Incidental Learning and Task Boundaries

Michael Freedberg, Tana T. Wagschal, and Eliot Hazeltine
University of Iowa

For skill learning processes to be effective, they must encode associations that are inherent to the current task and avoid those that are spurious or particular to training conditions so that learning can transfer to novel situations. Some everyday contexts even require grouped responding to simultaneously presented stimuli. Here we test whether learning of these grouped responses depends on overlap in stimulus and/or response modality or on the conceptualization of the stimulus and response streams as belonging to a common task. In the present experiments, participants made 2 responses to 2 simultaneously presented stimuli, and learning was assessed by comparing performance on response combinations that had been practiced throughout training to performance on combinations that had been withheld. Experiments 1–4 paired the same visual-manual task with a 2nd task that differed in terms of the stimulus modality, the response modality, neither modality, or both modalities. Combination-specific learning was only observed when both the stimulus and response modalities were the same for the 2 tasks. However, Experiments 5 and 6 showed that combination-specific learning could occur with nonoverlapping stimulus modalities or response modalities if the 2 tasks were conceptually related. The results suggest that task representations provide top-down constraints on skill learning processes.

Keywords: associative learning, task integration, top-down processes, incidental learning, task boundaries

The human brain has a remarkable capacity for detecting contingencies and using them to optimize behavior. This ability fundamentally shapes human behavior across a range of domains, enabling us to benefit from previous experience and discover structure in a chaotic world. Given the welter of information confronting the brain, the number of possible contingencies associated with any given event may be nearly uncountable. Thus, there is a binding problem: How do we efficiently select the sources of information that are truly predictive of each other from among the countless spurious contingencies provided by chance?

Researchers have explored the boundaries restricting associative learning using tasks in which contingencies occur among stimuli along multiple modalities (Conway & Christiansen, 2006; Hazeltine, 2002; Hsiao & Reber, 2001; Keele, Ivry, Mayr, Hazeltine, & Heuer, 2003; Kramer, Wickens, & Donchin, 1985; Schmidtke & Heuer, 1997). In such a task, Conway and Christiansen (2006) observed that participants were able to encode sequential information from two streams (auditory and visual) simultaneously. Critically, accurate recognition of two grammars of distinct modalities (visual and auditory) during a

posttraining probe surpassed the recognition of two visual grammars (colors and symbols), two auditory grammars (tones and nonwords), and single-modality grammars (auditory or visual). Thus, the researchers concluded that learning is constrained within modalities and is guided by stimulus-specific rules rather than amodal representations.

This account shares critical features with the framework offered by Keele et al. (2003), who proposed that two learning systems are engaged to flexibly form associations between discrete events. According to this account, the *unidimensional* system integrates information within a dimension but not between dimensions, and operates independently of attention. This system may comprise a set of learning mechanisms each associated with a specific modality. Thus, irrelevant information occurring along distinct modalities should not interfere with the encoding of these sequences because that information is not gated to these modality-specific mechanisms. If it is assumed that the task used by Conway and Christiansen (2006) primarily taps the unidimensional system, possibly because the cross-modal transitions in their task were random, then their results are consistent with the Keele et al. model.

But the story is not so simple; crossmodal contingencies are encoded under certain task conditions (e.g., Curran & Keele, 1993; Schmidtke & Heuer, 1997). To accommodate such findings, Keele et al. (2003) proposed a second learning system, the *multidimensional* system, that is capable of forming associations across modalities, provided the events are attended. Thus, the key constraint on the multidimensional system is the degree to which attentional resources can be dedicated to the tasks. In this way, the model can account for the learning of a wide array of associations with the constraint that attention is required for the encoding of crossmodal contingencies.

While this account can explain why associations between events occurring along distinct modalities are formed in some cases and

This article was published Online First April 21, 2014.

Michael Freedberg, Tana T. Wagschal, and Eliot Hazeltine, Department of Psychology, University of Iowa.

The research was supported by National Institutes of Health Grant R03 DA031583-01A1. We wish to thank Robert Gaschler and Rich Ivry for their extremely helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Eliot Hazeltine, Department of Psychology, University of Iowa, 113 Seashore Hall E, Iowa City, IA 52242-1407. E-mail: Eliot-hazeltine@uiowa.edu

not in others, the availability of attention does not always provide a succinct account of when crossmodal contingencies are formed, especially if data from other experimental procedures are considered. For example, a dual-task study by [Hazeltine, Teague, and Ivry \(2002\)](#) tested the hypothesis that reductions in dual-task costs reflected the learning of specific crossmodal associations across the two tasks. Dual-task procedures are typically used to examine how performing one task interferes with another, not to study the boundaries of learning processes; all possible combinations of stimuli are usually presented, and it is assumed that the two tasks are learned independently. In the [Hazeltine et al.](#) study, participants performed a visual-manual task and an auditory-vocal task in which the stimuli for the two tasks were presented simultaneously and the responses were made as quickly as possible. Critically, unlike typical dual-task experiments, they withheld combinations of stimuli during training, and, after several sessions, introduced the withheld combinations. These withheld combinations were performed as well as practiced combinations, suggesting no cross-modal learning. Given that both stimuli required a response and therefore must have been attended, it is unclear why the cross-modal associations were not encoded by the multidimensional system.

The Chord Task

To resolve this puzzle, we turn to the chord task. The chord task requires participants to make two or more simultaneous responses, each indicated by a distinct stimulus ([Seibel, 1963](#)). In this way, the task is analogous to a pianist forming chords on a musical instrument by producing simultaneous keypresses. However, these conditions are generalizable to many everyday tasks in which we configure our limbs to perform a specific task, such as when we grasp a jar to open it or position our hands to swing a golf club.

Although it was initially believed that learning reflected more efficient translation of the individual stimulus–response associations within the chord ([Seibel, 1963](#)), subsequent work showed that a large portion of the learning-related improvements in performance were combination specific ([Hazeltine, Aparicio, Weinstein, & Ivry, 2007](#)). To show this, performance of novel combinations of responses was compared to practiced combinations, and the difference was taken as a measure of combination-specific learning, as in the [Hazeltine et al. \(2002\)](#) dual-task study. That is, even though the individual responses in the novel and practiced chords were the same, novel chords were performed more slowly than practiced chords, indicating that learning involved integration of the individual responses comprising each chord. Moreover, learning appeared to entail the encoding of both stimulus and response information ([Hazeltine et al., 2007](#); [Wifall, McMurray, & Hazeltine, 2012](#)), and as with the serial reaction time (SRT) task, learning occurred without explicit awareness. In fact, awareness was uncommon in the chord task even when participants are not performing a secondary task ([Hazeltine et al., 2007](#)).

In the present study, we use the chord task to examine how task demands affect the boundaries of learning processes by measuring combination-specific learning. Because individual responses can be signaled by arbitrary stimuli that can be presented along a range of modalities, the chord task is ideal for examining crossmodal learning.

Note that when just two responses are required in the chord task, the task demands are highly similar to those in a typical dual-task setting: Two stimuli are presented, and the participant must respond to both of them, although the chord task adds the constraint that the stimuli are presented simultaneously and the responses must be made simultaneously as well. The similarity between the two procedures poses the question, why was no combination-specific learning observed by [Hazeltine et al. \(2002\)](#) when combination-specific learning is consistently observed in chord tasks (e.g., [Hazeltine et al., 2007](#); [Wifall et al., 2012](#))? Differences in the deployment of attention do not seem to be a promising explanation given the similarity of the task demands and the fact that in all the experiments, all the stimuli must be attended because they all require a speeded response.

There are several important differences between the chord learning experiments and the dual-task experiments conducted by [Hazeltine et al. \(2002\)](#). First, participants in the dual-task study practiced for several 1-hr sessions, whereas the chord learning studies of [Hazeltine et al. \(2007\)](#) were single-session experiments. Thus, there was much less practice in the chord learning studies.¹ However, [Wifall et al. \(2012\)](#) obtained similar chord-specific learning after eight sessions, suggesting that this is not the critical factor. In the dual-task study, there were only nine possible response combinations, three of which were withheld, whereas the chord learning studies used larger stimulus–response sets, so there were at a minimum 16 possible combinations across the various experiments. Furthermore, although the stimuli were presented simultaneously for both sets of experiments, participants in the chord experiments were instructed to make their responses simultaneously. Participants in the dual-task experiment were instructed simply to make both responses as quickly as possible. On average, they produced the manual response 66 ms before the vocal response.

A fourth major difference between the chord and dual-task studies involves the modalities of the stimuli and responses of the various stimulus–response associations. In the chord experiments, all the stimuli were visual and all the responses were manual. In the dual-task experiments of [Hazeltine et al. \(2002\)](#), one stimulus on each trial was visual and the other was auditory, and one response was manual and the other was vocal. That is, both the stimulus and response modalities were different for the two tasks in the dual-task experiment, but they were the same for the chord experiments. If, as hypothesized by [Keele et al. \(2003\)](#) and [Conway and Christiansen \(2006\)](#), modalities play a role in the types of associations that can be formed, then this difference in the experimental procedures could explain the different patterns of learning. Thus, in the Experiments 1–4, we examine the role that modalities play in combination-specific learning.

Top-Down Contributions to Associative Learning

While the stimulus and response modalities may play a role in determining the boundaries of learning, several studies indicate that top-down factors may play a critical role (e.g., [Melchers, Shanks, & Lachnit, 2008](#); [Turk-Browne & Scholl, 2009](#)). For

¹ In [Seibel's \(1963\)](#) classic study, participants practiced for 75 half-hour sessions. However, he used different measures of learning and never compared practiced combinations to unpracticed combinations.

example, Melchers et al. (2008) reviewed evidence from associative learning studies providing evidence of flexible informational acquisition. They pointed out that traditional theories of associative learning typically adopt one of two views. Elemental theories posit that associations are learned based on the degree of similarity between individual properties of stimuli or their parts (Atkinson & Estes, 1963; Harris, 2006; Rescorla & Wagner, 1972). In contrast, configural theories hold that the strength of the association is a function of the combined properties of the stimulus, independent of its parts (Pearce, 1987, 1994; Wasserman & Miller, 1997). While both theories present inherent constraints to the limits of associative learning, Melchers et al. argued that task demands may promote flexibility to the learning process. They proposed that elemental or configural processing can be induced depending on the participant's representation of the task.

Evidence for this claim comes from several studies of associative learning. For example, Williams and Braker (1999) were able to encourage either elemental or configural strategies based on pretreatment demands. As noted by Melchers et al. (2008), similar biasing effects were observed by manipulating top-down processes such as prior experience (Mehta & Williams, 2002; Williams, Sagness, & McPhee, 1994), experimental instructions (Williams et al., 1994), stimulus properties (Lachnit, 1988), and stimulus organization (Glaetzer, 2002; Melchers, Lachnit, Üngör, & Shanks, 2005). Top-down processing can also influence visual (Chun & Jiang, 1999; Li, Piëch, & Gilbert, 2004) and auditory learning (Polley, Steinberg, & Merzenich, 2006; Tillman, Bharucha, & Bigund, 2000) and affect sequence learning (Halvorson, Ebner, & Hazeltine, 2013) and induce neural plasticity (Sigman & Dehaene, 2005; see also Gilbert & Sigman, 2007).

Configural learning is clearly consistent with the combination-specific learning that is observed in the chord task but not in dual-task experiments. Thus, the finding that the type of learning (configural vs. elemental) can depend on top-down factors is clearly relevant to our investigation. In particular, it is possible that the dual-task experiments (Hazeltine et al., 2002) and chord task experiments (Hazeltine et al., 2007) differ not just in terms of their stimulus and response modalities, but also in the how the modalities bias top-down mechanisms that influence learning. For instance, participants under dual-task conditions, and not chord task conditions, may have been unable to adopt a configural learning strategy because the two tasks under dual-task conditions were distinct and unrelated. That is, participants may have found it easier to integrate the chord task (Hazeltine et al., 2007) than in the dual-task experiment (Hazeltine et al., 2002) because of the greater similarities between the two responses in the chord task.²

In sum, stimulus and response modalities may contribute to the perceived relatedness of individual stimulus–response pairs; that is, the modalities may affect whether pairs belong to the same or different task representations. By doing so, they may indirectly determine the types of associations that can be formed by affecting the top-down processes that organize the tasks. In short, modalities may play a role in establishing task boundaries and thus affect learning (i.e., whether it is configural or elemental) by proxy. Therefore, while Experiments 1–4 investigate whether modalities can play a role in combination-specific learning, Experiments 5 and 6 pit the modality hypothesis against the task representation hypothesis by examining whether associations can be formed

between stimulus–response pairs that involve distinct modalities but are conceptually related.

Experiment 1

As a starting point, we eliminate some of the procedural differences between the dual-task and chord experiments to determine whether the modality pairings can play a role in the encoding of contingencies in the chord task. Participants performed two tasks: a visual-manual task in which three spatially differentiated stimuli were mapped to three button press responses and an auditory-vocal task in which three differently pitched tones were mapped to three single syllable words. Although the tasks were similar to those used in dual-task studies (e.g., Hazeltine et al., 2002; Pashler, 1994; Schumacher, Seymour, Glass, Kieras, & Meyer, 2001), the task requirements were identical to those used in the chord learning experiments of Hazeltine et al. (2007). Participants performed only a single session and were instructed to make both responses at the same time. A subset of stimulus combinations were withheld during training, and at the end of the session, these withheld combinations were presented to assess learning. If combination-specific learning is dependent on features of the task, such as the requirement that responses be made simultaneously or the amount of practice, then combination-specific learning should be observed in the experiment. If combination-specific learning depends on overlapping stimulus and response modalities for the two tasks, then no combination-specific learning should be observed. In short, the experiment uses the task structure from the chord experiments that show combination-specific learning but the stimulus and response modalities from the dual-task experiments that show no combination-specific learning.

Method

Participants. Sixteen participants were recruited from the undergraduate population at the University of Iowa. The participants performed for credit in an introductory psychology class. All participants agreed to participate in the study in accordance with the guidelines of the university's Committee for the Protection of Human Subjects.

Apparatus and stimuli. Participants responded manually to visual stimuli presented on the computer monitor of a PC-clone computer and vocally to auditory stimuli presented through the computer's speakers. The visual stimuli were white Xs presented against a black background. The Xs subtended 1° of visual angle and were centered 2°, 3.5°, and 5° to the left of a fixation cross. The locations of the Xs were indicated by place-holding lines just below where the Xs could appear. The placeholders and fixation cross were displayed throughout the trial. Only a single X appeared on any given trial and was extinguished as soon as the participant made the manual response. Manual responses were made on a computer keyboard using the *S*, *D*, and *F* keys with the ring, middle, and index fingers of the left hand, respectively. The mapping of the Xs to keys was spatially compatible, so the leftmost X was mapped to the *S* key, etc.

² Experiments 3–6 investigate exactly what is meant by “greater similarity.”

The auditory stimuli were pure tones lasting 250 ms. The tones were 220, 880, or 3520 Hz with the volume adjusted to the participants' comfort. Vocal responses were detected by a microphone placed by the participants' mouth and identified with the Microsoft speech recognition software that was integrated with the experimental program written in C++. Participants were instructed to say "bug" when they heard the 220-Hz tone, "food" when they heard the 880-Hz tone, and "tree" when they heard the 3520-Hz tone. The arbitrary response labels were used because they did not correspond with the visual stimuli and thus would not produce cross-task interference.

There are nine possible combinations of visual and auditory stimuli, and three combinations were withheld for each participant. To counterbalance the withheld combinations, the participants were divided into four groups. For each group, three combinations were chosen so that each individual stimulus was in exactly one of the withheld combinations. In this way, each of the stimuli remained equally probable during the practice blocks. As depicted in Figure 1A, this can be conceptualized as choosing three cells in a 3×3 grid. The actual correspondence between particular columns and visual stimuli, and between particular rows and auditory stimuli, is not arbitrary. Because the stimuli vary along a continuous dimension—location for the visual stimuli and pitch for the auditory stimuli—the ordinal relationships between the two stimulus–response combinations may lead to correspondence effects (e.g., Ben-Artzi & Marks, 1995). That is, an extreme value (e.g., leftmost or rightmost) on one dimension (e.g., location) might correspond to an extreme value (e.g., lowest or highest) on the other dimension (e.g., pitch). In fact, the data showed this pattern (see below).

For each of the four groups, two of the noncorresponding combinations were withheld and two were practiced (see Figure 1B), and the remaining withheld combination served as a sort of counterweight so that all stimuli and responses, when considered individually, would be equally likely. With this method, four of the six practiced combinations held some form of correspondence, compared to only one of the three withheld combinations. Moreover, some participants had withheld combinations with one form of correspondence (e.g., leftmost X with highest tone) and other participants had withheld combinations with a different form of correspondence (e.g., leftmost X with lowest tone). So that these

correspondence effects did not present any confounds, we restricted our primary analyses to consideration of only the four noncorresponding combinations, two of which were withheld and two of which were practiced. These are the combinations consisting of the middle value on one dimension and the nonmiddle value on the other dimension. Thus, to obtain our measure of learning, we conducted transfer blocks in which all combinations were presented and compared performance on the withheld noncorresponding trials to performance on the practiced noncorresponding trials.

Procedure. Each trial began with the presentation of the fixation cross for 500 ms. The stimuli were then presented for 250 ms. On dual-task trials, the visual and auditory stimuli were presented simultaneously. After the stimuli were extinguished, participants had 3,000 ms to produce their responses. Once a response had been detected for each presented stimulus or 3,000 ms had elapsed, the fixation cross was extinguished and the 500 ms intertrial interval began. The next trial was initiated after the completion of the intertrial interval.

Participants were instructed to make their responses as quickly and accurately as possible. They were also instructed to make both responses simultaneously on dual-task trials. They were not informed that any combinations were withheld or more likely than any others.

Participants completed 12 blocks of trials. The first two blocks entailed only one of each task. In these single-task blocks, there were 20 trials of each of the three possible stimuli of one modality, for a total of 60 trials. The next eight blocks (Blocks 3–10) were training blocks consisting of 30 dual-task trials and 30 single-task trials. For the single-task trials, there were 15 of each modality (five of each stimulus). For the dual-task trials, only six of the nine possible stimulus combinations were presented, with each of the six combinations presented five times in a block. As depicted in Figure 1B, the actual withheld combinations were counterbalanced across participants. The final two blocks (Blocks 11–12) were probe blocks used to assess learning. On these blocks, there were 36 dual-task trials, comprising four instances of each possible combination, including the withheld combinations, and 24 single-task trials, comprising four instances of each of the six single stimuli (three visual and three auditory). The order of the trials was random.

Participants were permitted to take short breaks between the blocks as needed, but there was no additional break between the training and probe blocks. None of the participants reported being aware of any difference between the training and probe blocks.

Results and Discussion

Reaction times. The first two blocks were treated as practice and not included in the primary analysis. The first trial of each block was also treated as practice. Moreover, trials with an incorrect response or a reaction time (RT) less than 250 ms or greater than 2,000 ms were removed from the analyses, eliminating 9% of the remaining data. For the comparisons between practiced and withheld combinations, we eliminated exact repetitions of the combinations.

The mean manual and vocal RTs across blocks are depicted in Figures 2A and 3A. The data are divided according to trial type, defined by the relationship between the two stimuli. Corresponding pairs consisted of the leftmost visual stimulus and the low-

Figure 1. (A) Example set of withheld combinations in which only six of the possible nine pairings are presented during practice. (B) The four sets of withheld combinations that were evenly distributed across the participants in each experiment. Stim = stimulus.

Figure 2. (A) Mean reaction times for the visual-manual task across the eight practice blocks for the four trial types in Experiment 1. (B) Mean reaction times for the visual-manual task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 1. The error bars represent standard errors of the mean.

pitched tone, the middle visual stimulus and the midpitched tone, and the rightmost visual stimulus and the high-pitched tone. Mirror pairs consisted of the leftmost visual stimulus and high-pitched tone and the rightmost visual stimulus and the low-pitched tone. Noncorresponding combinations included the four pairings of a middle stimulus on one modality and a nonmiddle stimulus on the other modality, as described above.

Training performance: Visual-manual task. The RTs for the visual-manual task were submitted to a two-way analysis of variance (ANOVA) with block (3–10) and trial type (single, corresponding, mirror, and noncorresponding) as factors. Both main effects—block, $F(7, 105) = 3.60$, $p < .005$, $\eta_p^2 = .19$; trial type, $F(3, 45) = 64.17$, $p < .001$, $\eta_p^2 = .81$ —and their interaction, $F(21, 315) = 1.73$, $p < .05$, $\eta_p^2 = .10$, were significant. As shown in Figure 2A, RTs generally decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials. Moreover, the decrements in RT with practice were smaller for the mirror trials (–11 ms) than for the single (165 ms), corresponding (94 ms), and noncorresponding (119 ms) trial types.

Post hoc comparisons revealed that the noncorresponding trials (mean RT = 831 ms) were performed significantly slower across blocks than any of the other trial types ($t_s > 2.69$, $p_s < .05$), and the single-task trials (565 ms) were performed significantly faster across blocks than any of the other trial types ($t_s > 9.04$, $p_s < .0001$). The RTs for the corresponding (756 ms) and mirror (778 ms) trials did not differ significantly across blocks, $t(15) = 1.30$, $p = .21$. The RT differences between the noncorresponding and corresponding (and mirror) trials indicate that the relationship between abstract properties of the stimulus–response pairs was sufficient to affect performance, despite the fact that there was no inherent similarity between the stimuli of the two tasks or between the responses of the two tasks. This finding is consistent with the notion that interresponse compatibility effects primarily stem from conceptual properties of the responses rather than motor codes (see Hazeltine, 2005).

Combination-specific learning: Visual-manual task. To determine whether participants encoded specific combinations of responses or simply were able to do the two tasks together more efficiently (i.e., reduce their dual-task costs), we compared the

performance of the noncorresponding practiced combinations to the performance of the noncorresponding withheld combinations on Blocks 11 and 12. To simplify the analyses and maximize power, we pooled the data across the two blocks. As shown in Figure 2B, there was little indication of any difference between the practiced (839 ms) and withheld (843 ms) combinations, $t(15) < 1$.³ There were also no significant differences between practiced and withheld combinations for the corresponding (739 ms vs. 702 ms; $t < 1$) and mirror (762 ms vs. 842 ms; $t < 1$) combinations. Thus, the RT reductions observed during training appear to reflect that the two tasks were learned independently. This finding is consistent with the results of another experiment pairing a visual-manual task with an auditory-vocal task in Hazeltine et al. (2002) but inconsistent with experiments pairing two visual-manual tasks (Hazeltine et al., 2007).

Training performance: Auditory-vocal task. The RTs for the auditory-vocal task were submitted to the same two-way Block $\times$ Trial Type ANOVA applied to the visual-manual data. Both main effects—block, $F(7, 105) = 8.24$, $p < .0001$, $\eta_p^2 = .36$; trial type, $F(3, 45) = 16.56$, $p < .001$, $\eta_p^2 = .53$ —were significant, but not their interaction, $F(21, 315) < 1$. As shown in Figure 3A, RTs generally decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials.

Unlike in the visual-manual task, the mean RT across blocks for single-task trials in the auditory-vocal task (793 ms) did not differ significantly from the mean RTs for the corresponding (793 ms) and mirror (794 ms) trials, which also did not differ from each other ($t_s < 1$). Thus, participants slowed their visual-manual responses so that they were performed at the same time as the auditory-

³ Given that participants may have begun to learn the withheld combinations once they were presented, we also performed the analyses considering only the data from Block 11. The pattern of data was nearly identical, and differences were not significant ($t < 1$), for either task; in fact, RTs for withheld combinations were slightly lower than for practiced combinations. We also performed analyses on Block 11 data only for the other experiments reporting no significant differences between practiced and withheld combinations (Experiments 3 and 4), and again found that there were no differences between practiced and withheld combinations in these experiments when only data from Block 11 were considered (all $t_s < 1$).

Figure 3. (A) Mean reaction times for the auditory-vocal task across the eight practice blocks for the four trial types in Experiment 1. (B) Mean reaction times for the auditory-vocal task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 1. The error bars represent standard errors of the mean.

vocal responses, which were not clearly slowed in the dual-task conditions. However, the mean RT across blocks for the noncorresponding trials (876 ms) was significantly greater than for all the other trial types ($t_s > 4.83$, $p_s < .0005$).

Combination-specific learning: Auditory-vocal task. As with the data from the visual-manual task, the auditory-vocal RTs were collapsed across Blocks 11 and 12, and the practiced noncorresponding trials (872 ms) were directly compared to the withheld noncorresponding trials (849 ms). The results were the same: There were no significant differences between the RTs, $t(15) < 1$ (see Figure 3B). There were also no significant differences between practiced and withheld combinations for the corresponding (862 ms vs. 713 ms; $t = 1.48$, $p > .15$) and mirror (657 ms vs. 677 ms; $t < 1$) combinations. Thus, there was no evidence of any learning of the specific combinations of stimuli or combinations of responses.

Accuracy. The proportions of correct responses were submitted to identical ANOVAs as the RTs. For the visual-manual task, the two-way ANOVA yielded a significant effect of trial type, $F(3, 45) = 4.18$, $p < .05$, $\eta_p^2 = .22$, but no effect of block, $F(7, 107) = 1.36$, $p = .23$, $\eta_p^2 = .08$, and no interaction, $F(21, 315) = 1.05$, $p = .40$, $\eta_p^2 = .07$. Follow-up t tests revealed that the single and corresponding trials were performed more accurately than the mirror and noncorresponding trials: single versus mirror, $t(15) = 2.01$, $p = .06$, $d = 0.70$; single versus noncorresponding, $t(15) = 2.64$, $p < .05$, $d = 0.97$; corresponding versus mirror, $t(15) = 2.21$, $p < .05$, $d = 0.76$; corresponding versus noncorresponding, $t(15) = 3.14$, $p < .01$, $d = 1.07$. The single and corresponding, $t(15) = 1.74$, $p = .10$, $d = 0.51$, and mirror and noncorresponding ($t < 1$) did not differ from each other.

For the auditory-vocal task, the results of the two-way ANOVA were similar: There was a significant effect of trial type, $F(3, 45) = 9.96$, $p < .001$, $\eta_p^2 = .40$, but no effect of block, $F(7, 107) = 1.17$, $p = .33$, $\eta_p^2 = .07$, and no interaction, $F(21, 315) < 1$. However, the pattern of accuracies averaged across blocks was slightly different. Single (.97) trials were significantly more accurate than corresponding trials (.94), $t(15) = 3.94$, $p < .005$, and noncorresponding trials (.90), $t(15) = 4.52$, $p < .0005$, but did not differ significantly from mirror trials (.96), $t(15) = 1.01$, $p = .33$.

Mirror and corresponding trials did not differ from each other, $t(15) = 1.62$, $p = .13$, but they both differed from noncorresponding trials: corresponding versus noncorresponding, $t(15) = 2.70$, $p < .05$; mirror versus noncorresponding, $t(15) = 3.02$, $p < .01$. In sum, the accuracy data are generally consistent with the RT data; single-task trials were performed best, noncorresponding trials were performed worst, and corresponding and mirror trials were performed at an intermediate level.

There was no evidence of any combination-specific learning in the accuracy rates. For the visual-manual task, the mean accuracies for the practiced combinations (.99) and withheld combinations (.99) were statistically indistinguishable ($t < 1$). The same pattern was observed for the auditory-vocal task, where the accuracies for the practiced (.91) and withheld (.91) combinations did not differ ($t < 1$). This pattern is completely consistent with the conclusions drawn from the RT data.

Summary. The results of Experiment 1 indicate that there was no learning of specific combinations of stimuli or responses across tasks. Although there were significant decrements in RT with training, these improvements in performance appeared to reflect learning related to the individual tasks separately, or simply accommodating to dual-task conditions, rather than the integration of information across tasks. The findings are consistent with those from Hazeltine et al. (2002) and indicate that multiple sessions are not necessary to achieve functional independence between the two simultaneously performed tasks. However, the findings differ from those for chord learning tasks, in which participants are presented with two or more stimuli and required to make a corresponding number of simultaneous manual responses. In chord experiments, RTs were slower by over 100 ms for withheld combinations than for practiced combinations, even under conditions in which all the chords consisted of a single keypress on each hand (Hazeltine et al., 2007).

Experiment 2

One obvious account of the differences between the present findings and those obtained in the chord learning studies is that the stimulus and response modalities were the same for all stimulus–response combinations in the chord tasks but distinct

for the two tasks in Experiment 1. Perhaps for the two tasks to be integrated, they must be sufficiently similar. Thus, when there is no dimensional overlap between the stimuli or responses of the two tasks, learning occurs for them separately. This is consistent with Conway and Christiansen's (2006) modality-specific learning account. It is less consistent with the two-system account proposed by Keele et al. (2003) unless one makes the additional assumption that the attentional demands of making two simultaneous responses prevents the multidimensional system from forming associations between the two tasks. This possibility cannot be ruled out, but it is unclear why this would be, given that both stimuli must be attended because they require a speeded response.

To confirm that combination-specific learning does occur under these task conditions when the modality pairings are more similar for the two tasks, we conducted a second experiment, identical to Experiment 1 in every respect except that the auditory-vocal task was replaced by a visual-manual task for the right hand. If the difference in the modality pairings of the two tasks is responsible for the absence of combination-specific learning in Experiment 1, then such learning should be observed in Experiment 2.

Method

Participants. Fifteen participants were recruited from the undergraduate population at the University of Iowa. The participants performed for credit in an introductory psychology class. All participants agreed to participate in the study in accordance with the guidelines of the university's Committee for the Protection of Human Subjects. Three participants were eliminated from the analysis for having high error rates or high RTs.

Apparatus and stimuli. The apparatus was identical to that of Experiment 1 except that no microphone was used because there were no vocal responses. The visual stimuli for the left-hand task were identical to those of Experiment 1: White Xs against a black background subtending 1° of visual angle and centered 2°, 3.5°, and 5° to the left of a fixation cross. The stimuli for the right-hand task were identical to those for the left-hand task except that they appeared to the right of fixation. Again, place holding lines just below the Xs were displayed throughout the trial. The stimuli for each hand were presented simultaneously and extinguished as soon as any response was made. Left-hand responses were made with the *S*, *D*, and *F* keys with the ring, middle, and index fingers, respectively. Right-hand responses were made with the *J*, *K*, and *L* keys with the index, middle, and ring fingers, respectively. The mapping of the Xs to keys was spatially compatible for both hands. As in Experiment 1, participants were instructed to make both responses at the same time.

The stimulus combinations were withheld exactly as in Experiment 1 (see Figure 1). Note that corresponding pairs now required spatially corresponding responses (e.g., leftmost finger of the left hand and leftmost finger of the right hand), and mirror pairs now required responses with homologous effectors (e.g., index finger of the left hand and index finger of the right hand).

Procedure. All other aspects of the procedure were identical to those in Experiment 1.

Results and Discussion

Reaction times. The data were analyzed as in Experiment 1. Trials with an incorrect response or an RT less than 250 ms or greater than 2,000 ms were removed from the analyses, eliminating 0.5% of the remaining data.

Training performance: Left-hand task. The RTs for the visual-manual task were submitted to a two-way ANOVA with block (3–10) and trial type (single, corresponding, mirror, and noncorresponding) as factors. Both main effects—block, $F(7, 77) = 4.85, p < .001, \eta_p^2 = .31$; trial type, $F(3, 33) = 58.52, p < .001, \eta_p^2 = .84$ —and their interaction, $F(21, 231) = 1.89, p < .05, \eta_p^2 = .15$, were significant. As shown in Figure 4A, RTs decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials. However, the decrements in RT with practice were larger for the noncorresponding trials (256 ms) than for the other trial types (71–140 ms; all $t_s > 2.5, p_s < .05$), as in Experiment 1. The decrements did not differ between any other trial types (all $t_s < 2.08, p_s > .06$).

Post hoc comparisons between the four trial types revealed that each differed from all the others ($t_s > 3.74, p_s < .005$). Single-task trials (537 ms) were the fastest, mirror compatible (643 ms) were the next fastest, corresponding (749 ms) were the next fastest, and noncorresponding (1,043 ms) were the slowest. Thus, the greater dimensional overlap between the two tasks increased the differences among the trial types and created an advantage for the mirror trials over the corresponding trials. The advantage for the mirror trials may stem from the fact that these trials required keypresses with homologous effectors.

Combination-specific learning: Left-hand task. Practiced and withheld noncorresponding trials were directly compared with RTs pooled across Blocks 11 and 12, as in Experiment 1. As depicted in Figure 4B, there was a 161-ms (practiced, 927 ms; withheld, 1,088 ms) difference between the practice and withheld combinations, $t(11) = 3.12, p < .001, d = 0.67$.⁴ The differences between practiced and withheld combinations for the corresponding (699 ms vs. 814 ms; $t = 3.33, p < .05$) and mirror (629 ms vs. 741 ms; $t = 2.76, p < .05$) combinations were also significant. Thus, unlike in Experiment 1, there was clear evidence of combination-specific learning. Overlap in the stimulus and response modalities for the two tasks appears to play a critical role in task integration. That is, when the two tasks used the same modalities, we saw robust combination-specific learning, but when they used distinct modalities, we saw no combination-specific learning. This finding suggests that combination-specific learning is highly dependent on the relationship between the two tasks.

Training performance: Right-hand task. Given the strong similarity between the two tasks, it is not surprising that the pattern of right-hand RTs is nearly identical to that of the left-hand RTs. Both main effects—block, $F(7, 77) = 4.47, p < .001, \eta_p^2 = .29$;

⁴ We also examined whether learning-related differences emerged for the withheld combinations for Blocks 11 and 12. The t tests revealed no significant differences for the withheld RTs across the two blocks for both tasks ($t_s < 1$). We performed identical analyses for the other experiments showing combinations-specific learning (Experiments 5 and 6) and again found no significant differences (all $t_s < 1$). Thus, learning of withheld combinations from Block 11 to 12 appears to be minimal for all the experiments showing combination-specific learning.

Figure 4. (A) Mean reaction times for the left-hand task across the eight practice blocks for the four trial types in Experiment 2. (B) Mean reaction times for the left-hand task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 2. The error bars represent standard errors of the mean.

trial type, $F(3, 33) = 55.58$, $p < .001$, $\eta_p^2 = .84$ —as well as the interaction, $F(21, 231) = 2.42$, $p < .005$, $\eta_p^2 = .18$, were significant. As shown in Figure 5A, RTs generally decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials. As with the left-hand task, the decrements in RT with practice were larger for the noncorresponding trials (259 ms) than for the other trial types (29–154 ms; all $t_s > 2.2$, $p_s < .05$). Moreover, the decrement for the mirror trials (154 ms) was significantly greater than the decrement for single-task trials (29 ms) trials, $t(11) = 3.51$, $p < .01$.

Post hoc comparisons between the four trial types revealed that each differed from all the others ($t_s > 3.74$, $p_s < .005$). Single-task trials (513 ms) were the fastest, mirror compatible (648 ms) were the next fastest, corresponding (756 ms) were the next fastest, and noncorresponding (1,048 ms) were the slowest. The pattern is nearly identical to that observed for the left hand, which is not surprising given that participants were instructed to make their responses simultaneously.

Combination-specific learning: Right-hand task. The practiced (929 ms) and withheld (1,094 ms) trials differed by 165 ms, $t(11) = 2.81$, $p < .05$, $d = 0.67$, indicating that combination-specific learning had occurred. The differences between practiced and withheld combinations for the corresponding (705 ms vs. 813 ms; $t = 3.23$, $p < .05$) and mirror (635 ms vs. 736 ms; $t = 2.47$, $p < .05$) combinations were also significant. These findings support the conclusions drawn from left-hand task.

Accuracy. The proportions of correct responses were submitted to identical ANOVAs as the RTs. For the training data from the left-hand task, there was a significant effect of trial type, $F(3, 33) = 8.49$, $p < .001$, but no effect of block, $F(7, 77) < 1$, or interaction, $F(21, 231) = 1.46$, $p = .09$. For the right-hand task, there was a significant effect of trial type, $F(3, 33) = 17.08$, $p < .001$, and block, $F(7, 77) = 2.22$, $p < .05$, but no interaction, $F(21, 231) = 1.45$, $p = .10$. For both tasks, accuracy was lowest for the noncorresponding trials (.96–.97) and very high in the remaining conditions (.99–1.00).

Figure 5. (A) Mean reaction times for the right-hand task across the eight practice blocks for the four trial types in Experiment 2. (B) Mean reaction times for the right-hand task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 2. The error bars represent standard errors of the mean.

For the comparison of the practice and withheld combinations on Blocks 11 and 12, the mean accuracies were lower for the withheld combinations for both the left hand (.92 vs. .98) and right hand (.95 vs. .98). The difference was significant for the left hand, $t(11) = 3.07, p < .05$, but not the right hand, $t(11) = 1.42, p = .18$. In sum, the accuracies were consistent with the RTs.

Summary. Experiments 1 and 2 indicate that the stimulus and response modalities affect whether combination-specific learning occurs. Although the decrements in RT over practice were observed in both experiments, combination-specific learning was observed only in Experiment 2, which used the same procedure as Experiment 1, except that both tasks were visual-manual rather than visual-manual and auditory-vocal. This suggests that stimulus modality and/or response modality plays a critical role in determining whether associations are encoded between simultaneously performed stimulus-response pairs for the two tasks. Modality may play a role because learning mechanisms are modality specific or because modality provides a strong cue to top-down processes about the relatedness of the two tasks and these top-down processes in turn influence learning. Therefore, in Experiments 3 and 4, we examine whether stimulus or response modality affects learning, and in Experiments 5 and 6, we determine whether modality provides a hard constraint because learning processes are modality specific or rather simply acts as a salient cue to top-down processes.

Experiment 3

An obvious question emerging from Experiments 1 and 2 is whether overlap in the stimulus modalities or overlap in the response modalities is sufficient, on its own, to lead to task integration. We address this question directly in Experiments 3 and 4, in which we use separate stimulus modalities but a shared response modality (Experiment 3) or a shared stimulus modality but separate response modalities (Experiment 4).

Method

Participants. Twenty-seven participants were recruited from the undergraduate population at the University of Iowa. The par-

ticipants performed for credit in an introductory psychology class. All participants agreed to participate in the study in accordance with the guidelines of the university's Committee for the Protection of Human Subjects. Three participants were eliminated from the analysis for having high error rates or high RTs.

Apparatus and stimuli. The apparatus was identical to that of Experiment 1 except that no microphone was used because there were no vocal responses. The stimuli were identical to those used in Experiment 1. The responses were identical to those used in Experiment 2. Thus, participants responded to the low, middle, and high tones, with the index, middle, and ring fingers of their right hands. The stimulus combinations were withheld exactly as in Experiments 1 and 2.

Procedure. All other aspects of the procedure were identical to those in Experiments 1 and 2.

Results and Discussion

Reaction times. The data were analyzed as in the previous experiments. Trials with an incorrect response or an RT less than 250 ms or greater than 2,000 ms were removed from the analyses, eliminating 7% of the remaining data. The mean RTs across blocks are depicted in Figures 6A and 7A.

Training performance: Left-hand task. Both main effects—block, $F(7, 161) = 9.53, p < .001, \eta_p^2 = .29$; trial type, $F(3, 69) = 118.54, p < .001, \eta_p^2 = .84$ —were significant, but their interaction was not, $F(21, 483) = 1.10, p = .35$. As shown in Figure 6A, RTs generally decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials. Post hoc comparisons between the four trial types revealed that each differed from all the others ($ts > 4.66, ps < .001$). Single-task trials (713 ms) were the fastest, mirror compatible (974 ms) were the next fastest, corresponding (1,052 ms) were the next fastest, and noncorresponding (1,211 ms) were the slowest. Although the overall RTs are slower, the pattern is similar to that of Experiment 2.

Combination-specific learning: Left-hand task. In contrast to Experiment 2, which also required participants to make two simultaneous manual responses, there was little evidence of a difference between the practiced (1,130 ms) and withheld (1,147 ms) combinations, $t(23) < 1$ (see Figure 6B). There were also no

Figure 6. (A) Mean reaction times for the left-hand task across the eight practice blocks for the four trial types in Experiment 3. (B) Mean reaction times for the left-hand task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 3. The error bars represent standard errors of the mean.

Figure 7. (A) Mean reaction times for the right-hand task across the eight practice blocks for the four trial types in Experiment 3. (B) Mean reaction times for the right-hand task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 3. The error bars represent standard errors of the mean.

significant differences between practiced and withheld combinations for the corresponding (965 ms vs. 1,069 ms; $t = 1.55$, $p > .14$) and mirror (1,016 ms vs. 956 ms; $t < 1$) combinations. Thus, it appears that the overlap in the response modality for the two tasks is not sufficient to produce combination-specific learning. Even when the two tasks require simultaneous, manual responses, combination-specific learning does not necessarily take place.

Training performance: Right-hand task. The patterns of right-hand RTs were nearly identical to those of the left-hand RTs. Both main effects—block, $F(7, 161) = 10.85$, $p < .001$, $\eta_p^2 = .32$; trial type, $F(3, 69) = 79.30$, $p < .001$, $\eta_p^2 = .78$ —as well as the interaction, $F(21, 483) = 1.87$, $p < .05$, $\eta_p^2 = .08$, were significant. As shown in Figure 7A, RTs generally decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials. As with the left-hand task, each differed from all the others ($t_s > 4.46$, $p_s < .001$). Single-task trials (919 ms) were the fastest, mirror (991 ms) were the next fastest, corresponding (1,078 ms) were the next fastest, and noncorresponding (1,244 ms) were the slowest. Again, the interaction reflected larger improvements across blocks for slower trial types: single-task, 80 ms; mirror, 150 ms; corresponding, 205 ms; and noncorresponding, 284 ms.

Combination-specific learning: Right-hand task. The practiced (1,162 ms) and withheld (1,187 ms) trials differed by just 25 ms, $t(23) < 1$, indicating that little combination-specific learning had occurred (see Figure 7B). There were also no significant differences between practiced and withheld combinations for the corresponding (999 ms vs. 1,041 ms; $t = 1.38$, $p > .19$) and mirror (962 ms vs. 958 ms; $t < 1$) combinations. These findings support the conclusions drawn from left-hand task.

Accuracy. The proportions of correct responses were submitted to identical ANOVAs as the RTs. For the training data from the left-hand task, there was a significant effect of trial type, $F(3, 69) = 11.65$, $p < .001$, but no effect of block, $F(7, 161) < 1$, or interaction, $F(21, 483) < 1$. The same pattern was observed for the right-hand task: trial type, $F(3, 69) = 40.79$, $p < .001$; block, $F(7, 161) = 1.30$, $p = .26$; interaction, $F(21, 483) = 1.29$, $p = .18$. For both tasks, accuracy was lowest for the noncorresponding trials

(left hand, .90; right hand, .98) and very high in the remaining conditions (.99–1.00).

For the comparison of the practiced and withheld combinations on Blocks 11 and 12, the mean accuracies were lower for the withheld combinations for both the left hand (.98 vs. .97) and right hand (.93 vs. .88). The difference was significant for the right hand, $t(11) = 2.11$, $p < .05$, but not the left hand, $t(11) < 1$. Thus, the 5% difference between the practiced and withheld combinations suggests that there may be some encoding of specific combinations, but it is qualitatively different from the learning observed in Experiment 2, in which robust RT costs were observed.

Summary. In Experiment 3, we investigated whether a common response modality with distinct stimulus modalities was sufficient to produce combination-specific learning. No combination-specific learning was observed, indicating that simultaneously produced combinations of manual responses are not always learned. This finding suggests an important constraint on motor learning not captured by existing theories of motor learning, including those that have been applied to the chord task (e.g., Logan, 1988; Rosenbloom & Newell, 1988). It is not sufficient for motor responses to be produced at the same time for them to be encoded as a single chunk. It appears that the responses have to be linked prior to the execution of a motor command for combination-specific learning to occur; that is, task integration occurs before the programming of the motor command.

Experiment 4

In Experiment 4, we tested whether overlap in the stimuli are sufficient to produce response integration. Thus, we used the same stimuli as in Experiment 2, in which combination-specific learning was observed, but the same responses as in Experiment 1, in which combination-specific learning was not observed. In this way, Experiment 4 fills the fourth cell in a 2×2 design in which the secondary task stimulus modality (auditory or visual) and response

modality (vocal or manual) are manipulated to promote or discourage intertask relatedness.

Method

Participants. Seventeen participants were recruited from the undergraduate population at the University of Iowa. The participants performed for credit in an introductory psychology class. All participants agreed to participate in the study in accordance with the guidelines of the university's Committee for the Protection of Human Subjects. One participant was eliminated from the analysis for having high error rates.

Apparatus and stimuli. The apparatus was identical to that of Experiment 1 except that no auditory stimuli were presented. Instead, the stimuli were identical to those used in Experiment 2: two rows of three dashes to the left and right of fixation, with one of the dashes replaced by an X. The responses were identical to those used in Experiment 1. Thus, participants responded to the Xs to the right of fixation by saying the words "bug," "food," and "tree." The stimulus combinations were withheld exactly as in Experiments 1–3.

Procedure. All other aspects of the procedure were identical to those in Experiments 1–3.

Results and Discussion

Reaction times. The data were analyzed as in the previous experiments. Trials with an incorrect response or an RT less than 250 ms or greater than 2,000 ms were removed from the analyses, eliminating 7% of the remaining data.

Training performance: Left-hand task. Both main effects—block, $F(7, 105) = 9.36, p < .0001, \eta_p^2 = .38$; trial type, $F(3, 45) = 65.67, p < .0001, \eta_p^2 = .81$ —were significant, along with their interaction, $F(21, 315) = 2.34, p < .005, \eta_p^2 = .14$. As shown in Figure 8A, RTs generally decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials. Post hoc comparisons between the four trial types revealed that each differed from all the others ($t_s > 4.66, p_s < .001$). Single-task trials (657 ms) were the fastest, mirror compatible (1,119 ms) were the next fastest, corresponding (1,123 ms)

were the next fastest, and noncorresponding (1,259 ms) were the slowest. The interaction reflected smaller improvements across blocks for the single-task trials: single-task, 106 ms; mirror, 207 ms; corresponding, 169 ms; and noncorresponding, 207 ms.

Combination-specific learning: Left-hand task. In contrast to Experiment 2, which also used two simultaneous, visually presented spatial stimuli, there was little evidence of a difference between the practiced (1,206 ms) and withheld (1,218 ms) combinations, $t(17) < 1$ (see Figure 8B). There were also no significant differences between practiced and withheld combinations for the corresponding (1,073 ms vs. 1,102 ms; $t < 1$) and mirror (1,141 ms vs. 1,113 ms; $t < 1$) combinations.

Training performance: Vocal task. Both main effects were significant—block, $F(7, 105) = 16.29, p < .0001, \eta_p^2 = .52$; trial type, $F(3, 45) = 22.75, p < .0001, \eta_p^2 = .60$ —as well as the interaction, $F(21, 315) = 2.09, p < .005, \eta_p^2 = .12$. As shown in Figure 9A, RTs decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials. Single-task trials (1,014 ms) were the fastest, mirror (1,140 ms) were the next fastest, corresponding (1,145 ms) were the next fastest, and noncorresponding (1,268 ms) were the slowest. Again, the interaction reflected the smallest improvements across blocks for single-task trials: single-task, 114 ms; mirror, 296 ms; corresponding, 240 ms; and noncorresponding, 211 ms.

Combination-specific learning: Vocal task. The vocal RTs for practiced trials (1,140 ms) were similar to those for the withheld trials (1,141 ms), $t(15) < 1$, indicating that little combination-specific learning had occurred (see Figure 9B). There were also no significant differences between practiced and withheld combinations for the corresponding (1,220 ms vs. 1,253 ms; $t < 1$) and mirror (1,319 ms vs. 1,303 ms; $t < 1$) combinations. These findings support the conclusions drawn from the left-hand task.

Accuracy. The proportions of correct responses were submitted to identical ANOVAs as the RTs. For the training data from the left-hand task, there was a significant effect of trial type, $F(3, 45) = 6.68, p < .001$, but no effect of block, $F(7, 105) = 1.18, p = .32$, or interaction, $F(21, 315) < 1$. The participants were highly accurate for all trial types, but slightly less accurate for the single-task (.988) trials than for the other trial types (mirror, .998;

Figure 8. (A) Mean reaction times for the left-hand task across the eight practice blocks for the four trial types in Experiment 4. (B) Mean reaction times for the left-hand task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 4. The error bars represent standard errors of the mean.

Figure 9. (A) Mean reaction times for the vocal task across the eight practice blocks for the four trial types in Experiment 4. (B) Mean reaction times for the vocal task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 4. The error bars represent standard errors of the mean.

corresponding, .996; noncorresponding, .998). For the vocal task, the mean accuracy was .981, and there were no main effects or interactions (all F s < 1). For the comparison of the practice and withheld combinations on Blocks 11 and 12, the mean accuracies were similar for the withheld combinations for both the left hand (.98 vs. .98; $t < 1$) and vocal responses (.99 vs. .97), $t(15) = 1.78$, $p = .10$.

Summary. In Experiment 4, participants selected two responses based on two sets of visual stimuli. Despite the overlap in stimulus modality, there was no combination-specific learning, indicating that neither stimulus overlap (Experiment 4) nor response overlap (Experiment 3) on its own is sufficient to allow the encoding of associations across the two tasks. Thus, the findings from Experiments 1–4 suggest a major limitation in the ability to learn contingencies between stimulus–response associations across tasks. Intertask combinations appear to be learned only when there is overlap between the stimulus and response modalities associated with the two tasks; it is not simply enough that the stimuli overlap in terms of dimension for task integration to occur. Because both the input and output modalities play a critical role, the locus of learning must involve central processes that link sensory modalities with motor modalities.

However, the means by which input and output modalities constrain learning is unclear. It is possible that this form of learning takes place within modality-specific processes, such that associations can only be formed between stimulus–response pairs sharing the same stimulus and response modalities. If the stimulus and response modalities are broadly defined, the number of combinations of stimulus and response modalities may not be overly large to allow for a reasonable number of learning mechanisms to capture the vast array of human behaviors. This hypothesis is consistent with the Conway and Christiansen (2006) account, with the added constraint that the responses also play a role in determining the boundaries of the associative learning processes. Note that in the Conway and Christiansen experiments, no overt responses were made to the stimuli during training, as in many statistical learning tasks (e.g., Saffran, Johnson, Aslin, & Newport, 1999).

An alternative account is that it is not the relationship of the modalities of the tasks but rather the overall relatedness of the two

tasks that determines whether combination-specific learning occurs. That is, simultaneously produced responses may be associated when they are conceptualized as part of a common task. In Experiment 2, not only were both stimuli visual and both responses manual, but the stimulus–response mappings for the two tasks were closely related. Indeed the dual-task condition could be thought of as a single chord-like task. In the other experiments, the tasks were less clearly related, so they were likely conceptualized as separate. This may have influenced learning processes through top-down mechanisms along the lines described by Melchers et al. (2008). These two alternative accounts are tested in Experiments 5 and 6.

One issue to note is that in the experiments that did not show combination-specific learning, the single-task RTs for the two tasks differed considerably (Experiment 1, 565 ms vs. 793 ms; Experiment 3, 713 ms vs. 919 ms; Experiment 4, 657 ms vs. 1,014 ms). In contrast, in the experiment that did show combination-specific learning, the single-task RTs were similar for the two tasks (Experiment 2, 537 ms vs. 513 ms). Thus, it is possible that the similar RTs facilitate combination-specific learning. On dual-task trials, participants were instructed to produce both responses at the same time, so there is little difference in RT on these trials; that is, the differences in RT are mostly confined to the single-task conditions. However, it is possible that when one response requires more time to select than another, individuals engage in a different strategy that makes combination-specific learning less likely. For example, response selection might occur in a more serial fashion when the two tasks differ in duration, with the first response being withheld until the second is selected. This approach might reduce combination-specific learning, because selection occurs independently and control processes imposing the serial order of the task operations limit the interactions between the two tasks (e.g., Meyer & Kieras, 1997; see also Schumacher & Schwarb, 2009). Thus, the role of single-task RT differences in combination-specific learning remains unclear. We return to this issue in Experiments 5 and 6.

Experiment 5

In Experiment 5, we investigate whether the effects observed in Experiments 1–4 reflect “hard” boundaries such that associations between distinct modalities cannot be formed under the conditions

imposed by chord learning tasks. This would be consistent with theories proposing that response selection is performed by separate modality-specific modules (e.g., Cohen & Shoup, 1997; Goschke & Bolte, 2012; Magen & Cohen, 2007). An alternative account holds that the boundaries are “soft” and do not reflect limitations that are determined strictly by the modalities. Rather, combination-specific learning is contingent on whether the two tasks are sufficiently related such that they can be conceptualized as a single task.

Here we manipulate the relatedness of the two tasks by increasing dimensional overlap among the stimuli while keeping the modality separate. If top-down processes can overcome the boundaries of crossmodal learning, combination-specific learning should be observed by increasing the relatedness of the two tasks being performed. In contrast, if manipulating the relatedness of the two tasks does not induce combination-specific learning, then it may be the case that modality represents a stronger force in defining the rules associative learning.

To test this, we used two manual tasks with stimuli from distinct modalities as in Experiment 3. However, they were all numbers: Visually presented numbers 1–3 indicated the left-hand responses, and aurally presented numbers 4–6 indicated the right-hand responses. In this way, the two tasks could be conceptualized as a single task, with the responses for both hands falling on a single number line.

Method

Participants. Thirteen participants were recruited from the undergraduate population at the University of Iowa. The participants performed for credit in an introductory psychology class. All participants agreed to participate in the study in accordance with the guidelines of the university’s Committee for the Protection of Human Subjects. One participant was eliminated from the analysis for having high error rates.

Apparatus and stimuli. Visually presented numbers 1, 2, and 3, subtending 1° of visual angle vertically and 0.6° horizontal, were presented 2.4° to the left of the fixation cross. The auditory numbers were presented by playing .wav files, created with the AT

& T text-to-speech website (<http://www.research.att.com/~ttsweb/tts/demo.php>) and modified with Praat software (<http://www.praat.org>). The .wav files generated by the website were trimmed and shortened so that each number word lasted exactly 250 ms. Participants were able to adjust the volume of the auditory stimuli to a comfortable level.

Procedure. All other aspects of the procedure were identical to those in Experiment 3, which also required two manual responses.

Results and Discussion

Reaction times. The data were analyzed as in the previous experiments. Trials with an incorrect response or an RT less than 250 ms or greater than 2,000 ms were removed from the analyses, eliminating 5% of the remaining data. The mean RTs across blocks are depicted in Figures 10A and 11A.

Training performance: Left-hand task. The effect of trial type was significant, $F(3, 33) = 37.31, p < .0001, \eta_p^2 = .77$, but the main effect of block, $F(7, 77) = 1.68, p = .13$, and its interaction with trial type, $F(21, 231) < 1$, was not. As shown in Figure 10A, RTs generally decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials. Post hoc comparisons between the four trial types revealed that each type differed from all the others ($t_s > 3.43, p_s < .005$), except for the corresponding and mirror trials ($t < 1$). Single-task trials (780 ms) were the fastest, corresponding (948 ms) were the next fastest, mirror (972 ms) were the next fastest, and noncorresponding (1,136 ms) were the slowest.

Combination-specific learning: Left-hand task. In contrast to Experiment 3, in which participants also performed a visual-manual and an auditory-manual task, there was clear evidence of learning (see Figure 10B). The withheld combinations (1,240 ms) were performed significantly more slowly than the practiced combinations (1,120 ms), $t(11) = 2.36, p < .05, d = 1.00$. The differences between practiced and withheld combinations for the corresponding (929 ms vs. 1,083 ms; $t = 3.43, p < .05$) and mirror (859 ms vs. 1,060 ms; $t = 5.13, p < .005$) combinations were also significant. Despite the fact that distinct stimulus modalities were

Figure 10. (A) Mean reaction times for the left-hand task across the eight practice blocks for the four trial types in Experiment 5. (B) Mean reaction times for the left-hand task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 5. The error bars represent standard errors of the mean.

Figure 11. (A) Mean reaction times for the right-hand task across the eight practice blocks for the four trial types in Experiment 5. (B) Mean reaction times for the right-hand task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 5. The error bars represent standard errors of the mean.

used, the responses for the two tasks were conceptually related, and this in turn appears to allow for the integration of responses, as indicated by the cost for new combinations.

Training performance: Right-hand task. Both main effects—block, $F(7, 77) = 3.52, p < .005, \eta_p^2 = .24$; trial type, $F(3, 33) = 52.46, p < .0001, \eta_p^2 = .83$ —were significant, but not the interaction, $F(21, 231) < 1$. As shown in Figure 11A, RTs generally decreased with practice. As with the left-hand task, each trial type differed from all the others ($t_s > 4.00, p_s < .005$), except the corresponding and mirror trial types ($t < 1$). Single-task trials (778 ms) were the fastest, corresponding (980 ms) were the next fastest, mirror (997 ms) were the next fastest, and noncorresponding (1,186 ms) were the slowest.

Combination-specific learning: Right-hand task. The right hand, like the left hand, showed clear evidence of combination-specific learning (see Figure 11B). The practiced noncorresponding combinations (1,149 ms) were performed significantly faster than the withheld noncorresponding combinations (1,275 ms), $t(11) = 2.99, p < .05, d = 0.92$. The differences between practiced and withheld combinations for the corresponding (943 ms vs. 1,082 ms; $t = 3.04, p < .05$) and mirror (862 ms vs. 1,119 ms; $t = 4.48, p < .01$) combinations were also significant. These findings support the conclusions drawn from left-hand task.

The single-task RTs for the two tasks were much more similar to each other (left hand, 778 ms; right hand, 780 ms) than the single-task RTs in the experiments in which no learning was observed. Thus, similar single-task RTs may be a factor contributing to the integration of stimulus–response combinations.

Accuracy. For the training data from the left-hand task, there was a significant effect of trial type, $F(3, 33) = 7.99, p < .001$, but no effect of block, $F(7, 77) = 1.36, p = .24$, or interaction, $F(21, 231) < 1$. The same pattern was observed for the right-hand task: an effect of trial type, $F(3, 33) = 10.81, p < .001$, but not of block, $F(7, 77) = 1.69, p = .13$, and no interaction ($F < 1$). For both tasks, accuracy was lowest for the noncorresponding trials (left hand, .97; right hand, .93) and very high in the remaining conditions (.98–.99).

For the comparison of the practice and withheld combinations on Blocks 11 and 12, the accuracies were lower for the

withheld combinations for both the left hand (.98 vs. .93) and right hand (.96 vs. .92). This difference was marginally significant for the left hand, $t(11) = 2.21, p = .07$, but not for the right hand ($t < 1$). In sum, the accuracies were consistent with the RTs.

Summary. In Experiment 5, we pitted the overlapping-modality hypothesis against the related-task hypothesis. Robust combination-specific learning was observed, similar in magnitude to what was observed in Experiment 2, indicating that overlapping stimulus modalities are not necessary for associations to be formed across the two tasks. The pattern observed in previous experiments does not appear to reflect neuroanatomical connections between sensory and motoric regions. Instead, combination-specific learning appears to depend on the relationship between the two tasks. That is, the boundaries of combination-specific learning appear to be soft rather than hard, at least with regard to the stimulus boundaries. Intriguingly, the single-task RTs were similar (visual task, 780 ms; auditory task, 778 ms), as in the other experiment (Experiment 2) producing robust learning.

Experiment 6

Experiment 5 demonstrated that when the two tasks are conceptually related, combination-specific learning is observed, even when the stimuli were presented along distinct modalities. In Experiment 6, we asked whether combination-specific learning can be observed for conceptually related tasks even when the responses are made along indistinct modalities. If combination-specific learning is observed under these experimental conditions, it would suggest that top-down processes can determine the types of contingencies that the brain is able to form. However, if combination-specific learning is not observed, this may suggest that the types of contingencies the brain can encode are limited to the constraints of response modality.

To test this, we paired a visual-vocal task with a visual-manual task, where responses were executed between distinct modalities as in Experiment 4. Contrary to Experiment 4, the stimuli in Experiment 6 were all numbers: Visually presented numbers 1, 3, and 5 indicated the left-hand responses, and visually presented

numbers 2, 4, and 6 indicated the vocal responses. The numbers were interleaved to encourage participants to think of them as belonging to the same task even though they required responses along distinct modalities.

Method

Participants. Twenty-two participants were recruited from the undergraduate population at the University of Iowa. The participants performed for credit in an introductory psychology class. All participants agreed to participate in the study in accordance with the guidelines of the university's Committee for the Protection of Human Subjects. Six participants were eliminated from the analysis because of malfunctions recording auditory responses.

Apparatus and stimuli. The apparatus was identical to that of Experiment 4 except that the dashes to the left and right of the fixation cross were replaced by numbers. Visually presented numbers 1, 3, and 5, subtending 1° of visual angle vertically and 0.6° horizontal, were presented 2.4° to the left of the fixation cross. Likewise, visually presented numbers 2, 4, and 6, subtending 1° of visual angle vertically and 0.6° horizontal, were presented 2.4° to the right of the fixation cross. Participants responded to the numbers on the left as in Experiment 5. Participants responded to the numbers 2, 4, and 6 on the right, with the vocal response "bug," "food," or "tree," respectively. Responses were interleaved to encourage the conceptualization of both tasks as one.

Procedure. All other aspects of the procedure were identical to those in Experiment 4, which also required one manual and one vocal response.

Results and Discussion

Reaction times. The data were analyzed as in previous experiments. Trials with an incorrect response or an RT less than 250 ms or greater than 2,000 ms were removed from the analyses, eliminating 8.9% of the remaining data. The mean RTs across blocks are depicted in Figures 12A and 13A.

Training performance: Left-hand task. The effect of trial type, $F(3, 33) = 82.71$, $p < .001$, $\eta_p^2 = .87$, and block, $F(7, 77) =$

12.05 , $p < .001$, $\eta_p^2 = .50$, were significant, but not the interaction, $F(21, 231) < 1$. As shown in Figure 12A, RTs decreased with practice and were shorter for the single-task trials and longest for the noncorresponding trials. Post hoc comparisons between the four trial types revealed that each type differed from all the others ($ts > 3.49$, $ps < .001$), except for the noncorresponding and mirror trials ($t < 1$). Single-task trials (777 ms) were the fastest, corresponding (1,001 ms) were the next fastest, mirror (1,129 ms) were the next fastest, and noncorresponding (1,163 ms) were the slowest.

Combination-specific learning: Left-hand task. In contrast to Experiment 4, in which participants also performed a visual-manual and a visual-vocal task, there was clear evidence of combination-specific learning (see Figure 12B). The withheld combinations (1,263 ms) were performed significantly more slowly than the practiced combinations (1,137 ms), $t(15) = 2.902$, $p < .05$, $d = 0.52$. The difference between practiced and withheld combinations for the corresponding combinations was significant (894 ms vs. 1,174 ms; $t = 3.72$, $p < .005$), but the difference for the mirror combinations was not (980 ms vs. 1,114 ms; $t = 1.52$, $p = .15$), although the trend was in the same direction as the other combinations. Despite the fact that distinct response modalities were used, the stimuli for the two tasks were conceptually related, and this in turn appears to allow for the integration of responses, as indicated by the cost for new combinations.

Training performance: Vocal task. Both main effects—block, $F(7, 77) = 4.99$, $p < .001$, $\eta_p^2 = .294$; trial type, $F(3, 33) = 23.78$, $p < .001$, $\eta_p^2 = .665$ —were significant, but not the interaction, $F(21, 231) = 1.343$. As shown in Figure 13A, RTs generally decreased with practice. As with the left-hand task, each trial type differed from all the others ($ts > 3.02$, $ps < .01$), except the noncorresponding and mirror trial types ($t < 1$). Single-task trials (795 ms) were the fastest, corresponding (849 ms) were the next fastest, mirror (944 ms) were the next fastest, and noncorresponding (963 ms) were the slowest.

Combination-specific learning: Vocal task. The vocal task, like the left-hand task, showed clear evidence of combination-specific learning (see Figure 13B). The practiced noncorresponding

Figure 12. (A) Mean reaction times for the left-hand task across the eight practice blocks for the four trial types in Experiment 6. (B) Mean reaction times for the left-hand task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 6. The error bars represent standard errors of the mean.

Figure 13. (A) Mean reaction times for the right-hand task across the eight practice blocks for the four trial types in Experiment 6. (B) Mean reaction times for the right-hand task on the noncorresponding withheld and practiced combinations in Blocks 11–12 of Experiment 6. The error bars represent standard errors of the mean.

combinations (922 ms) were performed significantly faster than the withheld noncorresponding combinations (1,052 ms), $t(15) = 2.83$, $p < .02$, $d = .50$. The difference between practiced and withheld combinations for the corresponding combinations was significant (766 ms vs. 1,014 ms; $t = 3.37$, $p < .005$), but the difference for the mirror combinations was not (917 ms vs. 1,051 ms; $t = 1.67$, $p = .12$), although the trend was in the same direction as the other combinations. These findings support the conclusions drawn from left-hand task.

Accuracy. For the training data from the left-hand task, there was a significant effect of block, $F(7, 77) = 17.40$, $p < .001$, $\eta_p^2 = .537$, and trial type, $F(3, 33) = 2.128$, $p < .05$, $\eta_p^2 = .124$, but not the interaction, $F(21, 231) = 1.070$. For the right-hand task, there was an effect of trial type, $F(3, 33) = 4.04$, $p < .05$, $\eta_p^2 = .212$, but not block, $F(7, 77) < 1$, or the interaction ($F < 1$). For both tasks, accuracy was lowest for noncorresponding trials for the left hand task (93%) and mirror trials (86%) for the right hand task. Accuracy was between 97% and 99% for the left hand task, and between 87% and 92% for the right hand in the remaining conditions.

For the comparison of the practice and withheld combinations on Blocks 11 and 12, the mean accuracies were lower for the withheld combinations for both the left hand (.94 vs. .89) and vocal task (.86 vs. .82). This difference was not significant for the left hand, $t(15) = 1.918$, or the vocal task ($t = 1.375$).

Summary. Experiment 6 demonstrated combination-specific learning between tasks with nonoverlapping response modalities. Taken together the results of Experiments 5 and 6 demonstrate that the boundaries of associative learning are not constrained by differences in stimulus or response modality. Rather it appears that the conceptual relatedness of the tasks plays a major role in how the task is learned. This provides further evidence of flexible associative learning in humans. The single-task RTs for the two tasks were similar to each other (left hand, 777 ms; right hand, 795 ms) as in the other experiments showing combination-specific learning and unlike the experiments where no learning was observed.

General Discussion

To optimally guide behavior, learning processes need to capture invariance in the task while being flexible enough to apply across a range of situations that are nonidentical to what has previously been experienced. Therefore, it is necessary to distinguish between contingencies that are inherent to the task at hand and those that emerge because of the particular setting in which the task is being learned. In Experiments 1, 3, and 4, associations between simultaneous events were not encoded despite the fact that they were task relevant, suggesting that contingencies are not encoded automatically. Instead, learning depended on the relationship between the two sources of information. Such findings are particularly notable given the wealth of studies demonstrating the remarkable sensitivity of the human brain to contingencies (e.g., Chun & Jiang, 1998; Jiménez & Méndez, 1999; Miller, 1987; Nissen & Bullemer, 1987; Schneider & Shiffrin, 1977; Ziessler, 1998). Whereas combination-specific learning can be viewed as reflecting the brain's ability to flexibly encode associations across stimuli and effectors to facilitate the performance of complex tasks, the absence of combination-specific learning can be viewed as reflecting the brain's ability to shield learning processes from spurious associations that might disrupt performance when the task is required in other contexts (see Dreisbach & Haider, 2009; Halvorson, Wagschal, & Hazeltine, 2013).

This is not to say that combination-specific learning necessarily indicates strengthened stimulus–response associations. It is possible that the learning scores reflect the participants becoming aware of the novel stimulus–stimulus combinations and their performance being slowed by an orienting response to a surprising stimulus (e.g., Knight, 1996). Previous work has attempted to rule out this possibility by demonstrating that there is little awareness of the practiced versus withheld combinations as assessed by postexperimental questionnaires and that the magnitude of the learning score is unrelated to awareness (Hazeltine et al., 2007). Indeed, in the present study, awareness of the withheld combinations was very uncommon across all experiments, as indicated by

informal interview. However, if the learning score is driven by an orienting response, it remains to be explained why its effects would be moderated by our modality and conceptual-relatedness manipulations.

Learning and Attention

It is difficult to formulate a principled explanation for why attention was required for both tasks when combination-specific learning was observed (Experiments 2, 5, and 6) but not required for both tasks when it was not observed (Experiments 1, 3, and 4). Thus, the present findings indicate that attending to both stimuli is insufficient on its own to cause learning of cross-modal associations.⁵ Given that the left-hand task was the same in Experiments 1–4, it is reasonable to ask whether the right-hand task demands more attention in the experiment that produced combination-specific learning (Experiment 2) than the tasks accompanying the left-hand task in those that showed no combination-specific learning (Experiments 1, 3, and 4). The right-hand task in Experiment 2 was a spatially compatible visual-manual task with numerically shorter RTs than the right-hand/vocal tasks in the other experiments, so it is unlikely that it would demand more attentional resources than these other tasks. Thus, we conclude that greater attentional requirements in the right-hand tasks are not responsible for cross-task integration.

Instead, we conclude that top-down control, based on the conceptualization of the tasks, determines which associations are encoded. This conclusion is further supported by Experiments 5 and 6. The exact boundary conditions of “conceptual-relatedness” may be difficult to establish and may depend on levels of practice and individual differences. The critical point is that top-down processes guide the associations that can be formed, even when all the stimuli require responses and thus, presumably, attention. This flexibility is ideal given the diversity of the tasks humans perform—there should be no modality-based limits to the kinds of tasks we learn. As many theorists have noted, allowing for this tremendous range of associations leads to the binding problem. Although the mechanism theorists typically invoke to resolve the binding problem is selective attention (Cohen, Ivry, & Keele, 1990; Kahneman & Treisman, 1984; Keele et al., 2003; Logan, Taylor, & Etherton, 1999; Rosenbloom & Newell, 1988), the present data suggest that simply attending to a stimulus is insufficient for driving the encoding of associations on its own. Top-down processes appear to exert control so that all attended information is not encoded equally (e.g., shielding; Dreisbach & Haider, 2009). Such an approach is optimal given that the information demanding attention may not always be related to the current task.

Because the stimuli in all the experiments required responses and thus should have received attention, we considered other task properties that may have led to integration. In Table 1, we list the single-task and noncorresponding dual-task RTs from the six experiments to determine whether mean RT may have played a role. The experiments that produced robust combination-specific learning (Experiments 2, 5, and 6) used tasks that had similar single-task RTs, whereas the experiments that did not produce significant combination-specific learning (Experiments 1, 3, and 4) involved tasks with dissimilar single-task RTs.⁶ It is possible that the more similar single-task RTs facilitated combination-specific learning, but it is also possible that participants treating the sets of stimulus–response associ-

ations as belonging to a common task led to the more similar RTs. Note that the left-hand task is identical in Experiments 1–4, yet the RTs vary depending on the other task. Thus, it is possible that the single-task RTs are similar when combination-specific learning is observed because the trials producing those RTs involve items thought to belong to the same task. In any case, it is clear that neither overlap in the stimulus or response modalities nor long or short single- or dual-task RTs determine whether combination-specific learning is observed.

Task Structure and the SRT Task

Although we did not formally measure explicit knowledge of which combinations were withheld, our informal interviews with participants indicated that there was little awareness of this manipulation. Moreover, Hazeltime et al. (2007) reported that participants were mostly unaware of the introduction of novel combinations after training. Thus, while we cannot make strong claims that the combination-specific learning we observed was implicit, we believe it is worthwhile to compare these results to those from implicit learning tasks.

One widely studied model task for implicit learning is the SRT task, which is often performed with a secondary task, which reduces, but usually does not eliminate, sequence learning (e.g., Cohen et al., 1990; Curran & Keele, 1993; Frensch, Lin, & Buchner, 1998; Hazeltime, Grafton, & Ivry, 1997; Mayr, 1996; Stadler, 1995). In a typical SRT experiment involving a secondary task, the two tasks are interleaved, and the stimuli for the two tasks are often presented along separate modalities (as in Experiment 1). In such cases, there are usually no contingencies across the two tasks, yet learning in the SRT task proceeds, despite the randomness of the secondary task.

However, some SRT studies have addressed the role of the relationship of the two tasks and its effect on learning. For example, Frensch, Buchner, and Lin (1994) manipulated the interval between the stimulus for the SRT task and the stimulus for the random, distracting task. Learning scores were smaller for participants who appeared to integrate the two tasks rather than treat them separately. Thus, it is consistent with the task structure account that they should show the greatest disruptions in learning, given that the distracting task was random. Similarly, Stadler (1995) found that the effects of varying the interval between stimuli for the SRT task were as disruptive to learning as adding a random distracter task. This led him to conclude that learning was not affected by limits in attentional capacity but rather changes in task organization. This proposal with its emphasis on the integrity of the task is consistent with the present study.

⁵ It is important to note the distinction between attention as a resource and selective attention. Jiménez and Méndez (1999) observed that selective attention was critical for implicit sequence learning, but that implicit sequence learning was not affected by the attentional load of the secondary task (attention as a resource). In the current study, both tasks required that participants attend to both stimuli on each trial (each task demanded speeded responses). However, similarly to Jiménez and Méndez, we argue that learning in these experiments was not contingent on attentional requirements (attention as a resource; see below).

⁶ Participants were instructed to produce both responses at the same time on dual-task trials, so the RTs for the two tasks were consistently similar for the two tasks on those trials, except for Experiment 6.

Table 1
Reaction Times (in Milliseconds) and Mean Learning Scores for the Six Experiments

Experiment	LH task	O task	L SRT	O SRT	L DRT	O DRT	Learning
1	V(s)-M	A(t)-V	565	793	831	876	-10
2	V(s)-M	V(s)-M	537	513	1,043	1,048	163 ^a
3	V(s)-M	A(t)-M	713	919	1,211	1,244	21
4	V(s)-M	V(s)-V	657	1,014	1,259	1,268	1
5	V(n)-M	A(n)-M	780	778	1,136	1,186	123 ^a
6	V(n)-M	V(n)-V	777	795	1,162	963	128 ^a

Note. LH = left hand (all the experiments included a task with visual stimuli that required button presses with the left hand); O = other; L SRT = left single-task reaction time (RT); O SRT = other (vocal or right hand) single-task RT; DRT = dual-task RT from the noncorresponding trials; Learning = mean learning score averaged across the two tasks (withheld minus practiced RTs); V(s) = visual spatial stimuli; V(n) = visual number stimuli; A(t) = auditory tone stimuli; A(n) = auditory number stimuli; M = manual responses; V = vocal responses.

^a Cost was significant.

Further evidence for the role of task integration in SRT learning comes from [Schmidtke and Heuer \(1997\)](#), who devised a variant of the secondary task procedure in which both tasks followed a fixed sequence. The length of the sequence for the secondary task was either the same as or different from the sequence for the primary task. When the lengths were the same, the two tasks could be thought of as from a single 12-element sequence, whereas when they were different, the sequence was likely too long to be learnable (30 elements). Learning was robust when the two sequences were the same length but diminished when the sequences were different lengths. This finding indicates that participants learned an integrated sequence that incorporated information from both tasks when the sequences were equal length.

More recent work has more directly examined the role of top-down factors using instructional manipulations. For example, [Halvorson et al. \(2013\)](#) had two groups of participants perform the same SRT task with a secondary random distractor task. One group was instructed to integrate the two tasks, while the other was instructed to view the two tasks as separate. Thus, the only difference between the groups was their conceptualization of the task—the timing of events was identical for the two groups. Because the former group integrated the distractor task, learning was significantly impaired compared to the latter, separate group. [Gaschler, Frensch, Cohen, and Wenke \(2012\)](#) also divided participants into two groups that received different task instructions: One group was told to respond to shape stimuli that were mapped to colored keys, and the other was told to respond to the same stimuli that were mapped to key locations. Critically, the two groups showed different patterns of transfer and interference, suggesting that learned representations guiding online behavior was altered by the task instructions. Similarly, [Kirsch and Hoffman \(2011\)](#) had participants practice a sequence, with one group being told to memorize a sequence of keys and another to memorize a sequence of finger movements. The former group was not able to transfer sequence knowledge to homologous fingers of the opposite hand, while latter group was. These results demonstrate that the task representation can have a profound impact on what is learned and/or expressed ([Abrahamse, Jiménez, Verwey, & Clegg, 2010](#); [Wenke & Frensch, 2005](#)).

With regard to the present findings, task representations may impose boundaries to facilitate learning and protect it from unre-

lated events that might lead to the formation of association that would be detrimental in other contexts (see [Gaschler et al., 2012](#)). Such a conclusion bears on the current debate about the boundaries of statistical learning processes and whether or not they are modality specific (e.g., [Conway & Christiansen, 2006](#); [Glicksohn & Cohen, 2013](#)). Task representations may affect which associations are formed regardless of the availability of attentional resources.

It is difficult to disentangle whether the differences among the results of the present experiments reflect how top-down factors affect the learning or the retrieval of combination-specific information. That is, do top-down processes help to encode information, or do they facilitate the expression what is learned? [Frensch and colleagues \(Frensch et al., 1998; Frensch, Wenke, & Rüniger, 1999\)](#) observed suppression of the expression of sequence learning when using a secondary counting task (learning was observed when the counting task was removed; see also [Gaschler & Nattkemper, 2012](#)). Thus, it may be that encoded associations between the two responses are not expressed when they are considered to belong to distinct tasks, even if the underlying learning mechanisms are not affected by the task representation.

Task Integration

Previous work has examined the factors that lead to the integration of different sources of information into a single task ([Cohen & Shoup, 1997](#); [Colzato, Raffone, & Hommel, 2006](#); [Hommel, 2004](#); [Magen & Cohen, 2007](#); [Meiran, Kessler, & Adi-Japha, 2008](#)). [Magen and Cohen \(2007\)](#) argued that response selection is not performed by a single, unitary process but rather is implemented by distinct processes acting on different stimulus modalities (e.g., color, shape, etc.). Such an account can readily explain the findings of Experiments 1–4 if it is further assumed that response modality also plays a role in determining the response selection module. It is less clear how this framework could be applied to address the findings of Experiments 5 and 6. It might be that the numerical codes used to link the two tasks in Experiments 5–6 caused them both to engage a numerical response selection module. However, if the active module can be determined by abstract properties rather than stimulus and response modalities, then it must be resolved how modules are instantiated (i.e., do they reflect the operation of hard-wired neural regions, or

are they flexibly set up based on the current task demands?) and how the appropriate module is selected when there are multiple possibly relevant modules (e.g., shape and number).

Another theoretical framework for understanding how sensory and motor information is integrated is the theory of event coding (Hommel, Müsseler, Aschersleben, & Prinz, 2001). Hommel et al. (2001) proposed that stimulus and response information is bound together in event codes. An event code consists of the integrated features of a stimulus, its response, and the environmental consequences of that response. An expanded view of these representations that allowed for them to include multiple, related events might offer a promising account of the current findings, especially if the events that were bound together into the “task code” were at least partially under top-down control. Along these lines, Colzato et al. (2006) proposed multiple levels of binding, beginning with low-level temporary bindings that are short term and context sensitive and higher level bindings that are long term and guide online processing. The higher level bindings might include task goals so that other actions are not included unless they are relevant to the same task. Their formation might also be facilitated by top-down control.

What Is a Task?

The notion of “task” pervades the cognitive psychology literature, underpinning theoretical accounts of phenomena such as dual-task costs and task switching. However, as other theorists have noted, the task construct remains poorly defined (e.g., Rogers & Monsell, 1995). The chord task, which involves multiple stimuli and multiple responses, appears to us to provide an intriguing test case: Should a chord trial be thought of as a single- or dual- (or multiple-) task condition? The answer, based on the patterns of learning, depends on how the participants conceive of their task.

The performance of a task is often interrupted by events that require attention but do not provide contingencies that can support learning. For example, when learning to ride a bicycle, the rider must attend to the road and traffic. These events (e.g., turns in the road and oncoming traffic) require responses that are essentially unrelated to the basic bike-riding procedure. The demands placed on cyclists by these unpredictable events should not interfere with their ability to learn. Thus, the actions required by the oncoming car or obstacle should not be confused by the learner as integral to bike riding but rather should be attributed to the particular circumstances when one attempts to ride a bike in a different environment.

Thus, it is useful to differentiate between contingencies that are inherent to the task and events that occur in atypical contexts. The top-down mechanisms identified by the present experiments are clearly able to play that role. The data suggest that the boundaries of a task may be in the eye of the performer rather than determined by primitive properties of the stimuli and responses. Such an arrangement would be extremely useful given the diverse endeavors that we attempt to master every day.

References

- Abrahamse, E. L., Jiménez, L., Verwey, W. B., & Clegg, B. A. (2010). Representing serial action and perception. *Psychonomic Bulletin & Review*, 17, 603–623. doi:10.3758/PBR.17.5.603
- Atkinson, R. C., & Estes, W. K. (1963). Stimulus sampling theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 121–268). New York, NY: Wiley.
- Ben-Artzi, E., & Marks, L. E. (1995). Visual–auditory interaction in speeded classification: Role of stimulus difference. *Perception & Psychophysics*, 57, 1151–1162. doi:10.3758/BF03208371
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory visual context guides spatial attention. *Cognitive Psychology*, 36, 28–71. doi:10.1006/cogp.1998.0681
- Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, 10, 360–365.
- Cohen, A., Ivry, R. B., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 17–30. doi:10.1037/0278-7393.16.1.17
- Cohen, A., & Shoup, R. (1997). Perceptual dimensional constraints in response selection processes. *Cognitive Psychology*, 32, 128–181. doi:10.1006/cogp.1997.0648
- Colzato, L. S., Raffone, A., & Hommel, B. (2006). What do we learn from binding features? Evidence for multilevel feature integration. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 705–716. doi:10.1037/0096-1523.32.3.705
- Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities pitting abstract against stimulus-specific representations. *Psychological Science*, 17, 905–912. doi:10.1111/j.1467-9280.2006.01801.x
- Curran, T., & Keele, S. W. (1993). Attention and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 189–202. doi:10.1037/0278-7393.19.1.189
- Dreisbach, G., & Haider, H. (2009). How task representations guide attention: Further evidence for the shielding function of task sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 477. doi:10.1037/a0014647
- Frensch, P. A., Buchner, A., & Lin, J. (1994). Implicit learning of unique and ambiguous serial transactions in the presence and absence of a distractor task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 567–584. doi:10.1037/0278-7393.20.3.567
- Frensch, P. A., Lin, J., & Buchner, A. (1998). Learning versus behavioral expression of the learned: The effects of a secondary tone-counting task on implicit learning in the serial reaction task. *Psychological Research*, 61, 83–98. doi:10.1007/s004260050015
- Frensch, P. A., Wenke, D., & Rüniger, D. (1999). A secondary tone-counting task suppresses expression of knowledge in the serial reaction task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 260–274. doi:10.1037/0278-7393.25.1.260
- Gaschler, R., Frensch, P. A., Cohen, A., & Wenke, D. (2012). Implicit sequence learning based on instructed task set. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1389–1407. doi:10.1037/a0028071
- Gaschler, R., & Nattkemper, D. (2012). Instructed task demands and utilization of action effect anticipation. *Frontiers in Psychology*, 3, 578. doi:10.3389/fpsyg.2012.00578
- Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron*, 54, 677–696. doi:10.1016/j.neuron.2007.05.019
- Glautier, S. (2002). Spatial separation of target and competitor cues enhances blocking of human causality judgments. *Quarterly Journal of Experimental Psychology: Section B. Comparative and Physiological Psychology*, 55, 121–135. doi:10.1080/02724990143000207
- Glicksohn, A., & Cohen, A. (2013). The role of cross-modal associations in statistical learning. *Psychonomic Bulletin & Review*, 20, 1161–1169. doi:10.3758/s13423-013-0458-4
- Goschke, T., & Bolte, A. (2012). On the modularity of implicit sequence learning: Independent acquisition of spatial, symbolic, and manual se-

- quences. *Cognitive Psychology*, 65, 284–320. doi:10.1016/j.cogpsych.2012.04.002
- Halvorson, K. M., Ebner, H., & Hazeltine, E. (2013). Investigating perfect timesharing: The relationship between IM-compatible tasks and dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 39, 413–432. doi:10.1037/a0029475
- Halvorson, K. M., Wagschal, T. T., & Hazeltine, E. (2013). Conceptualization of task boundaries preserves implicit sequence learning under dual-task conditions. *Psychonomic Bulletin & Review*, 20, 1005–1010. doi:10.3758/s13423-013-0409-0
- Harris, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological Review*, 113, 484–605. doi:10.1037/0033-295X.113.3.584
- Hazeltine, E. (2002). The representational nature of sequence learning: Evidence for goal-based codes. In W. Prinz & B. Hommel (Eds.), *Attention and Performance XIX* (pp. 673–689). Oxford, England: Oxford University Press.
- Hazeltine, E. (2005). Response–response compatibility during bimanual movements: Evidence for the conceptual coding of action. *Psychonomic Bulletin & Review*, 12, 682–688. doi:10.3758/BF03196758
- Hazeltine, E., Aparicio, P., Weinstein, A., & Ivry, R. B. (2007). Configural response learning: The acquisition of a nonpredictive motor skill. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 1451–1467. doi:10.1037/0096-1523.33.6.1451
- Hazeltine, E., Grafton, S. T., & Ivry, R. B. (1997). Attention and stimulus characteristics determine the locus of motor sequence encoding: A PET study. *Brain*, 120, 123–140. doi:10.1093/brain/120.1.123
- Hazeltine, E., Teague, D., & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 527–545. doi:10.1037/0096-1523.28.3.527
- Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, 8, 494–500. doi:10.1016/j.tics.2004.08.007
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24, 849–878. doi:10.1017/S0140525X01000103
- Hsiao, A. T., & Reber, A. S. (2001). The dual-task SRT procedure: Fine-tuning the timing. *Psychonomic Bulletin & Review*, 8, 336–342. doi:10.3758/BF03196170
- Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 236–259. doi:10.1037/0278-7393.25.1.236
- Kahneman, D., & Treisman, A. M. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29–61). Orlando, FL: Academic Press.
- Keele, S. W., Ivry, R. B., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review*, 110, 316–339. doi:10.1037/0033-295X.110.2.316
- Kirsch, W., & Hoffmann, J. (2011). Impact of instruction on the acquisition of sequence knowledge in a sensorimotor task. *Acta Psychologica*, 138, 85–91.
- Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature*, 383, 256–259. doi:10.1038/383256a0
- Kramer, A. F., Wickens, C. D., & Donchin, E. (1985). Processing of stimulus properties: Evidence for dual-task integrality. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 393–408. doi:10.1037/0096-1523.11.4.393
- Lachnit, H. (1988). Convergent validation of information processing constructs with Pavlovian methodology. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 143–152. doi:10.1037/0096-1523.14.1.143
- Li, W., Piëch, V., & Gilbert, C. D. (2004). Perceptual learning and top-down influences in primary visual cortex. *Nature Neuroscience*, 7, 651–657. doi:10.1038/nn1255
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492–527. doi:10.1037/0033-295X.95.4.492
- Logan, G. D., Taylor, S. E., & Etherton, J. L. (1999). Attention and automaticity: Toward a theoretical integration. *Psychological Research*, 62, 165–181.
- Magen, H., & Cohen, A. (2007). Modularity beyond perception: Evidence from single task interference paradigms. *Cognitive Psychology*, 55, 1–36. doi:10.1016/j.cogpsych.2006.09.003
- Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 350–364. doi:10.1037/0278-7393.22.2.350
- Mehta, R., & Williams, D. A. (2002). Elemental and configural processing of novel cues in deterministic and probabilistic tasks. *Learning and Motivation*, 33, 456–484. doi:10.1016/S0023-9690(02)00008-5
- Meiran, N., Kessler, Y., & Adi-Japha, E. (2008). Control by action representation and input selection (CARIS): A theoretical framework for task switching. *Psychological Research*, 72, 473–500. doi:10.1007/s00426-008-0136-8
- Melchers, K. G., Lachnit, H., Üngör, M., & Shanks, D. R. (2005). Prior experience can influence whether the whole is different from the sum of its parts. *Learning and Motivation*, 36, 20–41. doi:10.1016/j.lmot.2004.06.002
- Melchers, K. G., Shanks, D. R., & Lachnit, H. (2008). Stimulus coding in human associative learning: Flexible representations of parts and wholes. *Behavioural Processes*, 77, 413–427. doi:10.1016/j.beproc.2007.09.013
- Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I. Basic mechanisms. *Psychological Review*, 104, 3–65. doi:10.1037/0033-295X.104.1.3
- Miller, J. (1987). Priming is not necessary for selective-attention failures: Semantic effects of unattended, unprimed letters. *Perception & Psychophysics*, 41, 419–434. doi:10.3758/BF03203035
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1–32. doi:10.1016/0010-0285(87)90002-8
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, 116, 220–244. doi:10.1037/0033-2909.116.2.220
- Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, 94, 61–73. doi:10.1037/0033-295X.94.1.61
- Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, 101, 587–607. doi:10.1037/0033-295X.101.4.587
- Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *Journal of Neuroscience*, 26, 4970–4982. doi:10.1523/JNEUROSCI.3771-05.2006
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Rogers, R., & Monsell, S. (1995). The costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 124, 207–231. doi:10.1037/0096-3445.124.2.207
- Rosenbloom, P. S., & Newell, A. (1988). An integrated computational

- model of stimulus–response compatibility and practice. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21, pp. 1–52). San Diego, CA: Academic Press. doi:10.1016/S0079-7421(08)60025-0
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52. doi:10.1016/S0010-0277(98)00075-4
- Schmidtke, V., & Heuer, H. (1997). Task integration as a factor in secondary-task effects on sequence learning. *Psychological Research*, 60, 53–71. doi:10.1007/BF00419680
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1–66. doi:10.1037/0033-295X.84.1.1
- Schumacher, E. H., & Schwarb, H. (2009). Parallel response selection disrupts sequence learning under dual-task conditions. *Journal of Experimental Psychology: General*, 138, 270–290. doi:10.1037/a0015378
- Schumacher, E. H., Seymour, T. L., Glass, J. M., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central attentional bottleneck. *Psychological Science*, 12, 101–108. doi:10.1111/1467-9280.00318
- Seibel, R. (1963). Discrimination reaction time for a 1,023-alternative task. *Journal of Experimental Psychology*, 66, 215–226. doi:10.1037/h0048914
- Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: A characterization of the mind's bottleneck. *PLoS Biology*, 3, e37. doi:10.1371/journal.pbio.0030037
- Stadler, M. A. (1995). The role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 674–685. doi:10.1037/0278-7393.21.3.674
- Tillmann, B., Bharucha, J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, 107, 885–913. doi:10.1037/0033-295X.107.4.885
- Turk-Browne, N. B., & Scholl, B. J. (2009). Flexible visual statistical learning: Transfer across space and time. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 195–202. doi:10.1037/0096-1523.35.1.195
- Wasserman, E. A., & Miller, R. R. (1997). What's elementary about associative learning? *Annual Review of Psychology*, 48, 573–607. doi:10.1146/annurev.psych.48.1.573
- Wenke, D., & Frensch, P. A. (2005). The influence of task instruction on action coding: Constraint setting or direct coding? *Journal of Experimental Psychology: Human Perception and Performance*, 31, 803–819. doi:10.1037/0096-1523.31.4.803
- Wifall, T., McMurray, B., & Hazeltine, E. (2012). Perceptual similarity affects the learning curve (but not necessarily learning). *Journal of Experimental Psychology: General*, 143, 312–331. doi:10.1037/a0030865
- Williams, D. A., & Braker, D. S. (1999). Influence of past experience on the coding of compound stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, 25, 461–474. doi:10.1037/0097-7403.25.4.461
- Williams, D. A., Sagness, K. E., & McPhee, J. E. (1994). Configural and elemental strategies in predictive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 694–709. doi:10.1037/0278-7393.20.3.694
- Ziessler, M. (1998). Response-effect learning as a major component of implicit serial learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 962–978. doi:10.1037/0278-7393.24.4.962

Received September 4, 2013

Revision received February 13, 2014

Accepted February 14, 2014 ■