Concurrent Task Performance Enhances Procedural Skill 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

March 7, 2014

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

Skill learning is often plagued by the development of bad habits that are very difficult to fix. Moreover, the most problematic aspect of all current treatments for addiction and other maladaptive states is relapse. Thus, a variety of perspectives highlight the need for effective intervention protocols that can erase bad habits and replace them with stable good behaviors. Relapse is often estimated in laboratory settings as the amount of savings in relearning following an intervention protocol that causes some trained behavior to disappear. Savings is often inferred by observing that relearning occurs more quickly than original learning We recently reported the efficacy of various intervention protocols in unlearning procedural categorization (Crossley, Ashby & Maddox, 2013, J. of Experimental Psychology: General). This work demonstrated that an intervention protocol composed of random feedback effectively reduced accuracy during intervention, but did not eliminate savings in relearning (i.e., intervention failed to induce true unlearning). However, an intervention protocol composed of a mixture of random and veridical feedback did eliminate savings (i.e., caused true unlearning). This article shows that concurrently performing a demanding working memory task during an intervention phases consisting of random feedback reduces or fully eliminates savings in relearning. This suggests that the mechanism that protects learning during random feedback relies on prefrontal mechanisms.

Introduction

Skill learning is often plagued by the development of bad habits that are notoriously difficult to fix. For example, people who try to teach themselves golf (or any other skilled behavior for that matter) will usually end up with several gross technique flaws that can take months or even years of coaching to fix. Even then, every athlete that has tried to overcome a bad habit will tell you that the process is grueling, and often unsuccessful. Moreover, the neurobiological and psychological basis of skill learning overlaps with that of behaviors commonly used to study drug addiction and other maladaptive states. Thus, it may be the case that discovering treatments that help eliminate bad habits in skill learning, will also help to eliminate these dangerous and tragic states. The present experiment shows that engaging in an explicit dual task during treatment can enhance the efficacy of an intervention protocol intended to change a procedural skill.

The propensity for a behavior (good or bad) to persist is often estimated experimentally by measuring savings in relearning. These experiments first train a target behavior, and then expose participants to an intervention protocol that causes this behavior to disappear (e.g., a lever press in simple instrumental conditioning paradigms). Savings of the original learning is then inferred by observing that relearning occurs more quickly than original learning (e.g., rapid reacquisition). Savings in relearning is incredibly common across all forms of procedural skill learning, which reflects our anecdotal observation that bad habits in skill learning are notoriously difficult to fix.

The present experiment examined savings in relearning in an information-integration (II) category-learning task. A basic category learning task involves learning a many-to-one mapping between a multitude of stimuli and two or more responses. Stimuli are presented one per trial and category membership is learned through trial-and-error across many trials. II tasks are those in which stimuli are assigned to categories in such a way that it is difficult or impossible infer via explicit logical reasoning. More technically, II tasks are those in which accuracy is maximized only if information from two or more non-commensurable stimulus dimensions is integrated at some pre-decisional stage (Ashby & Gott, 1988). The II task we used, as well as illustrations of example trials, are shown in Figure 1. Note that each stimulus is a single black line that varies in length and orientation. The diagonal lines denote the category boundaries. Note that no simple verbal rule (e.g., short lines belong to category A, etc.) 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 an acquisition (300 trials), intervention (400 trials), and a reacquisition (150 trials) phase. These phases differed in the nature of the feedback provided after each response. During acquisition and reacquisition, feedback indicated whether each response was correct or incorrect. The feedback was random during the intervention phase – indicating a correct response with probability ¼ and an incorrect response 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 is 95%. In these ways, this task is identical to the task used in Experiment 1 of Crossley et al. (2013).

Crossley et al. (2013) found that an intervention protocol composed of purely random feedback did not eliminate savings. This implies that the brain possesses a mechanism through which it can detect when feedback has become random, and protect learning from being erased or overwritten during this time. The goal of the present experiment was to determine if this mechanism depends on prefrontal resources. Our approach was to attempt to interfere with the efficacy of this protective mechanism by reducing the neural resources available to it. Specifically, we introduced a concurrent numerical Stroop task (Algom, Dekel, & Pansky, 1996; Besner & Coltheart, 1979; Stroop, 1935) during a subset of acquisition and intervention categorization trials. Concurrent secondary tasks have long been thought to reduce resources available for a primary task (Pashler, 1994). In this case, the primary task would be detecting the onset of random feedback and protecting learning. Moreover, the evidence is strong that variants of the Stroop task rely on prefrontal resources (Bench et al., 1993). Thus, if this mechanism depends on prefrontal resources, then concurrent numerical Stroop performance will impair it’s ability to protect learning during random feedback. In this case, random feedback will lead to true unlearning. On the other hand, if the protective mechanism is independent of prefrontal resources then we should observe massive savings independent of the presence or absence of the Stroop task. We included a concurrent Stroop component on trials 251-350 in condition 1, 251-450 in condition 2, 251-550 in condition 3 and 400-650 in condition 4.

INSERT FIGURE 1 ABOUT HERE

Methods

*Participants*

There were 30 participants in Conditions 1 and 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. Note that these trials correspond to the last 50 trials of initial learning in which no dual task was present across all four conditions. Second, any participant whose mean overall performance on the Stroop task was less than 85% correct was excluded. 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. The experimental design for each condition is shown in Figure 2.

INSERT FIGURE 2 ABOUT HERE

*Stimuli*

Stimuli were black lines that varied across trials only in length (pixels) and orientation (degrees counterclockwise rotation from horizontal). The stimuli are illustrated graphically in Figure 1, and were identical to those used by Crossley et al. (2012).

*Procedure*

Participants in all conditions were told that they were to categorize lines on the basis of their length and orientation, that there were four equally-likely categories, and that high levels of accuracy could be achieved. At the start of each non-Stroop trial, a fixation point was displayed for 1 second and then the stimulus appeared. The stimulus remained on the screen until the participant generated a response by pressing the “Z” key for category A, the “W” key for category B, the “/” key for category C, or the “P” key for category D. Written instructions informed participants of the category label to button mappings. An “invalid key” message was displayed if any other button was pressed. The word “Correct” was presented for 1 second if the response was correct or the word “Wrong” was presented for 1 second if the response was incorrect (except during the intervention phase in which feedback was completely random).

Stroop trials began with a fixation point that was displayed for 1 second. The category stimulus and the Stroop stimuli (numbers flanking the category stimulus) were displayed simultaneously. After 200 ms the Stroop stimuli were replaced by white rectangles which remained on the screen until they made a category response. Responses emitted before the Stroop stimuli were replaced by white rectangles were not accepted. Feedback about the category response was given immediately in the same fashion as on non-Stroop trials. The word “value” or “size” then appeared on the screen prompting participants to indicate which side contained the numerically larger or the physically larger number. Participants pressed the ‘F’ key to choose the number on the left or the ‘J’ key to choose the number on the right. The word “Correct” was then again presented for 1 second if the response to the Stroop task was correct or the word “Wrong” was presented for 1 second if the response was incorrect. See Figure 1 for example trials both including and excluding the Stroop component.

Participants were instructed to try their hardest on both task components but to prioritize performance on the Stroop task. Both the category learning task and the Stroop task were explained to participants prior to beginning the experiment, and on screen messages warned them when the Stroop component would begin, and again when it would end. These messages read, “You will now perform both the categorization task and the paired numbers task simultaneously. Keep trying your hardest!” and “You have now finished the section with the paired numbers task. You will now be shown only the line categorization task. Keep trying your hardest.” 85% of Stroop trials the numerically larger number was physically smaller. The proportion of Stroop trials that prompted “size” or “value” was split 50/50. Accuracy on the numerical Stroop task was indicated at the top of the screen when they received feedback regarding their performance on the concurrent task on each trial. This score was displayed in green if it was above 80% and red if it was below 80%.

Results

*Accuracy-based results*

The top panel of Figure 3 displays the mean accuracy for each 25-trial block in all conditions. 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, with the exception of the concurrent Stroop task components. Note that participants from all conditions were able to learn the categories, and that accuracy dropped significantly during intervention. Figure 4 shows that conditions 1, 2 and 4 trended towards savings, but failed to reach significance, and condition 3 showed an interference.

INSERT FIGURE 3 ABOUT HERE

All effects in a 4 condition (1, 2, 3, 4) × 3 phase (Acquisition, Intervention, Reacquisition) repeated measures ANOVA were significant [Condition: F(3, 69) = 6.56, p < 0.001; Phase: F(2, 2596) = 120.05, p < 0.001; Interaction: F(6, 2596) = 4.93, p < 0.001]. The key result was that savings (estimated as the first 6 blocks of acquisition - all blocks of reacquisition) in conditions 1, 2, and 4 trended towards savings but failed to reach significance, and condition 3 showed a significant impairment. [1: t(2586) = -0.11, p = 0.91; 2: t(2586) = -0.74, p = 0.46; 3: t(2586) = 2.35, p < 0.05; 4: t(2586) = -1.82, p = 0.07]. Between-condition comparison of savings is complicated by the results of serval pairwise comparisons that show significant differences between the groups during acquisition and intervention (See Table 1 for details). For instance, accuracy during acquisition in condition 3 was higher than it was in condition 1 or 2. Accuracy during intervention was higher in condition 3 than it was in conditions 1 or 2. Accuracy in condition 4 was higher than condition 1 during intervention. Accuracy in conditions 3 and 4 were higher than in conditions 1 and 2 during reacquisition. However, note that none of these result will have any effect on our within-condition estimate of savings.

INSERT FIGURE 4 ABOUT HERE

INSERT TABLES 1 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, and some strategies can even generate accuracy levels similar to those observed in our experiment. 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. Note that the proportion of II users during the last block of acquisition in every condition is about the same (69%) and that this proportion is reduced by a similar amount in every condition (about 45%). In conditions 1, 2, and 3 this drop in II users is driven almost entirely by an increase in the number of guessers, but in condition 4 this drop is driven by an increase in the number RB users. This suggests that while condition 4 showed the best evidence for savings (though it failed to reach significance), this most likely reflects a shift to suboptimal RB responding and does not reflect a true recovery of the response strategy used during acquisition.

INSERT TABLE 2 ABOUT HERE

Discussion

We previously reported savings in procedural categorization after an intervention phase consisting of the presentation of random feedback (crossley et al. 2013). The current results show that a concurrent numerical Stroop task can significantly reduce or else eliminate savings in this paradigm. The trend towards savings in condition 4 suggests that it is critical for the Stroop task to be introduced prior to the transition from the acquisition phase to the intervention phase. The significant interference in condition 3, but not in conditions 1 or 2, suggests that the longer the Stroop task remains present during the intervention, the smaller the savings.

Crossley et al. (2013) proposed a model that assumed that procedural category learning is instantiated via plasticity at cortical-striatal synapses and that this plasticity is gated by striatal cholinergic interneurons (called TANs for tonically active neurons). They assumed that the TANs act as a gate on procedural learning, and that during periods of random feedback they learn to prevent the modification and expression of procedural skills. This model successfully accounted for a broad array of savings-based phenomena, while simultaneously respecting a range of neurobiological constraints including single-cell recordings from striatal projection neurons and TANs. This model did not, however, specify a neurobiological mechanism through which the TANs could become sensitive to the presence of random feedback. This article was aimed at testing the hypothesis that this mechanism depends on prefrontal resources, and our results are indeed consistent with this idea.

References – Needs updating

Algom, D., Dekel, A., & Pansky, A. (1996). The perception of number from the separability of the stimulus: The Stroop effect revisited. Memory & Cognition, 24(5), 557-572.

Bench, C., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S. J., & Dolan, R. J. (1993). Investigations of the functional anatomy of attention using the Stroop test. Neuropsychologia, 31(9), 907-922.

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 and Table Legends

Table 1. Post-hoc pairwise comparisons. Shaded cells indicate significant differences.

Table 2. 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.

Figure 1. Figure 1. Top Left: Sequence of events on an example trial without the Stroop component. Top Right: Sequence of events on an example trial including the Stroop component. Bottom 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. See text for further details.

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 defined as the mean accuracy during the the first 150 trials of acquisition - the mean accuracy during reacquisition. Error bars represent 95% confidence intervals on the mean accuracy per condition.

Table 1. Post-hoc pairwise comparisons. Shaded cells indicate significant differences.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Acquisition Phase | Intervention Phase | Reacquisition Phase |
| 1 – 2 | *t*(77) = 0.71, *p = 0.48* | *t*(70) = -1.31, *p = 0.19* | *t*(104) = -0.19, *p = 0.85* |
| 1 – 3 | *t*(74) = -2.95, *p < 0.01* | *t*(68) = -4.36, *p < 0.001* | *t*(100) = -2.11, *p < 0.05* |
| 1 – 4 | *t*(62) = -0.86, *p = 0.40* | *t*(58) = -2.98, *p < 0.01* | *t*(82) = -2.17, *p < 0.05* |
| 2 – 3 | *t*(269) = -4.39, *p < 0.001* | *t*(238) = -3.66, *p*  < 0.001 | *t*(409) = -2.21, *p < 0.05* |
| 2 – 4 | *t*(75) = -1.54, *p = 0.13* | *t*(69) = -1.80, *p* = 0.08 | *t*(101) = -2.02, *p < 0.05* |
| 3 – 4 | *t*(73) = 1.90, *p = 0.06* | *t*(67) = 1.03, *p* = 0.31 | *t*(98) = -0.24, *p* = 0.81 |

Table 4. 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.

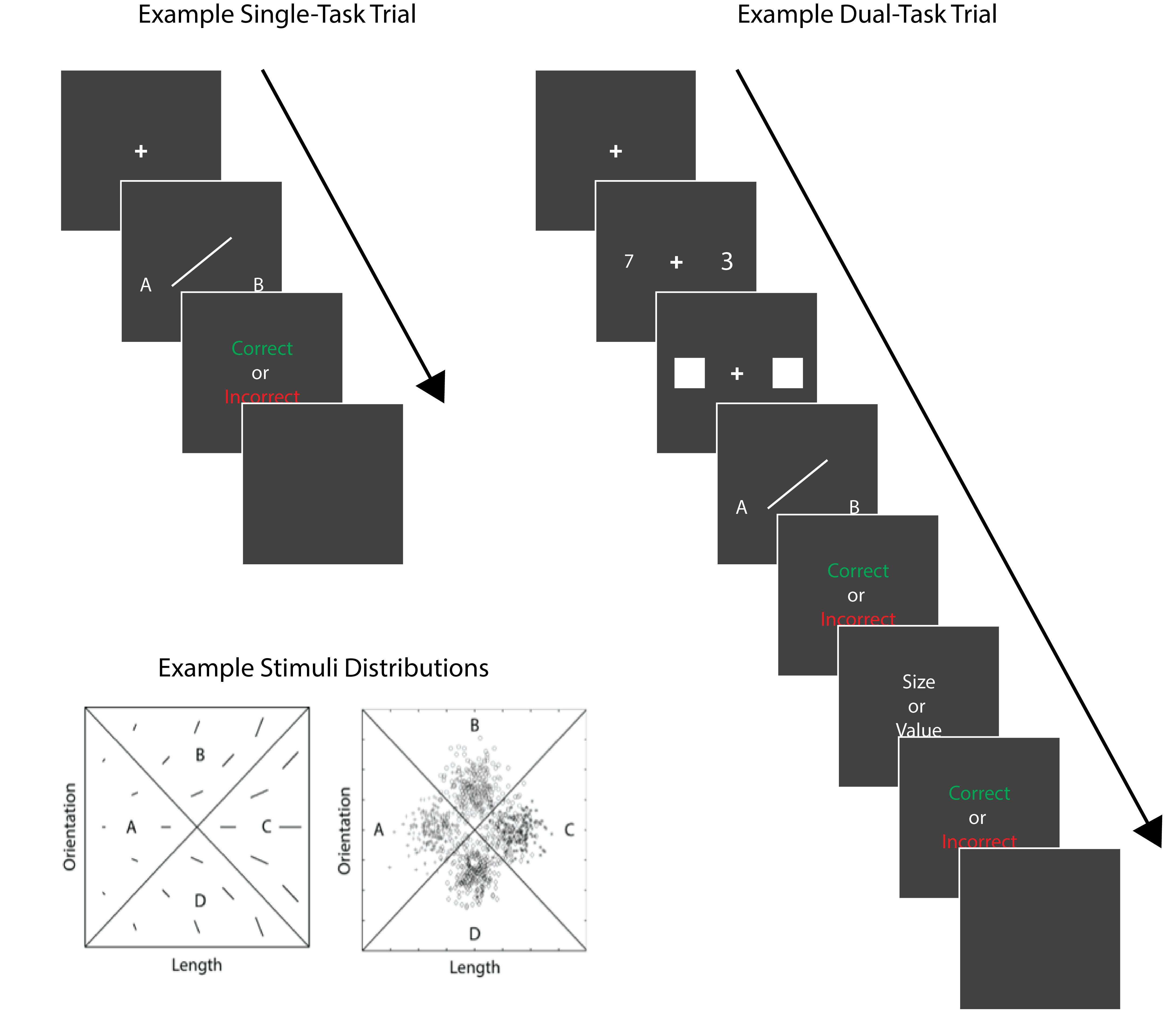


Figure 1. Top Left: Sequence of events on an example trial without the Stroop component. Top Right: Sequence of events on an example trial including the Stroop component. Bottom Left: A few examples of stimuli that might be used in an information-integration (II) category-learning experiment. Right: The category distributions used here.

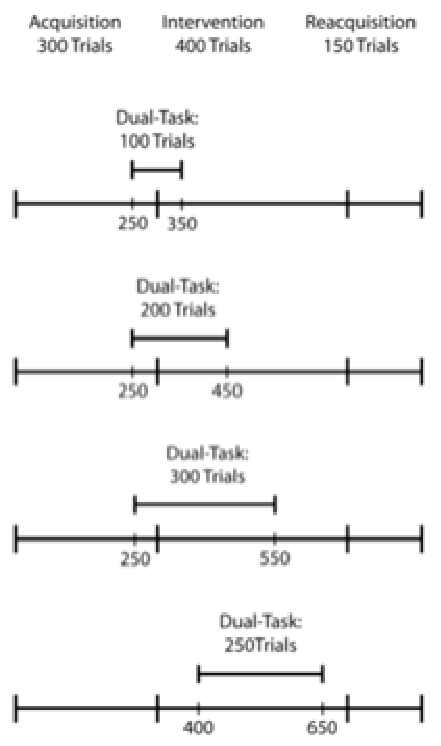


Figure 2. Experimental design. See text for further details.

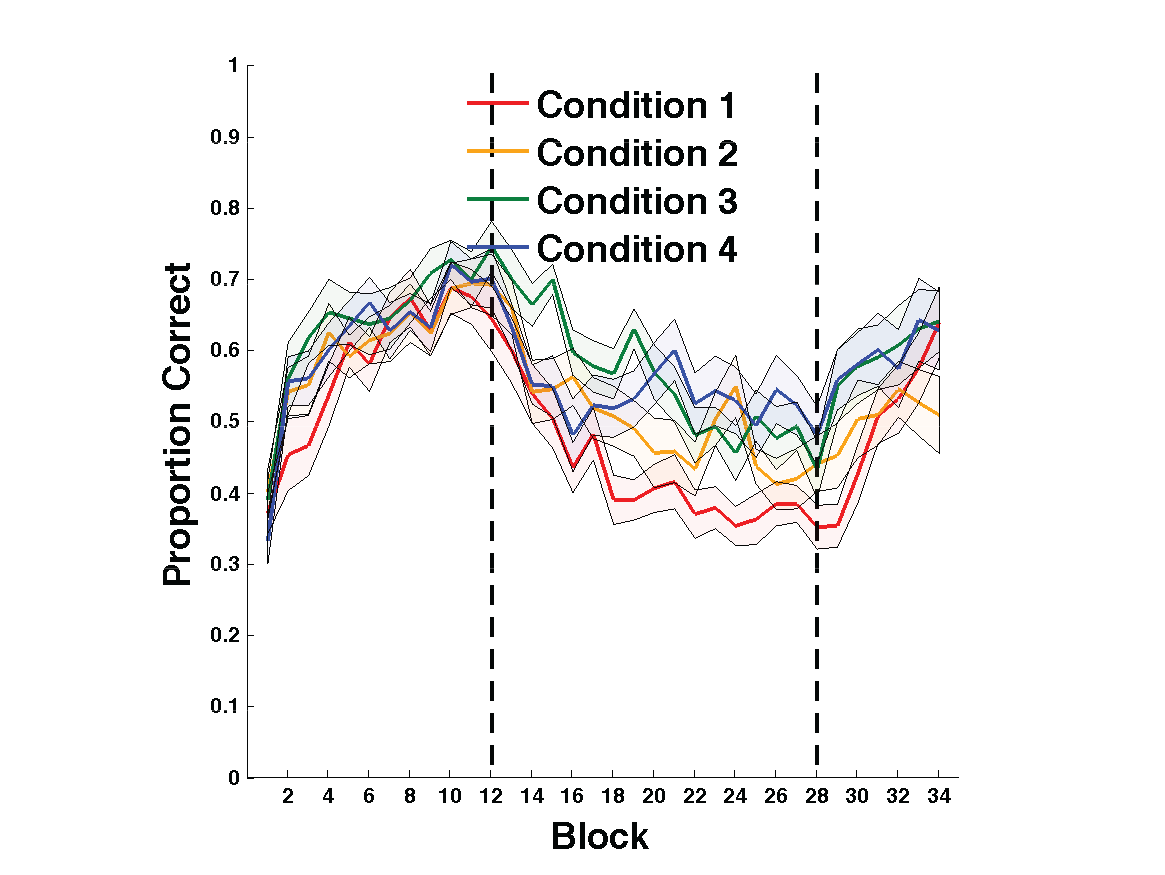


Figure 3. Experiment 1 accuracy. Each block includes 25 trials, and bands represent SEM. Accuracy in every condition throughout the entire experiment. Blocks 1-12 were in the acquisition phase, blocks 13-28 were in the intervention phase, and blocks 29-34 were in the reacquisition phase.

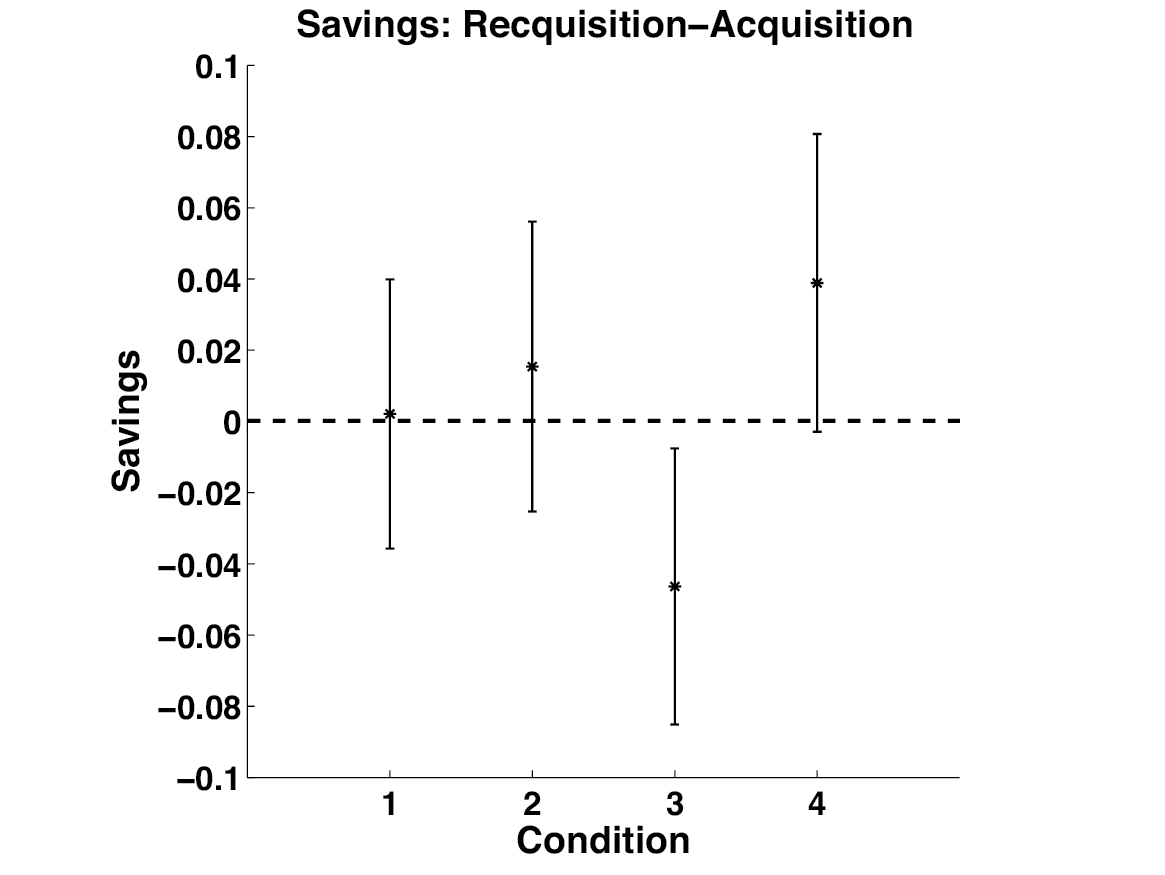


Figure 4. Savings in each condition defined as the mean accuracy during the the first 150 trials of acquisition - the mean accuracy during reacquisition. Error bars represent 95% confidence intervals on the mean accuracy per condition.