that induced true unlearning
4. Idea is that TANs are “tricked” into thinking that this is still a rewarding context
5. Focus of current report is to determine whether this mechanism is prefrontally mediated.
6. Test with concurrent WM task
7. Find support for prefrontal mediation
8. Make point that this suggests an alternative to continuing to give valid rewards during intervention.

**References – will need updating**

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105*, 442-481.

Ashby, F. G., & Crossley, M. J. (2011). A computational model of how cholinergic interneurons protect striatal-dependent learning. Journal of Cognitive Neuroscience, *23*, 1549-1566.

Ashby, F., Ell, S., & Waldron, E. (2003). Procedural learning in perceptual categorization. *Memory & Cognition*, *31*, 1114–1125.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory and cognition, 14*, 33-53.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. Annual Review of Psychology, 56, 149-178.

Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin & Review*, 6, 363-378.

Ashby, F. G., Waldron, E. M., Lee, W. W., & Berkman, A. (2001). Suboptimality in human categorization and identification. *Journal of Experimental Psychology:* *General*, 130, 77-96.

Bienenstock, E.L., Cooper, L.N., & Munro, P.W. (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, 2, 32–48.

Bouton, M. E., Todd, T. P., Vurbic, D., & Winterbauer, N. E. (2011). Renewal after the extinction of free operant behavior*. Learning & behavior*, 39(1), 57-67.

Crossley, M. J., Ashby, F. G., & Maddox, W. T. (2012). Erasing the engram: The unlearning of procedural skills. *Journal of Experimental Psychology: General.*

Ermentrout, B. (1996). Type i membranes, phase resetting curves, and synchrony. *Neural Computation*, *8*, 979–1001.

Gershman, S., Blei, D., & Niv, Y. (2010). Context, learning, and extinction. *Psychological Review*, *117*, 197-209.

Higgins, S. T., Budney, A. J., & Bickel, W. K. (1995). Outpatient behavioral treatment for cocaine dependence: One-year outcome. *Experimental and Clinical Psychopharmacology, 3*, 205-212.

Izhikevich, E. (2007). *Dynamical systems in neuroscience: The geometry of excitability and bursting*. Cambridge, MA: The MIT press.

Kirkwood, A., Rioult, M.G. & Bear, M.F. (1996). Experience-dependent modification of synaptic plasticity in visual cortex. *Nature*, 381, 526-528.

Kruschke, J. K. (2011). Models of attentional learning. In E. M. Pothos and A. J. Wills (eds.), *Formal Approaches in Categorization*, pp. 120-152. Cambridge University Press.

Lewandowsky, S., & Kirsner, K. (2000). Expert knowledge is not always integrated: A case of cognitive partition. *Memory & Cognition, 28,* 295–305.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, 53, 49-70.

Maddox, W., Ashby, F., Ing, A., & Pickering, A. (2004). Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Memory & Cognition*, *32*, 582–591.

Maddox, W., Bohil, C., & Ing, A. (2004). Evidence for a procedural-learning based system in perceptual category learning. *Psychonomic Bulletin & Review*, *11*, 945–952.

Matsumoto, N., Minamimoto, T., Graybiel, A., & Kimura, M. (2001). Neurons in the thalamic cm-pf complex supply striatal neurons with information about behaviorally signiﬁcant sensory events. *Journal of Neurophysiology*, *85*, 960-976.

Nakajima, S., Tanaka, S., Urshihara, K., & Imada, H. (2000). Renewal of extinguished lever-press responses upon return to the training context. *Learning & Motivation, 31*, 416–431.

Nakajima, S., Urushihara, K., & Masaki, T. (2002). Renewal of operant performance formerly eliminated by omission or noncontingency training upon return to the acquisition context. *Learning and Motivation*, *33*, 510–525.

Rall, W. (1967). Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic input*. Journal of Neurophysiology, 30*(5), 1138-1168.

Redish, A. D., Jensen, S., Johnson, A., & Kurth-Nelson, Z. (2007). Reconciling reinforcement learning models with behavioral extinction and renewal: Implications for addition, relapse, and problem gambling. *Psychological Review*, *114*, 784-805.

Sanborn, A.N., Griffiths, T.L., & Navarro, D.J. (2010). Rational approximations to rational models: Alternative algorithms for category learning. Psychological Review, 117, 1144-1167.

Schultz, W., Dayan, P., & Montague, P. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593-1599.

Tobler, P., Dickinson, A., & Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *The Journal of Neuroscience*, *23*, 10402-10410.

Waldschmidt, J., & Ashby, F. G. (2011). Cortical and striatal contributions to automaticity in information-integration categorization. *Neuroimage*, 56, 1791-1802.

Yang, L.-X., & Lewandowsky, S. (2004). Knowledge partitioning in categorization: Constraints on exemplar models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 1045-1064.

**Author Notes**

We especially thank Micajah Spoden for help with programming and members of the Maddox Lab for all data collection. This research was supported in part by the U.S. Army Research Office through the Institute for Collaborative Biotechnologies under grant W911NF-09-D-0001, by grants P01NS044393 from the National Institute of Neurological Disorders and Stroke and DA032457 from the National Institute of Drug Abuse, and by grant FA9550-12-1-0355 from AFOSR.

**Figure Legends**

**Insert legends here**

**Table 1**. Within-condition comparison between phases via repeated measures mixed design 2 factor (Condition x Block) ANOVAs (1 per phase).

|  |  |
| --- | --- |
| Acquisition | |
| Condition | F(3,81) = 1.5, P = 0.2 |
| Block | F(11,826) = 38.8, P < 0.001 |
| Interaction | F(33,826) = 0.9, P = 0.6 |
| Intervention | |
| Condition | F(3,75) = 6.49, P < 0.001 |
| Block | F(15,1117) = 15.47, P < 0.001 |
| Interaction | F(45,1117) = 2.02, P < 0.001 |
| Intervention Posthoc Tukey HSD Pairwise Comparison Between Conditions | |
| 1 - 2 | P = 0.47 |
| 1 - 3 | P < 0.005 |
| 1 - 4 | P < 0.05 |
| 2 - 3 | P < 0.05 |
| 2 - 4 | P = 0.44 |
| 3 - 4 | P = 0.88 |
| Reacquisition | |
| Condition | F(3,73) = 2.74, P < 0.05 |
| Block | F(5,358) = 9.91, P < 0.001 |
| Interaction | F(15,358) = 2.02, P < 0.05 |
| Reacquisition Posthoc Tukey HSD Pairwise Comparison Between Conditions | |
| 1 - 2 | P = 0.99 |
| 1 - 3 | P = 0.26 |
| 1 - 4 | P = 0.45 |
| 2 - 3 | P = 0.07 |
| 2 - 4 | P = 0.31 |
| 3 - 4 | P = 1.0 |

Shaded cells indicate significant differences. Note that we report effect size for all repeated measures t-tests as described by Gibbons et al. (1993).

**Table 2**. Within-condition comparison between phases via repeated measures mixed design 2 factor (Condition x Block) ANOVAs (1 per phase).

|  |  |
| --- | --- |
| Savings: Reacquisition - Acquisition | |
| Condition | F(3,37) = 0.64, P = 0.6 |
| Block | F(11,395) = 8.74, P < 0.001 |
| Interaction | F(33,395) = 1.34, P = 0.1 |
| Intervention Posthoc Tukey HSD Pairwise Comparison Between Conditions | |
| 1 - 2 | P = 0.75 |
| 1 - 3 | P = 0.79 |
| 1 - 4 | P = 1.0 |
| 2 - 3 | P = 1.0 |
| 2 - 4 | P = 0.73 |
| 3 - 4 | P = 0.78 |

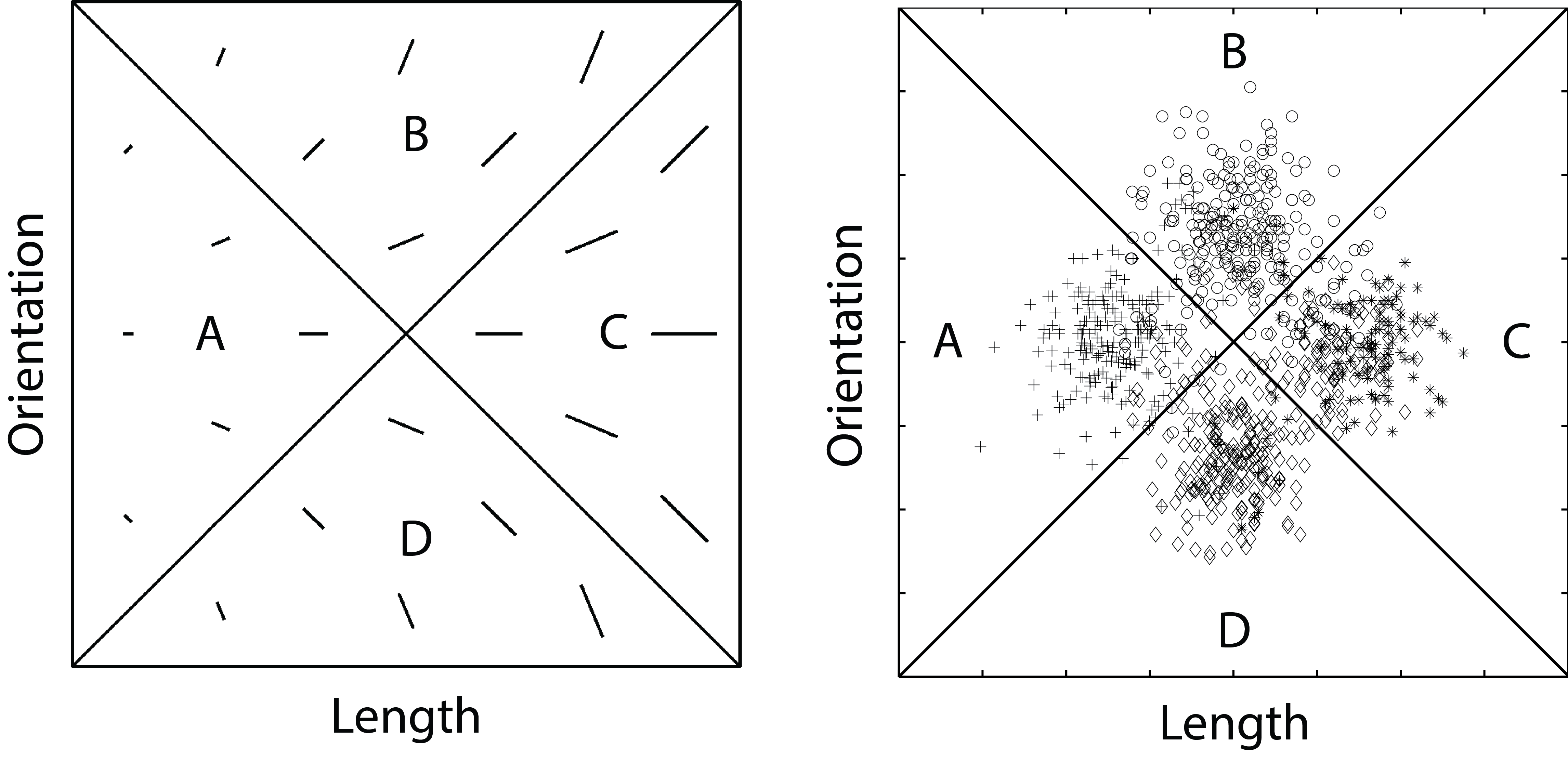
The take-home here is that – while Figure 4 shows that Conditions 1 and 4 have a significant impairment (i.e., negative savings) and conditions 2 and 3 aren’t impaired but also don’t show savings – savings (or absence thereof) isn’t significantly different between groups. In a sense this isn’t ideal because we would love for savings to be rank ordered with dual-task exposure duration. However, they key is that we don’t get savings… which is a big deal because Crossley et al. (2012) totally did… so the dual-task is doing something.

**Table 3**. Number of participants in the four conditions of the Experiment whose responses were best accounted for by a model that assumed an information-integration (II) decision strategy, a rule-based strategy, or random guessing, and the percentage of responses accounted for by those models.

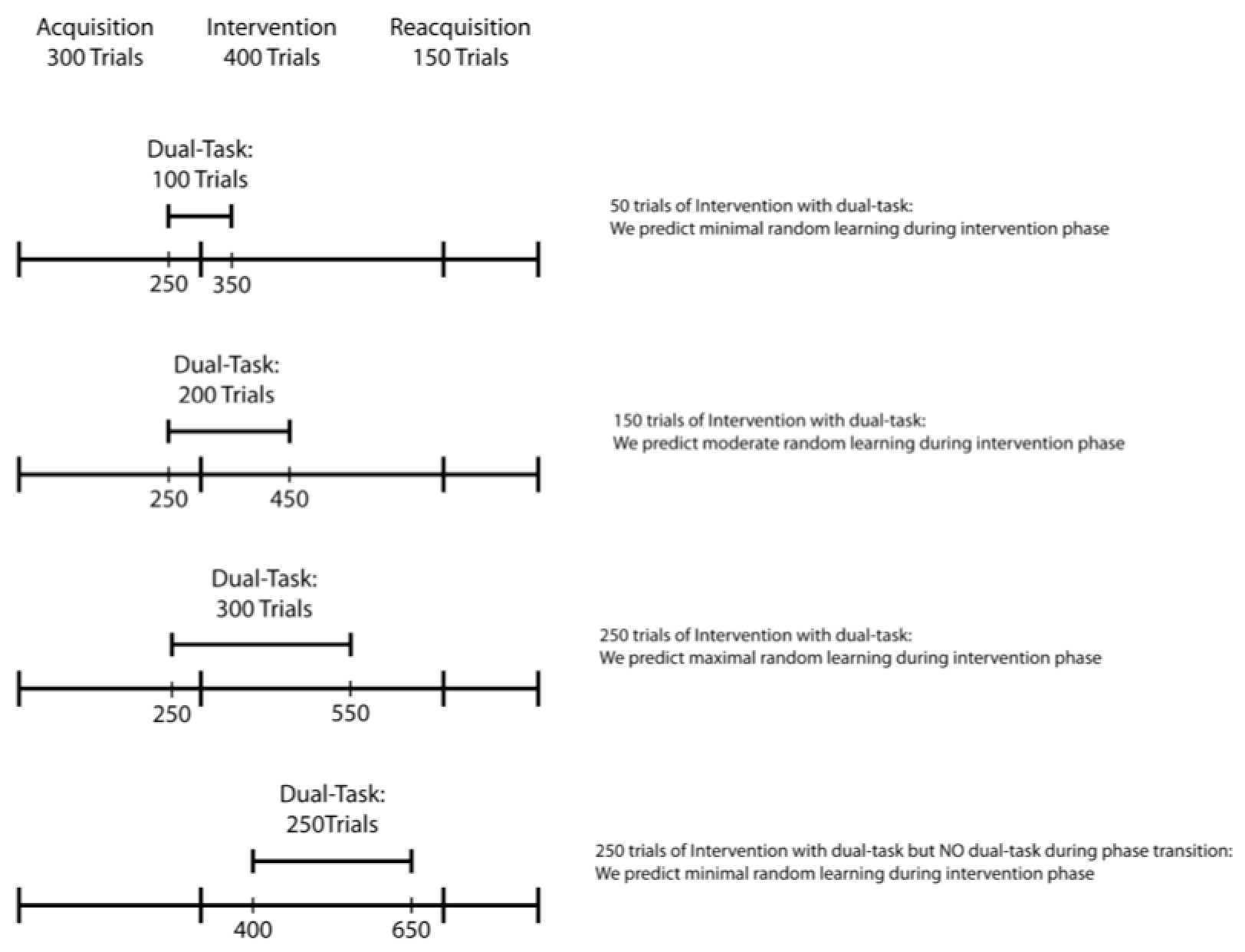
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Condition 1** | | | | **Condition 2** | | | |
|  | A3 | | R1 | | A3 | | R1 | |
|  | N | % | N | % | N | % | N | % |
| **II** | 14 | 68.0 | 11 | 53.4 | 13 | 70.3 | 8 | 65.1 |
| **RB** | 8 | 62.0 | 5 | 46.4 | 6 | 75.3 | 6 | 61.3 |
| **Guessing** | 0 | NA | 6 | NA | 0 | NA | 5 | NA |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Condition 3** | | | | **Condition 4** | | | |
|  | **A3** | | **R1** | | **A3** | | **R1** | |
|  | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** |
| **II** | 15 | 68.0 | 10 | 58.2 | 13 | 71.2 | 7 | 60.3 |
| **RB** | 6 | 65.8 | 6 | 56.3 | 5 | 62.0 | 10 | 57.2 |
| **Guessing** | 0 | NA | 5 | NA | 0 | NA | 1 | NA |

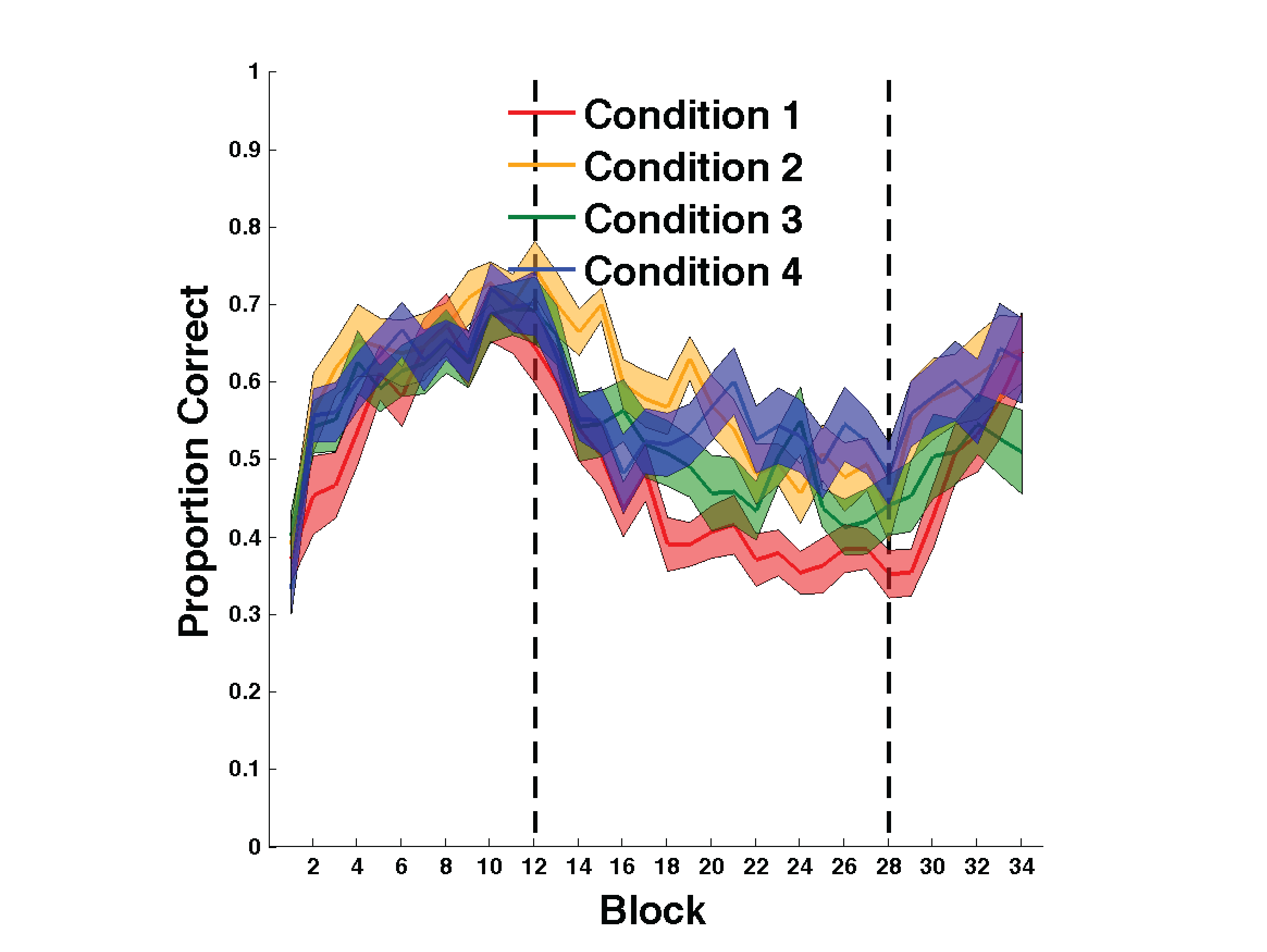
*Note*. A3 = the last 100 trials of acquisition, R1 = the first 100 trials of reacquisition, N = the number of participants contained in a given cell, and % = the percentage of responses accounted for by a particular model.

****

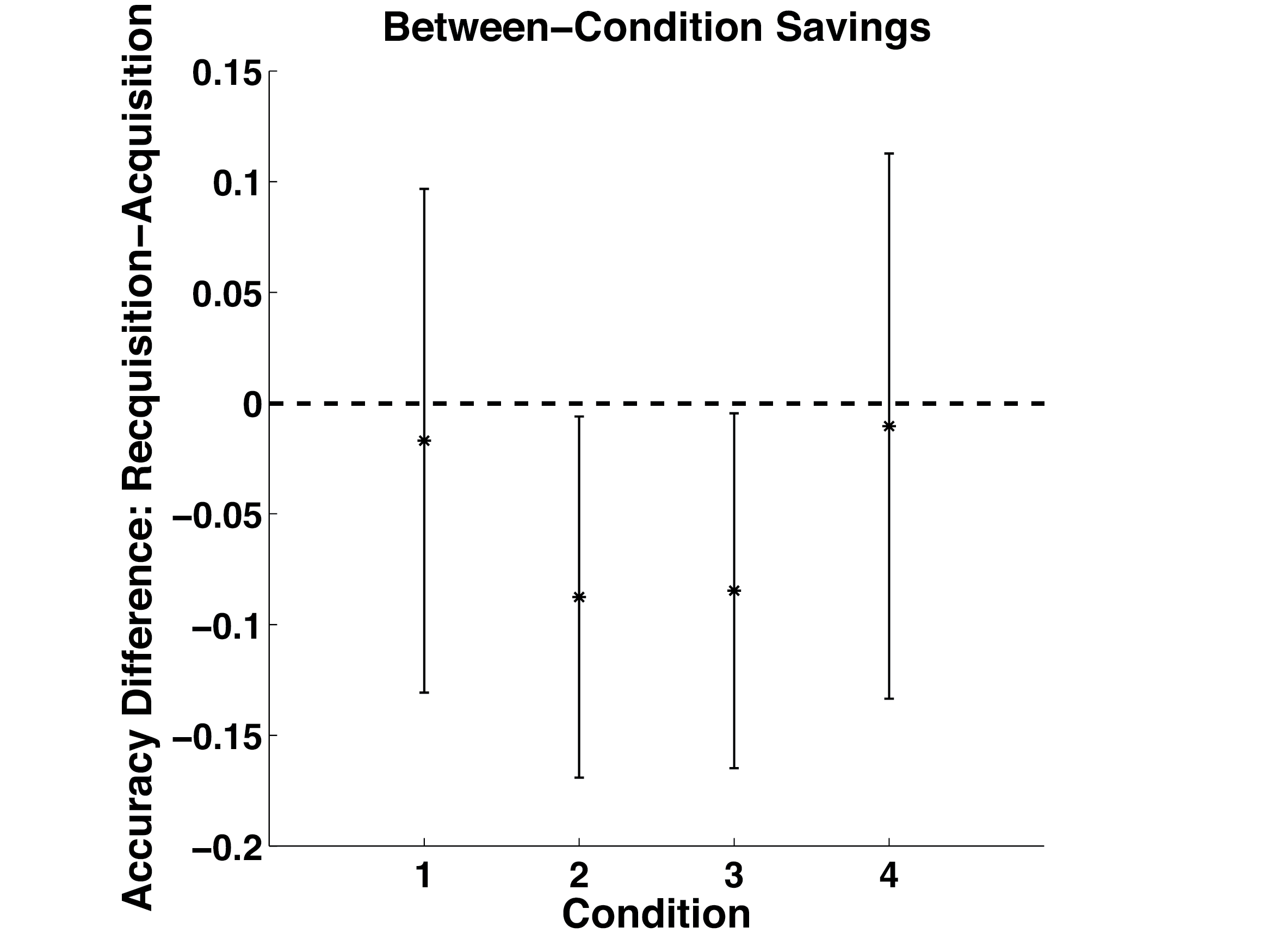
**Figure 1.** Left: A few examples of stimuli that might be used in an information-integration (II) category-learning experiment. Right: The category distributions used here.



**Figure 2.** Experiment design.



**Figure 3.** Experiment 1 accuracy. Each block includes 25 trials, and bands represent SEM. Top Panel: Accuracy in every condition throughout the entire experiment. Bottom Panel: Simulated results obtained from the model shown in Figure 5. Blocks 1-12 were in the acquisition phase, blocks 13-28 were in the intervention phase, and blocks 29-34 were in the reacquisition phase.



**Figure 4.** Savings in each condition and pairwise comparisons between all conditions. Significance of pairwise comparisons was determined using Tukey’s method. Error bars represent 95% confidence intervals on the mean accuracy per condition.