Temporal Yoking in Continuous Multitasking

Yuhong V. Jiang
University of Minnesota

Khena M. Swallow
Cornell University

Continuous tasks such as baggage screening often involve selective gating of sensory information when "targets" are detected. Previous research has shown that temporal selection of behaviorally relevant information triggers changes in perception, learning, and memory. However, it is unclear whether temporal selection has broad effects on concurrent tasks. To address this question, we asked participants to view a stream of faces and encoded faces of a particular gender for a later memory test. At the same time, they listened to a sequence of tones, pressing a button for specific pitched tones. We manipulated the timing of temporal selection such that target faces and target tones could be unrelated, perfectly correlated, or anticorrelated. Temporal selection was successful when the temporally coinciding stimuli were congruent (e.g., both were targets), but not when they were incongruent (i.e., only 1 was a target). This pattern suggests that attentional selection for separate tasks is yoked in time—when the attentional gate opens for 1 task it also opens for the other. Temporal yoking is a unique form of dual-task interaction.

Keywords: attention, target selection, temporal gating, dual-task processing

Selective processing of perceptual input is a foundational topic in attention research. Many early studies asked how people selectively attend to one channel of information, rather than another (e.g., Broadbent, 1958; for reviews see Lavie & Tsal, 1994; and Pashler, 1998). Still others examined situations in which both channels must be processed, emphasizing the costs of doing so (for a review, see Meyer & Kieras, 1997; Pashler, 1994). More recently, however, research has focused on selective processing within a single channel of input. This research recognizes that human performance often extends over time, and behaviorally relevant input occurs only some of the time. The behaviorally relevant input must be selected when it occurs (temporal attentional selection), enhancing the processing of stimuli presented at some moments in time but not at others (e.g., as in the "attentional blink," Chun & Potter, 1995; Dux & Marois, 2009; Raymond, Shapiro, & Arnell, 1992). The current study combines multichannel processing with single-channel temporal selection. Specifically, we ask, how does temporal selective attention for one task interact with selection for another concurrent task? Can people successfully select relevant information in one task while simultaneously filtering out information from the other task? Or does temporal selection in one task lead to a global attentional selection for stimuli from both tasks?

To make this question concrete, we will start by describing a dual-task scenario that is also used in this study. For one of the tasks, participants view a continuous visual stream of faces, but are asked to remember only a subset of the faces (e.g., male faces) for a later recognition memory test. All other faces can be ignored. For the other task participants listen to a continuous auditory sequence of high and low tones, but press a button only for tones of a particular pitch (e.g., the high tones). Both tasks are performed at the same time, with faces and tones occurring simultaneously and at a relatively slow pace of 1s/item. Importantly, each task requires temporal selection, and because the tones and faces are independent, selection could occur at different times for the different tasks. Can temporal selection on the faces occur independently from temporal selection on the tones? As will be reviewed next, our research question brings together two large bodies of literature that have until now being examined largely separately. Yet neither the dual-task processing literature nor the temporal attention literature makes clear predictions of how the temporal selection demands for two concurrent tasks will interact. Our study fills this gap.

Dual-Task Processing

Substantial research has demonstrated that multitasking produces interference: performance on one task usually suffers when conducted at the same time as a second task (Kinchla, 1992). This is observed both in discrete-trial tasks (Pashler, 1994) and in continuous tasks (Anderson et al., 2011; Horrey & Wickens, 2006; Kasper, Cecotti, Touryan, Eckstein, & Giesbrecht, 2014). However, the degree of interference depends on the similarity between the two tasks, whether they rely on the same sensory modality and motor effector, and instructions regarding which task should be given priority (Meyer & Kieras, 1997).

Several theories have been proposed to account for dual-task processing. Some theories focus on the overlap in processing machinery between concurrent tasks. For example, Wickens (2002) proposed that two tasks interfere more if they share processing demands on several dimensions (e.g., sensory modality,

format of coding). Similarly, Just and Varma (2007) and the Threaded Cognition theory (Salvucci & Taatgen, 2008) proposed that two tasks interfere more if they rely on the same cortical resources (Nijboer, Borst, van Rijn, & Taatgen, 2014). In the scenario described earlier, the face and tone tasks rely on different sensory modalities (visual vs. auditory), different output (encoding into memory vs. a motor response), and are both relatively simple. As a result, little dual-task interference should occur when performing these tasks at the same time rather than on their own.

The executive process-interactive control (EPIC) theory (Meyer & Kieras, 1997) also predicts relatively little interference. According to this theory, two tasks interfere if they share the same peripheral processing device (e.g., sensory processing or motor output). Because EPIC claims that central processing can be performed in parallel, tasks that use different modalities and produce different responses should interfere little. As for selection, EPIC only includes eye fixations as a visual perceptual filtering mechanism, and temporal selection is not considered separately from other cognitive processes. It is therefore likely that this theory predicts no interference between the face and tone tasks.

In contrast to EPIC and related theories, others have proposed that one or more central bottlenecks exist between perceptual processing and action (Pashler, 1994). For example, Pashler has proposed that the psychological refractory period (PRP) results from a central bottleneck in mapping a stimulus onto a response. Even a go/no-go task relies on the response selection bottleneck: the PRP is greater following a "go" response than following a "no-go" response (Bertelson & Tisseyre, 1969; Smith, 1967; as cited in Pashler, 1994). Central bottleneck theories have been tested primarily in situations in which both tasks require a speeded response. They nevertheless suggest that interference could occur for both the face/tone tasks described earlier. Although the face task does not require a speeded response, the encoding and consolidation of visual input to memory can slow responses to subsequent input (Chun & Potter, 1995; Jolicoeur & Dell'Acqua, 1998). Because a go response produces a larger PRP, face processing also should be impaired more when a target tone rather than a nontarget tone is presented. In addition, when a target face is presented, its memory load can interfere with participants' ability to prepare for the tone task (Logan, 1978; Pashler, 1994). This should result in impaired tone performance when a target face, rather than a nontarget face, is presented.[1]

In summary, existing dual-task processing theories do not specifically address how the temporal selection demands of two concurrent tasks may interact. Because the face and tone tasks do not share sensory modalities or motor effectors, and one task involves memory encoding rather than speeded decision, dual-task interference between the two tasks should be negligible (*perfect time-sharing hypothesis*). To the degree that interference is observed, the prediction may be this: face processing is more impaired when a target tone rather a nontarget tone occurs. Conversely, tone processing is more impaired when a target face rather than a nontarget face occurs (*target-induced interference hypothesis*).

The target-induced interference hypothesis is consistent with the two-target cost (Duncan, 1980). When asked to detect a character in horizontal locations and another character in vertical locations, people were successful if one of the locations contained a target, but the other did not. However, detecting a target in one location substantially impaired one's ability to detect a target in the other locations. Two-target costs have been observed in visual, auditory, and cross-modal target detection tasks (for a review, see Pashler, 1998). They suggest that selection is limited to one target at a time, supporting the idea that dual-task interference should be greatest when targets are presented at the same time. However, most accounts of dual-task performance have not generally considered processes that occur when information is selected in time, a topic examined by the temporal selection literature.

Temporal Selection and Attention in Time

Most daily activities are continuous over time but their demands on selective processing vary over time. For example, when driving it is more important to pay attention to pedestrians when approaching an intersection or crosswalk. Much like the control of spatial attention (Egeth & Yantis, 1997), the control of temporal attention in these situations could have both "bottom-up" and "top-down" components. For example, rhythmic neuronal activity in sensory cortical areas entrains (or "phase locks") to rhythmic sensory input (Cravo, Rohenkohl, Wyart, & Nobre, 2013; Schroeder & Lakatos, 2009). This entrainment, however, is more strongly driven by goal-relevant rather than ignored stimuli. When presented with regularly interleaved visual and auditory streams, neurons throughout the human brain entrain to the attended stimulus stream (Besle et al., 2011). Temporal selection is further demonstrated by the ability to direct attention to the moment in time when a stimulus is predicted to appear (Barnes & Jones, 2000; Miller, Carlson, & McAuley, 2013; Nobre, 2001) and to respond to behaviorally relevant stimuli that are unpredictably presented in a stream of distracting stimuli (Swallow & Jiang, 2013). Thus, sensory input and behavioral goals jointly contribute to the selection of information over time.

Several studies have shown that selecting a target in time leads to changes in perception, learning, and memory. Detecting a change in a display led to substantial increases in activity in dorsal frontoparietal regions important for attention, and in visual areas involved in perceiving the stimuli (Beck, Rees, Frith, & Lavie, 2001). Target detection is also important for learning. When repeatedly presented with the same visual displays in a change detection task, performance on these repeated displays improved only if the change was detected (Shen & Jiang, 2006). Similarly, contextual cueing, an implicit learning mechanism for repeated spatial contexts, occurs on target-present trials (Chun & Jiang, 1998) but not on target-absent trials (Kunar & Wolfe, 2011). Target detection also enhances visual long-term memory. When asked to remember all faces presented in a long stream but to press a button for only some of them (e.g., female faces), long-term memory for the target faces was better than that for the nontarget faces (Makovski, Jiang, & Swallow, 2013).

Single-unit neurophysiological studies have revealed possible neural correlates of target detection. When monkeys monitor a sequence of vertical and horizontal bars for targets (e.g., vertical

---

[1] Although inhibition is a primary factor in tasks such as go/no-go, it is negligible in our paradigm. As demonstrated in one recent study, when target, nontarget, and blank trials were compared, detecting a target enhanced memory for background images, but rejecting a nontarget did not inhibit memory, relative to blank trials (Swallow & Jiang, 2014b).

bars), detecting the target stimulus produces a transient increase in the activity of neurons in the Locus Coeruleus (for a review, see Aston-Jones & Cohen, 2005). These neurons release norepinephrine to cortical regions. Target detection and the associated LC-NE response could be considered a gating mechanism for information delivered over time (Nieuwenhuis, Aston-Jones, & Cohen, 2005; Swallow & Jiang, 2011, 2013).

The temporal gating idea is echoed in theories that were initially proposed to account for the attentional blink. When two targets are presented in an RSVP stream, detection of the first target (T1) impairs the processing of the second target (T2) if it appears approximately 200–500 ms later (Raymond et al., 1992). Curiously, the attentional blink is often absent if T2 follows T1 immediately. Several theories have been proposed to account for the attentional blink and "lag-1 sparing." In the Boost and Bounce Theory (Olivers & Meeter, 2008), target detection triggers an excitatory feedback response to perceptual areas that represent the target. In the simultaneous type, serial token model (ST^2), target detection triggers an attentional "blaster" that facilitates the maintenance of the target stimulus in working memory (Bowman & Wyble, 2007). These theories are consistent with experimental findings showing that temporal attention varies in response to behaviorally relevant events (Cravo et al., 2013; Large & Jones, 1999; Schroeder & Lakatos, 2009; Swallow & Jiang, 2013). However, these theories were developed to account for selective attention to an item in a single channel of information. They say little about how the attentional gating requirements for two temporally extended, but independent, tasks might combine in dual-task situations.

Temporal Selection in Concurrent Tasks

Are people able to select targets in one channel of input while simultaneously filtering out information from the other? Although this question is relevant to everyday multitasking, few studies have directly addressed it. As reviewed earlier, research on multitasking typically examines the overall costs of dual-task interference on primary task performance rather than on selective gating of information over time. Moreover, neuroscience studies on temporal attention typically examine attention to a single stimulus stream and therefore do not make clear predictions about how temporal selection works in dual-task situations.

One line of behavioral work has hinted at the difficulty in maintaining separate profiles of temporal attention. In the *attentional boost effect* (Swallow & Jiang, 2013), participants encode a continuous stream of pictures (e.g., faces) into memory. At the same time they monitor a second stimulus stream for occasional targets, such as a high-pitched tone (target) in a stream of low-pitched tones (nontargets). Because all faces are equally important, the encoding task requires participants to maintain a relatively constant amount of attention to the faces. In contrast, because some tones (targets) are more important than others, selective processing occurs in the tone task when target tones appear. If participants are able to maintain two independent attentional profiles for the two tasks, then selecting a target tone should have little effect on the face encoding task. In contrast to this prediction, memory for pictures encoded with target tones is enhanced relative to those encoded with nontarget tones (Lin, Pype, Murray, & Boynton, 2010; Swallow & Jiang, 2010, 2014a). These data suggest that temporal selective attention may be yoked across tasks in dual-task situations: it may not be possible to open the temporal attentional gate to one task without also letting in concurrent information from another task (the *temporal yoking hypothesis*).

The temporal yoking hypothesis implies that the attentional boost effect reflects an inability to maintain separate profiles of attentional gating for the picture encoding and tone detection tasks. Yet, the picture task requires constant attention to a large number of briefly and relatively rapidly presented background images, a task that is unlikely to be wholly successful (Potter, 1975). There may not be any incentive to prevent the temporal gating of the tone task from influencing image encoding. If this is the case, then the attentional boost effect does not necessarily indicate that people are unable to maintain separate temporal gating. Strong evidence for the temporal yoking hypothesis must come from studies that require participants to engage in temporal selection in both tasks.

To investigate whether people can maintain separate profiles of temporal selection, we performed several experiments using continuous streams of visual and auditory stimuli (which are less likely to compete for perceptual resources: Bonnel & Haftser, 1998; Duncan, Martens, & Ward, 1997; Treisman & Davies, 1973). Participants selectively encoded half of the visual stimuli (pictures of faces) and responded to half of the auditory stimuli (tones). We manipulated the temporal relationship between the visual and auditory targets. They could occur independently (Experiment 1), perfectly in-sync (Experiment 2A), or be anticorrelated (Experiment 2B). An additional experiment examined selective encoding when participants performed just the visual task (Experiment 3). These experiments allowed us to contrast the perfect time sharing, target-induced interference, and temporal yoking predictions.

Experiment 1

In Experiment 1 participants viewed a stream of faces at the same time that they listened to a sequence of tones. Faces and tones onset at the same time at a rate of 1s/item. Participants were asked to remember half of the faces (e.g., female faces) and to respond to half of the tones (e.g., high tones). However, the status of the face was unrelated to the status of the coinciding tone. This design produced four types of trials: target face + target tone (25%), target face + nontarget tone (25%), nontarget face + target tone (25%), and nontarget face + nontarget tone (25%). Participants' ability to select target tones was measured with a buttonpress response during the dual-task encoding phase. Their ability to selectively encode target faces was measured in a subsequent recognition memory phase. Output conflict for the two tasks was therefore minimized. We examined whether face processing is influenced by the tone task and vice versa.

Method

Participants. A prespecified sample size of 16 was used in all experiments. This sample size was comparable to those used in our previous dual-task studies (e.g., Swallow & Jiang, 2011, 2012). Participants were students at the University of Minnesota between the ages of 18 and 35 years old. They had normal or corrected-to-normal visual acuity, normal hearing, and passed a color blindness test. All participants were naïve to the purpose of the study

and completed just one experiment. Participants received extra course credit or $10 for their participation.

There were 12 females and four males in Experiment 1 with a mean age of 20.7 years. Data from two additional participants were dropped because their performance on the face recognition task was not higher than chance.

Equipment. Participants were tested individually in a room with normal interior lighting. They sat at an unrestrained distance of about 40 cm away from a 17" CRT monitor (screen resolution: 1024 × 768 pixels; vertical refresh rate: 75 Hz). The experiment was programmed with Psychtoolbox (Brainard, 1997; Pelli, 1997) implemented in MATLAB (*http://www.mathworks.com*) on a Macintosh computer. Tones were played through the computer's loudspeaker adjacent to the CRT monitor.

Stimuli. The auditory stimuli were high (650Hz) and low (350Hz) tones with a duration of 120 ms. The visual stimuli were front-view color photographs of famous faces chosen from the Internet (Swallow & Jiang, 2011). We used famous faces rather than novel faces because memory for novel faces was close to chance in a previous study with similar presentation rates (Makovski et al., 2013). Each face subtended approximately 11.5° × 11.5° and was surrounded by a red or blue outline frame (12° × 12°). The entire stimulus set contained 159 female faces and 307 male faces. A random subset of 312 faces—156 in each gender—was drawn for use in the main experiment. Randomization was performed for each participant separately. Another set of 44 black-and-white unfamiliar faces was used for practice.

Design and procedure. Following practice, participants completed (a) the dual-task encoding phase and (b) a recognition memory phase.

Dual-task encoding phase. The dual-task encoding phase was divided into six blocks of trials. Each block contained 156 trials presented continuously at a rate of 1s/trial. A face and a tone were presented concurrently for the first 120 ms of the trial, followed by a silent period of 880 ms during which a scrambled version of the preceding face was displayed (Swallow & Jiang, 2011). Male and female faces were equal in number (78 each), so were high and low tones (78 each). To ensure that memory for faces was not at floor, we showed the same faces six times, once per block. Although the stimuli were presented in a random order in each block, a specific face always coincided with the same type of tones (e.g., high tone) in all blocks.

For the face task, participants were asked to remember faces of a specific gender—either female or male—counterbalanced across participants. To enhance perceptual selectivity of the faces, different colored frames surrounded male and female faces. We told participants that they could rely on the frame color to help select the target faces. Participants made no overt responses to the faces. To reinforce the importance of selectively encoding just the male (or female) faces, at the end of each block we tested face memory for one of the faces. On this trial, two probe faces of the same gender were presented side-by-side, one of which participants had seen before. The "old" face was always one of the target faces. Different faces were probed after each of the six blocks.

For the tone task, participants responded to tones of a specific pitch by pressing the spacebar as quickly as they could. Responses must be made before the onset of the next tone. We encouraged participants to maintain perfect accuracy on the tone task. We displayed tone detection accuracy at the end of each block.

The gender of the face was unrelated to the pitch of the tone. Therefore, trials were divided randomly and evenly into four types: target face + target tone, target face + nontarget tone, nontarget face + target tone, and nontarget face + nontarget tone (see Figure 1). The gender of the target faces, the color of the frame surrounding the target faces, and the pitch of the target tones were fully counterbalanced across participants.

Readers can experience the dual-task encoding phase by viewing online demonstrations at the following website: http://jianglab .psych.umn.edu/TemporalYoking/TemporalYoking.html

Recognition memory phase. Two minutes after the encoding phase participants completed a two-alternative-forced-choice recognition memory test. At this point, participants were informed that, despite the instructions during the initial dual task, both target and nontarget faces would be tested and that they should try to be as accurate as possible. On each trial two faces of the same gender were presented side by side. Participants pressed a button to indicate which face they had seen during the encoding phase. Auditory feedback then informed them of their accuracy.

There were 144 recognition trials evenly divided among the target and nontarget faces. The six target faces that were tested during the encoding phase and six randomly selected nontarget faces were excluded from the recognition test.

Practice phase. Before the main experiment we familiarized participants with the dual-task encoding procedure. We used unfamiliar faces and shortened the practice block to 22 trials. All but one participant achieved at least 90% accuracy on the tone task after two practice blocks. The remaining participant completed three practice blocks to reach criterion (90% accuracy for tones).

Data analysis. Statistical analyses on the accuracy data were performed on both accuracy (proportion correct) and log-transformed accuracy to satisfy the linearity assumptions of ANOVA (Schweickert, 1985). The results were the same.

Results

Tone detection performance. In the dual-task encoding phase participants performed a go/no-go task on the tones. Table 1 shows the mean detection accuracy.

The target-induced interference hypothesis predicts that tone detection performance should be worse when the coinciding face was a target rather than a nontarget. To test this hypothesis, we calculated the tone-detection d' on trials with a target face (hit = 97.5%, false alarm = 5.3%), and tone-detection d' on trials with a
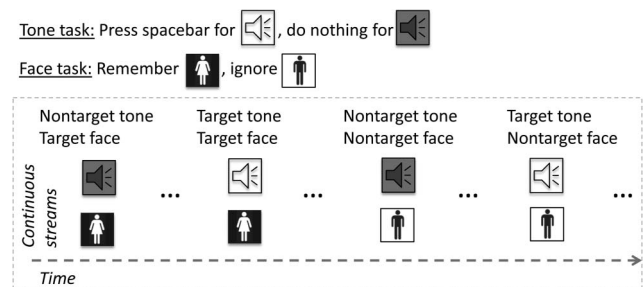

Figure 1. Design used in Experiment 1. Participants encoded faces of a specific gender to memory while monitoring tones of a specific pitch. The status of the face was unrelated to the status of the tone.

Table 1

Mean Response Rates to Target and Nontarget Tones in Experiment 1 (With the Between-Subjects Standard Error of the Mean)

Face condition	Target tone hit rate	Nontarget tone false alarm rate	RT hits (ms)	RT False alarms[*] (ms)
Target face	97.5% (0.5%)	5.3% (0.6%)	506 (18)	521 (30)
Nontarget face	96.3% (0.8%)	2.5% (0.6%)	521 (20)	395 (40)

[*] Reaction time (RT) in this column may be unreliable because of the small number of false alarm trials.

nontarget face (hit = 96.3%, false alarm = 2.5%). In contrast to the interference hypothesis, as shown in Figure 2A tone detection d' was numerically higher for target tones presented with target faces, rather than with nontarget faces. This difference was not statistically significant, $t(15) = 1.47$, $p > .15$. The data provided no evidence that the memory load associated with encoding a target face into memory interfered with tone detection response.

The temporal yoking hypothesis predicts that tone detection performance should be worse when the selection demands of the tone and face tasks were incongruent rather than congruent. To test this hypothesis, we calculated the tone-detection d' on trials when the tone and face stimuli were both targets or both nontargets (hit = 97.5%, false alarm = 2.5%), and the d' on trials when one of the stimuli was a target whereas the other was a nontarget (hit = 96.3%, false alarm = 5.3%). As shown in Figure 2B, d' was significantly higher on congruent trials than incongruent trials, $t(15) = 6.96$, $p < .001$. Reaction time (RT) was consistent with the accuracy results (see Table 1). These data provide initial support for the temporal yoking hypothesis. Because tone detection performance was influenced by the congruency between the tone and face stimuli, the data did not support the perfect time sharing hypothesis.

Face recognition memory. Figure 2C shows recognition memory for faces. If participants were able to selectively encode target faces and ignore nontarget faces, then recognition accuracy should be higher for target faces than nontarget faces overall. This was in fact the case, $F(1, 15) = 9.33$, $p < .008$, $\eta_p^2 = .38$ in

accuracy, $F(1, 15) = 9.49$, $p < .008$, $\eta_p^2 = .39$ in log-transformed accuracy.

Because go responses typically yield a larger PRP effect than do no-go responses (as reviewed in Pashler, 1994), the target-induced interference hypothesis predicts that the detection of the target tone should interfere with the encoding and consolidation of the coinciding face. To the contrary, our data showed that face memory was better, rather than worse, on target-tone trials than on nontarget-tone trials, $F(1, 15) = 6.78$, $p = .02$, $\eta_p^2 = .31$ in accuracy, $F(1, 15) = 7.38$, $p < .02$, $\eta_p^2 = .33$ in log-transformed accuracy.

In contrast, the temporal yoking hypothesis predicts that participants' ability to selectively process the faces depends on the congruency between the coinciding tone and face. When the tone and face were both targets or both nontargets, participants should be able to successfully select the target faces and ignore nontarget faces. But when the tone and face were incongruent (only one was a target), selective gating should fail. This leads to the counterintuitive prediction that memory for nontarget faces should be enhanced when the tones were selected. This should produce comparable recognition accuracy for target faces and nontarget faces.

As shown in Figure 2C, this prediction was supported by our data. Specifically, the status of the accompanying tone influenced performance on the face encoding task, resulting in a significant interaction between face status and tone status, $F(1, 15) = 13.99$, $p < .002$, $\eta_p^2 = .48$ in the accuracy data, and $F(1, 15) = 14.19$, $p < .002$, $\eta_p^2 = .49$ in log-transformed accuracy. Planned contrasts
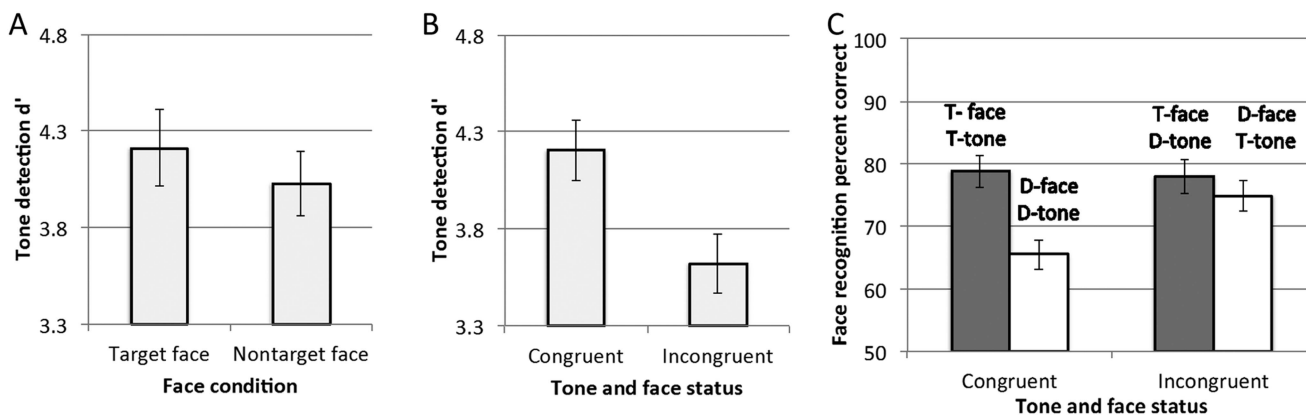

Figure 2. Results from Experiment 1. (A) Tone detection d' as a function of whether the coinciding face was a target or a nontarget face. (B) Tone detection d' as a function of whether the coinciding face was congruent (both targets or both nontargets) rather than incongruent with the tone (one target, the other nontarget). (C) Recognition accuracy for faces. T-Face: target face; T-tone: target tone; D-face: nontarget face; D-tone: nontarget tone. Error bars showed ± 1 between-subjects standard error of the mean.

showed that on congruent trials, target faces were better recognized than nontarget faces, $t(15) = 4.72$, $p < .001$ in accuracy, $t(15) = 3.98$, $p < .001$ in log-transformed accuracy. When the tone status was incongruent with the face status, however, memory for target faces was no better than memory for nontarget faces, $t(15) = 1.25$, $p > .20$ in the accuracy data, and $t(15) = 0.88$, $p > .35$ in log-transformed accuracy. Statistical significance remained the same with Bonferroni correction for multiple comparisons. Because tone status interacted with face performance, the data did not support the perfect time-sharing hypothesis.

Discussion

Experiment 1 revealed a unique form of dual-task interference. Rather than demonstrating target-induced dual-task interference, these data suggest an inability to selectively gate information from just one task. Specifically, it appears that when target tones were selected within the sequence of tones, the temporal attentional gate opened for the faces as well as the tones. As a result, both target and nontarget faces received additional processing. Conversely, when participants detected target faces, they were more likely to respond to the tone, whether it was a target or nontarget. Thus, selecting a target in one task led to the selection of stimuli presented in another task, regardless of whether it was a target or a nontarget. Attentional selection appears to be temporally yoked across tasks.

These data do not support the perfect time sharing hypothesis. They also contrast with standard dual-task interference effects observed in discrete trial designs. In those studies, dual-task interference manifests predominantly as a processing tradeoff between concurrent tasks. Furthermore, other research has shown that detecting a target in one set of spatial locations interferes with one's ability to detect a concurrent target in another set of spatial locations (Duncan, 1980; see also Pohlmann & Sorkin, 1976 for an example in auditory processing). How can we reconcile the discrepancy between the well-known two-target cost with the current finding? We believe there are two important differences. The first is that tasks that have been used to investigate the two-target cost often require similar kinds of processing for the targets. For example, in Duncan's classic experiments, both targets would have required visual analysis, identification, and consolidation at the same time. The dual-task and cognitive control literature accounts for these sources of interference. In contrast, the tasks used in Experiment 1 were designed to overlap as little as possible in modality and response. A second important difference is that Duncan's paradigm involves discrete trials that likely encourage a steady attentional state over time (Schroeder & Lakatos, 2009). In contrast, the multitasking paradigm used in Experiment 1 continuously presents a mixture of target and nontarget stimuli over time, and therefore necessitates selective gating of target stimuli. We believe that it is the selective gating of perceptual input in time that is yoked between concurrent tasks.

Temporal yoking highlights an important aspect of attention that has until recently been overlooked (Swallow & Jiang, 2013). Although concurrent tasks do compete for limited resources, selective attention is intrinsically dynamic. Some moments in time are more important than others. Consider the task of an airport baggage screener or someone who monitors the radar for specific signals. The visual input is continuous over time. Most of the time the input is "noise"; it is perceived, classified as "not of interest," and filtered out. But sometimes a suspicious object may be detected. This object, too, is perceived and classified, but its classification leads to additional processing and appropriate actions. This additional processing, variably described as a "boost" (Olivers & Meeter, 2008), "blaster" (Bowman & Wyble, 2007), or "temporal orienting" (Swallow & Jiang, 2011, 2013), can be collectively termed *temporal gating*. Data from Experiment 1 suggest that temporal gating is a global phenomenon. When we select a target from a continuous stream in one task, coinciding stimuli from the other task is selected as well.

Experiment 2

Many concurrent sources of information are correlated in their temporal structure. Speech, for instance, is conveyed both by spoken words and by facial expressions. In these situations, temporal yoking is likely to be advantageous—selective gating of correlated auditory and visual input should occur at similar times. Nonetheless, temporal selection demands in daily activities often vary for tasks that are unrelated, or for stimuli that are generated by independent sources. The temporal yoking hypothesis makes a strong prediction about when people can selectively gate sensory information in dual-task situations. Temporal selection should be successful when the two tasks have identical selection demands, but unsuccessful when they do not. Experiment 2 tested these predictions. By doing so, it also tests the strong version of the temporal yoking hypothesis—that selective gating cannot be independently applied to two tasks over time, even when the two tasks are perfectly anticorrelated.

Method

Participants. Thirty-two participants completed Experiment 2; 16 in Experiment 2A (11 females and 5 males, mean age 19.6 years), and 16 others in Experiment 2B (11 females and 5 males, mean age 19.2 years).

Design and procedure. This experiment was identical to Experiment 1 except for the temporal relationship between the target faces and target tones. On any given trial the status of the face was perfectly predicted by the status of the tone, and vice versa. In Experiment 2A the correlation was positive: target faces always co-occurred with target tones, and nontarget faces always co-occurred with nontarget tones. In Experiment 2B the correlation was negative: target faces always co-occurred with nontarget tones whereas nontarget faces always co-occurred with target tones (see Figure 3). Participants were not informed of the correlation. They received the same instructions as those of Experiment 1, including an emphasis on maintaining high accuracy in the tone task during the encoding phase.

Readers can experience the dual-task encoding phase by viewing the online demonstration movies: http://jianglab.psych.umn .edu/TemporalYoking/TemporalYoking.html

Results

Tone detection task. Table 2 shows the accuracy and RT in the tone detection task. The temporal yoking hypothesis predicts that the tone detection performance (d') should be higher when the
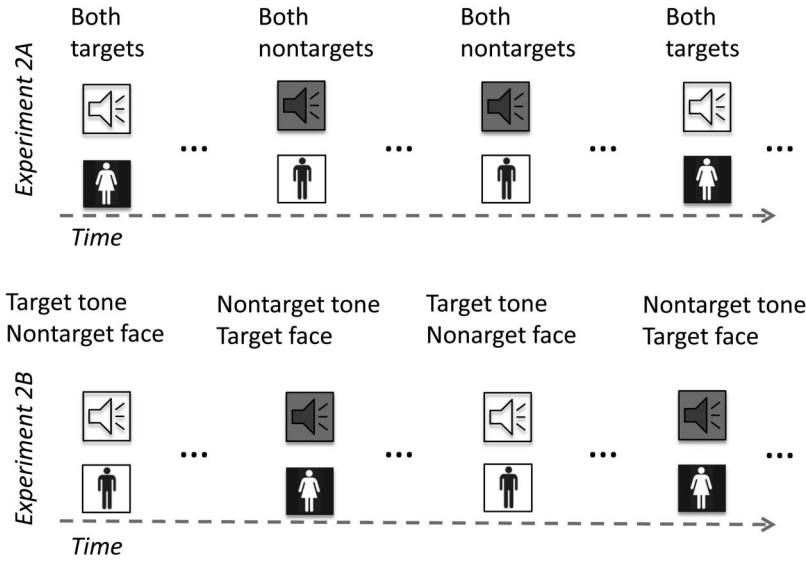
Figure 3. An illustration of the experimental designs for Experiments 2A and 2B. In Experiment 2A, target faces always coincided with target tones. In Experiment 2B target faces always coincided with nontarget tones.

tone and face stimuli are congruent (Experiment 2A) rather than incongruent (Experiment 2B). This was indeed the case (Figure 4A). Tone detection d' was significantly higher in Experiment 2A than in Experiment 2B, $t(30) = 2.51$, $p < .02$. Thus, the anticorrelation between target faces and target tones impaired people's ability to selectively attend to target tones.

Detection RT was comparable between the two experiments, $t(30) = 1.00$, $p > .30$, revealing no evidence for a speed–accuracy trade-off.

Face recognition memory. As shown in Figure 4B, participants' ability to selectively encode target faces was influenced by the congruency between the coinciding tone and face. An ANOVA on congruency (congruent—Experiment 2A vs. incongruent—Experiment 2B) and face status (target or nontarget) revealed a significant interaction, $F(1, 30) = 10.88$, $p < .003$, $\eta_p^2 = .27$ in accuracy, and $F(1, 30) = 11.77$, $p < .002$, $\eta_p^2 = .28$ in log-transformed accuracy. Planned contrasts showed that participants in Experiment 2A selectively encoded the target faces. Recognition accuracy was significantly higher for target faces than for nontarget faces, $t(15) = 5.93$, $p < .001$ in accuracy, $t(15) = 6.09$, $p < .001$ in log-transformed accuracy. In contrast, participants in Experiment 2B were unable to prioritize the target faces. Recognition accuracy was statistically indistinguishable between target faces and nontarget faces, $t(15) = 1.36$, $p > .15$ in accuracy, $t(15) = 1.27$, $p > .20$ in log-transformed accuracy. Statistical

significance remained the same with Bonferroni correction for multiple comparisons.

The overall level of accuracy was comparable between Experiments 2A and 2B, $F < 1$, suggesting that the congruency between faces and tones affected primarily selectivity (i.e., the difference between target and nontarget faces), rather than overall memory accuracy. These data showed that people successfully prioritized target faces if they always coincided with target tones, but failed to do so if they always coincided with nontarget tones.

Discussion

Experiment 2 evaluated two extreme cases of temporal correlation between concurrent tasks. In Experiment 2A the selection demands of both tasks were congruent: target tones coincided with target faces, and nontarget tones coincided with nontarget faces. Under these conditions selection was successful: tone detection was accurate and memory for the target faces was enhanced relative to the nontarget faces. In Experiment 2B the selection demands of the two tasks were incongruent: target tones always coincided with nontarget faces and nontarget tones always coincided with target faces. Performance on the tone detection task was impaired relative to Experiment 2A. Furthermore, participants in Experiment 2B failed to selectively encode target faces. Recognition memory was comparable between target and nontarget faces.

Table 2
Mean Response Rates to Target and Nontarget Tones in Experiment 2 (With the Between-Subjects Standard Error of the Mean)

Experiment	Target tone hit rate	Nontarget tone false alarm rate	RT hits (ms)	RT False alarms[*] (ms)
2A	98.8% (0.2%)	1.7% (0.4%)	477 (13)	457 (20)
2B	98.3% (0.3%)	3.3% (0.7%)	496 (14)	452 (25)

[*] Reaction time (RT) in this column may be unreliable because of the small number of false alarm trials.
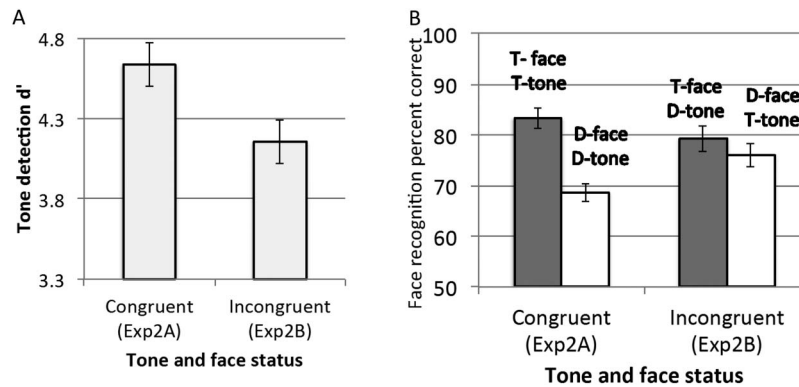
Figure 4. Results from Experiment 2. (A.) Tone detection *d'* in Experiments 2A and 2B. (B.) Face recognition accuracy in Experiments 2A and 2B. T-Face: target face; T-tone: target tone; D-face: nontarget face; D-tone: nontarget tone. Error bars show ± 1 between-subjects standard error of the mean.

Together, these two experiments provided compelling evidence for the temporal yoking hypothesis.

It may seem surprising that participants were unable to selectively encode the target faces in Experiment 2B. Much like Experiment 2A, the timing of targets in the tone task was redundant with the timing of targets in the face task in Experiment 2B. Every target tone signaled that the coinciding face should be ignored. Yet face selection was impaired, suggesting that participants could not take advantage of this correlational structure. The difficulty of maintaining independent temporal selection profiles appears to originate not from an uncertainty about when target faces might happen. Rather, it reflects an inability to selectively gate just one of the two concurrent tasks at any given time.

Experiment 3

The claim that temporal selection is yoked across tasks suggests that yoking can be detrimental to attentional selection. If so, then incongruent temporal attentional profiles could be a large source of dual-task interference. However, to interpret these data it is necessary to address one remaining question: How effectively can one select target faces for encoding when there are no other selection demands? Previous data suggest that incidental memory for visual objects may be as good as memory for intentionally encoded objects (Castelhano & Henderson, 2005). If this is the case, then it is essential to evaluate the degree to which selecting a face for encoding influences later memory.

In addition, although the data from Experiments 1 and 2 suggest that selection can be successful when the temporal profiles of the two tasks overlap, it is not clear whether overlapping attentional selection demands also interfere with task performance. Simply increasing working memory load increases interference from irrelevant stimuli in a spatial selection task (Lavie, Hirst, de Fockert, & Viding, 2004), raising the possibility that selective face encoding was impaired in Experiments 1 and 2A. Acquiring a measure of selection in the face encoding task under single-task conditions provides a basis for evaluating this possibility. It also allows us to assess the degree of interference that is present when the selection profiles are incongruent. Participants in Experiment 3 were therefore asked to encode only target faces, and to ignore all of the concurrently presented tones. If loading control processes impairs selection, then the selective encoding of target faces should be enhanced in Experiment 3.

Method

Participants. Sixteen new participants completed Experiment 3. There were nine females and seven males with a mean age of 20.8 years old.

Design and procedure. This experiment was similar to the first two experiments except that participants were told to ignore the tones and to concentrate on the face task. Participants intentionally encoded faces of a specific gender and ignored faces of the other gender. For half of the participants, the stimuli used in Experiment 3 were identical to those of Experiment 1: Female faces coincided with either low tones or high tones, and so did the male faces. For the other half of the participants, the stimuli used in Experiment 3 were identical to those of Experiment 2: a specific tone was always paired with a specific gender. This controlled for any low-level differences in stimulus presentation across experiments. However, because the results were identical between these two subgroups of participants (*ps* > .20), the data were pooled across all participants.

Results

Participants were able to selectively encode the target faces into memory. Recognition memory in Experiment 3 (see Figure 5) was significantly higher for target faces than nontarget faces, $t(15) = 5.13$, $p < .001$ in accuracy, $t(15) = 4.76$, $p < .001$ in log-transformed accuracy. If increasing the number of selection tasks impairs performance on its own (Experiment 2A), then overall accuracy should be lower in Experiment 2A than in Experiment 3. In addition, the difference between the target and nontarget faces should be smaller in Experiment 2A than in Experiment 3. This was not the case. Overall face memory was comparable between Experiment 2A and Experiment 3, $F(1, 30) = 1.37$, $p > .25$ in accuracy, and $F(1, 30) = 1.14$, $p > .25$ in log-transformed accuracy. In addition, target faces were remembered about 15% better than nontarget faces in both experiments, resulting in a lack of interaction between experiment and face condition, $F < 1$ in accuracy and log-transformed accuracy.
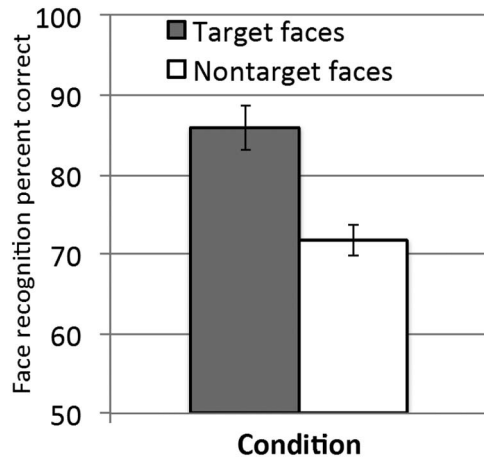
Figure 5. Face recognition accuracy in Experiments 3 (single task). Error bars show ± 1 between-subjects standard error of the mean.

If conflicting selection demands (Experiment 2B) interfered with temporal gating, then the difference between the target and nontarget faces should be greater in Experiment 3 than in Experiment 2B. This was confirmed by a significant interaction between experiment and face condition, $F(1, 30) = 8.73$, $p < .006$, $\eta_p^2 = .23$ in accuracy, and $F(1, 30) = 7.85$, $p < .009$, $\eta_p^2 = .21$ in log-transformed accuracy. This interaction was found in the absence of a main effect of experiment, $F(1, 30) < 1$ in accuracy and in log-transformed accuracy. These data showed that when performing concurrent visual and auditory tasks, the number of tasks per se does not significantly affect overall task performance. Rather, the congruency in temporal selection between the two tasks is a primary source of dual-task interaction.

Discussion

Experiment 3 provided a single-task baseline for gauging the success of selective encoding. When tones were ignored, participants successfully prioritized target faces over nontarget faces. However, selection was comparable to that shown by participants who were performing two temporally congruent selection tasks at the same time. In addition, the overall level of face memory in Experiment 3 was comparable to those observed in the dual-task experiments ($ps > .10$ when comparing Experiment 3 with Experiments 1, 2A, or 2B, individually or combined; bootstrapping analysis with 10,000 samples showed the same results; see also Bonnel & Haftser, 1998; Treisman & Davies, 1973). These data are consistent with the idea that people may be limited to imposing a single profile of temporal selection on all concurrent tasks.

Experiment 3 also demonstrates that the intention to remember an image can enhance visual long-term memory: Target faces were better remembered than nontarget faces. Previous research on verbal materials showed that long-term memory was sensitive primarily to the depth of encoding (Craik & Lockhart, 1972). Surprisingly, an intention to learn did not contribute to memory beyond the depth of the encoding task (Hyde & Jenkins, 1973). Consistent with the fact that scenes can be rapidly categorized when attention is diverted to other tasks and stimuli (e.g., Li, VanRullen, Koch, & Perona, 2002), intention also seems to have

little effect on the ability to remember objects presented in scenes (Castelhano & Henderson, 2005). However, other data suggest that intention can enhance visual long-term memory (Block, 2009) and that later memory for an image is impaired when attentional resources are diverted to other tasks (Strayer, Drews, & Johnston, 2003; Wolfe, Horowitz, & Michod, 2007). Our data support the claim that intention enhances visual long-term memory, at least under conditions that encourage competition between relevant and irrelevant images. In our experiments, the faces were presented relatively quickly and continuously, and encoding processes for one image may not be complete by the time the next image was presented. Processes that enhance memory for the target faces (e.g., consolidation and rehearsal) may continue when nontarget faces are presented.

General Discussion

This study demonstrates a distinct form of dual-task interference in continuous multitasking—temporal yoking. Selective encoding of a subset of faces is as effective when its selection demand aligns with that of a tone task as when the encoding task is performed on its own. However, when the temporal selection demands of the two tasks conflict, selection for both tasks is impaired. These data suggest that, in continuous tasks involving coinciding stimuli, attentional gating of sensory input occurs in unison.

The data from this study indicate that people have limited ability to restrict temporal selection to just one of two concurrent tasks. If multiple tasks differ in their temporal selection demands, then the primary task will exert the strongest control and the remaining tasks assume the temporal attentional profile of the dominant task. In our study, participants were asked to maintain high accuracy in the tone detection task and its selection profile dominated. Nonetheless, we observed evidence for mutual interference between the two tasks, suggesting that participants attempted to meet the selection demands of the face task. These attempts were largely unsuccessful, however; selective face encoding was unsuccessful when the status of the face conflicted with the status of the tone.

This finding differs from standard dual-task interference, in which increasing attention to one task impairs, or at best leaves unaffected, the concurrent performance of a second task (Kinchla, 1992; Pashler, 1994). It also differs from more recent data suggesting that temporal processing entrains to rhythmically presented stimuli (Large & Jones, 1999; Miller et al., 2013; Schroeder & Lakatos, 2009; Swallow & Jiang, 2013). To account for these data we first review their correspondence to other potentially related phenomena. We then reappraise current theories of dual-task performance.

Related Phenomena

To our knowledge this is the first demonstration of temporal yoking. However, other phenomena suggest that the ability to control attention in space and time may be fundamentally limited. Perhaps the best known examples of such limitations are in the ability to attend to multiple spatial locations at once. For example, in multiple object tracking studies people may be asked to maintain and update several (e.g., four) spatial attentional foci for an extended period of time (Pylyshyn, 1989; Pylyshyn & Storm, 1988). Some data from this literature suggest that multiple objects

cannot be independently tracked (Yantis, 1992). Rather, successful tracking may depend upon treating the objects as a single group, allowing a single updating process to track all of the objects. In another line of work on spatial attention, Spence, Driver, and their colleagues reported cross-modal control of spatial attention (for a review, see Driver & Spence, 1998). When presented with concurrent visual and auditory tasks, attending to the auditory input in one location facilitated processing of the visual stimuli from that location. It appears that spatial selection is yoked for visual and auditory stimuli. The degree of spatial yoking, however, is incomplete. Under appropriate conditions people were able to selectively attend to different locations for visual and auditory stimuli (Spence & Driver, 1996). Spatial attention and temporal yoking are clearly distinct, though they may be analogous effects in the spatial and temporal domains. Much like shared space, shared timing appears to be an important basis for attentional selection.

Within the temporal domain two phenomena that have more recently been described are likely to be related to temporal yoking. One of these is attentional entrainment to sensory stimuli (Large & Jones, 1999). In a recent study, participants were asked to saccade to a briefly presented visual stimulus (Miller et al., 2013). This stimulus was preceded by a sequence of 10 rhythmically presented tones. Although participants ignored the tones, their responses to the visual stimulus were facilitated if the visual stimulus was perfectly in sync with the auditory rhythm, relative to when it was offset by as little as 21ms. Miller et al. suggested that the auditory rhythm entrained attention, and consequently, increased the efficiency of visual processing at the moment the auditory tone was expected (see also, Cravo et al., 2013; Rohenkohl & Nobre, 2011). Similar to the current study, Miller et al.'s (2013) findings demonstrate a temporal correspondence between two channels of input. However, unlike the current study, attention was directed to a single visual stimulus and temporal selection was needed only at the end of the trial, if at all. The advantage for visual stimuli presented in sync with auditory rhythms therefore likely reflects attentional entrainment to rhythmic input (Schroeder & Lakatos, 2009). In our study, the stimuli were presented at a regular pace of 1s/item. However, the target stimuli were presented at irregular, nonrhythmic intervals. The rhythmic presentation of the stimuli may contribute to yoking (cf. Swallow & Jiang, 2013), but it cannot explain the selection advantage that is observed when the selection demands for two tasks occur at the same time.

Temporal yoking is perhaps most clearly related to the attentional boost effect, in which selecting a target for one task enhances later memory for a concurrently presented image. These studies typically require selection for just one of the two tasks, the detection task. The other task typically requires continuous encoding of the background stimuli, though these can also be incidentally encoded (Mulligan, Spataro, & Picklesimer, 2014; Swallow & Jiang, 2014a). Both the attentional boost effect and temporal yoking in the current study demonstrate that selecting a target can enhance the encoding of concurrent images. Yet, at least two important differences between these two phenomena exist. First, the attentional boost effect is larger when the background images are intentionally encoded, rather than incidentally encoded (at least when measured by explicit memory tests, see Mulligan et al., 2014). In contrast, selecting a target tone in this study had its clearest effect on concurrent image encoding when the coinciding image was a nontarget face (Experiment 1). Another important

difference is the fact that increasing the cognitive load of a task can make the attentional boost effect difficult to observe (Swallow & Jiang, 2010, 2014a), yet temporal yoking occurs even when multiple selection requirements must be maintained in working memory. Although neither of these differences make temporal yoking and the attentional boost effect incompatible, it will be important for future research to account for these differences in order to better characterize their relationship. Regardless of what that relationship might be, it is important to note that in the attentional boost effect temporal selection is required for just one of the two tasks. In contrast, selection is required for both tasks in the current study. Because yoked selection goes against task demand, the current study provides compelling evidence that temporal yoking is obligatory in dual-task processing.

Reappraisal of the Dual-Task Processing Theories

The experimental paradigm used here focuses on attentional gating over time, a topic that falls largely outside of traditional dual-task processing theories. The multiple resources view (Wickens, 2002), the cortical-resource theory (Just & Varma, 2007), and the EPIC theory (Meyer & Kieras, 1997) all predict negligible dual-task interference between our visual and auditory tasks. Consistent with these theories, overall face recognition accuracy was comparable between the dual-task experiments and the single-task experiment. However, the lack of an overall dual-task interference effect should not be taken as evidence for perfect time sharing. First, the task was designed to minimize the potential for interference in perceptual and response processing. Moreover, there were significant interactions across the two tasks: temporal gating for one task was successful only if its selection demands were congruent with those of the other task.

Other accounts of dual-task performance lead to the prediction that performance in one task should suffer if participants detect a target in the other task: the "go" response to a target produces a larger PRP effect than the "no-go" response, and encoding and consolidating a target face interferes with the preparation for the tone task (Pashler, 1994). In addition, the attentional demands increase when a target is detected, relative to when a nontarget is rejected (Duncan, 1980; Pohlmann & Sorkin, 1976). The pattern of dual-task interaction observed in the current study does not support the *target-induced interference account.* However, these dual-task theories were derived based on the attentional demand of detecting a target versus rejecting a nontarget. They were not designed to address issues associated with selective gating of sensory input at certain moments in time.

Given the different emphasis between our approach and the existing dual-task literature, one should not take our data as evidence against the existing dual-task processing theories. Many of these theories were not designed to account for conflicts in temporal selection. In fact, the concept of attentional gating is left out—most theories assume that the attentional gate is already open for both channels of input. The current study, however, required participants to determine when the gate should be open, and in so doing provided evidence suggesting that a single gating mechanism is used for all tasks.

What prevents people from maintaining independent temporal selection profiles for different tasks? The neurophysiological studies reviewed earlier shed some light on the underlying mechanism.

Detecting a target in one task produces a phasic increase in LC activity (Aston-Jones & Cohen, 2005; Nieuwenhuis et al., 2005). Because the LC projects broadly to cortical areas, its impact may not be specific to the processing of the target stimulus itself. Consistent with this possibility, detecting an auditory target increases activity in the primary visual cortex, and detecting a visual target increases activity in the primary auditory cortex (Swallow, Makovski, & Jiang, 2012).

At the behavioral level, the global nature of attentional gating may reflect the fundamental limit of human performance and the functional significance of selective attention. As noted by Allport (1989); Meyer and Kieras (1997), and others, the fundamental limit in human performance is most likely not in the amount of information that can be processed in parallel. But rather, people selectively process information to ensure that actions are coherent and goal directed. A major difference between the processing of behaviorally relevant information (e.g., suspicious objects on the radar screen) and irrelevant information (e.g., noise) is that relevant information demands an immediate, appropriate response. The data from the current study suggest that temporal gating may send a global "do something" signal to all channels and tasks when a target is detected in one channel or task. This could facilitate the coordination of different systems in the production of a coherent behavior.

An advantage of yoked temporal selection is to facilitate contextual learning. In real-world situations behaviorally relevant stimuli often occur within predictable contexts. For example, the turning of traffic signal from green to red often coincides with the change in traffic on the pedestrian walkway. Selective gating of all information at the time when a red light occurs may facilitate contextual learning of the lights and pedestrian traffic. In fact, spatial context learning in visual search tasks occurs when targets are detected (Chun & Jiang, 1998) but is absent when people do not detect a target (Kunar & Wolfe, 2011; Shen & Jiang, 2006).

Conclusion

Temporal yoking limits people's ability to correctly select sensory input in multiple tasks at the same time. However, it may also be important for successful cognition. Because visual and auditory stimuli in natural tasks tend to be semantically coherent and temporally coincident, yoked control may be adequate in most tasks that involve simultaneous streams of input. Even in situations when multiple sources of input are unrelated to one another, temporal yoking may be a beneficial way to facilitate learning of the context of important events (Chun & Jiang, 1998; Seitz & Watanabe, 2005). Thus, whereas temporal yoking may be problematic in some situations encountered in our modern world, it may also serve important adaptive functions.

References

Allport, A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). Cambridge, MA: MIT Press.

Anderson, J. R., Bothell, D., Fincham, J. M., Anderson, A. R., Poole, B., & Qin, Y. (2011). Brain regions engaged by part- and whole-task performance in a video game: A model-based test of the decomposition hypothesis. *Journal of Cognitive Neuroscience, 23,* 3983–3997. http://dx.doi.org/10.1162/jocn_a_00033

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal perfor-

mance. *Annual Review of Neuroscience, 28,* 403–450. http://dx.doi.org/10.1146/annurev.neuro.28.061604.135709

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology, 41,* 254–311. http://dx.doi.org/10.1006/cogp.2000.0738

Beck, D. M., Rees, G., Frith, C. D., & Lavie, N. (2001). Neural correlates of change detection and change blindness. *Nature Neuroscience, 4,* 645–650. http://dx.doi.org/10.1038/88477

Bertelson, P., & Tisseyre, F. (1969). Refractory period of c-reactions. *Journal of Experimental Psychology, 79,* 122–128. http://dx.doi.org/10.1037/h0026899

Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., . . . Schroeder, C. E. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *The Journal of Neuroscience, 31,* 3176–3185. http://dx.doi.org/10.1523/JNEUROSCI.4518-10.2011

Block, R. A. (2009). Intent to remember briefly presented human faces and other pictorial stimuli enhances recognition memory. *Memory & Cognition, 37,* 667–678. http://dx.doi.org/10.3758/MC.37.5.667

Bonnel, A. M., & Haftser, E. R. (1998). Divided attention between simultaneous auditory and visual signals. *Perception & Psychophysics, 60,* 179–190. http://dx.doi.org/10.3758/BF03206027

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review, 114,* 38–70. http://dx.doi.org/10.1037/0033-295X.114.1.38

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. http://dx.doi.org/10.1163/156856897X00357

Broadbent, D. (1958). *Perception and communication*. Oxford, UK: Oxford University Press. http://dx.doi.org/10.1037/10037-000

Castelhano, M., & Henderson, J. (2005). Incidental visual memory for objects in scenes. *Visual Cognition, 12,* 1017–1040. http://dx.doi.org/10.1080/13506280444000634

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71. http://dx.doi.org/10.1006/cogp.1998.0681

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 109–127. http://dx.doi.org/10.1037/0096-1523.21.1.109

Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684. http://dx.doi.org/10.1016/S0022-5371(72)80001-X

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *The Journal of Neuroscience, 33,* 4002–4010. http://dx.doi.org/10.1523/JNEUROSCI.4675-12.2013

Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences, 2,* 254–262. http://dx.doi.org/10.1016/S1364-6613(98)01188-7

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87,* 272–300. http://dx.doi.org/10.1037/0033-295X.87.3.272

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature, 387,* 808–810. http://dx.doi.org/10.1038/42947

Dux, P. E., & Marois, R. (2009). The attentional blink: A review of data and theory. *Attention, Perception, & Psychophysics, 71,* 1683–1700. http://dx.doi.org/10.3758/APP.71.8.1683

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48,* 269–297. http://dx.doi.org/10.1146/annurev.psych.48.1.269

Horrey, W. J., & Wickens, C. D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors, 48,* 196–205. http://dx.doi.org/10.1518/001872006776412135

Hyde, T., & Jenkins, J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior, 12,* 471–480. http://dx.doi.org/10.1016/S0022-5371(73)80027-1

Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36,* 138–202. http://dx.doi.org/10.1006/cogp.1998.0684

Just, M. A., & Varma, S. (2007). The organization of thinking: What functional brain imaging reveals about the neuroarchitecture of complex cognition. *Cognitive, Affective & Behavioral Neuroscience, 7,* 153–191. http://dx.doi.org/10.3758/CABN.7.3.153

Kasper, R. W., Cecotti, H., Touryan, J., Eckstein, M. P., & Giesbrecht, B. (2014). Isolating the neural mechanisms of interference during continuous multisensory dual-task performance. *Journal of Cognitive Neuroscience, 26,* 476–489. http://dx.doi.org/10.1162/jocn_a_00480

Kinchla, R. A. (1992). Attention. *Annual Review of Psychology, 43,* 711–742. http://dx.doi.org/10.1146/annurev.ps.43.020192.003431

Kunar, M. A., & Wolfe, J. M. (2011). Target absent trials in configural contextual cuing. *Attention, Perception, & Psychophysics, 73,* 2077–2091. http://dx.doi.org/10.3758/s13414-011-0164-0

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106,* 119–159. http://dx.doi.org/10.1037/0033-295X.106.1.119

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133,* 339–354. http://dx.doi.org/10.1037/0096-3445.133.3.339

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics, 56,* 183–197. http://dx.doi.org/10.3758/BF03213897

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America, 99,* 9596–9601. http://dx.doi.org/10.1073/pnas.092277599

Lin, J. Y., Pype, A. D., Murray, S. O., & Boynton, G. M. (2010). Enhanced memory for scenes presented at behaviorally relevant points in time. *PLoS Biology, 8,* e1000337. http://dx.doi.org/10.1371/journal.pbio.1000337

Logan, G. D. (1978). Attention demands of visual search. *Memory & Cognition, 6,* 446–453. http://dx.doi.org/10.3758/BF03197478

Makovski, T., Jiang, Y. V., & Swallow, K. M. (2013). How do observer's responses affect visual long-term memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1097–1105. http://dx.doi.org/10.1037/a0030908

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Pt. 1. Basic mechanisms. *Psychological Review, 104,* 3–65. http://dx.doi.org/10.1037/0033-295X.104.1.3

Miller, J. E., Carlson, L. A., & McAuley, J. D. (2013). When what you hear influences when you see: Listening to an auditory rhythm influences the temporal allocation of visual attention. *Psychological Science, 24,* 11–18. http://dx.doi.org/10.1177/0956797612446707

Mulligan, N. W., Spataro, P., & Picklesimer, M. (2014). The attentional boost effect with verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1049–1063. http://dx.doi.org/10.1037/a0036163

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin, 131,* 510–532. http://dx.doi.org/10.1037/0033-2909.131.4.510

Nijboer, M., Borst, J., van Rijn, H., & Taatgen, N. (2014). Single-task fMRI overlap predicts concurrent multitasking interference. *NeuroImage, 100,* 60–74. http://dx.doi.org/10.1016/j.neuroimage.2014.05.082

Nobre, A. C. (2001). Orienting attention to instants in time. *Neuropsychologia, 39,* 1317–1328. http://dx.doi.org/10.1016/S0028-3932(01)00120-8

Olivers, C. N. L., & Meeter, M. (2008). A boost and bounce theory of temporal attention. *Psychological Review, 115,* 836–863. http://dx.doi.org/10.1037/a0013395

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116,* 220–244. http://dx.doi.org/10.1037/0033-2909.116.2.220

Pashler, H. E. (1998). *The psychology of attention.* Cambridge, MA: MIT Press.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. http://dx.doi.org/10.1163/156856897X00366

Pohlmann, L., & Sorkin, R. (1976). Simultaneous three-channel signal detection: Performance and criterion as a function of order of report. *Perception & Psychophysics, 20,* 179–186. http://dx.doi.org/10.3758/BF03198598

Potter, M. C. (1975). Meaning in visual search. *Science, 187,* 965–966. http://dx.doi.org/10.1126/science.1145183

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32,* 65–97. http://dx.doi.org/10.1016/0010-0277(89)90014-0

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179–197. http://dx.doi.org/10.1163/156856888X00122

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18,* 849–860. http://dx.doi.org/10.1037/0096-1523.18.3.849

Rohenkohl, G., & Nobre, A. C. (2011). α oscillations related to anticipatory attention follow temporal expectations. *The Journal of Neuroscience, 31,* 14076–14084. http://dx.doi.org/10.1523/JNEUROSCI.3387-11.2011

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review, 115,* 101–130. http://dx.doi.org/10.1037/0033-295X.115.1.101

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences, 32,* 9–18. http://dx.doi.org/10.1016/j.tins.2008.09.012

Schweickert, R. (1985). Separable effects of factors on speed and accuracy: Memory scanning, lexical decision, and choice tasks. *Psychological Bulletin, 97,* 530–546. http://dx.doi.org/10.1037/0033-2909.97.3.530

Seitz, A., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences, 9,* 329–334. http://dx.doi.org/10.1016/j.tics.2005.05.010

Shen, Y. J., & Jiang, Y. V. (2006). Interrupted visual searches reveal volatile search memory. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1208–1220. http://dx.doi.org/10.1037/0096-1523.32.5.1208

Smith, M. C. (1967). The psychological refractory period as a function of performance of a first response. *The Quarterly Journal of Experimental Psychology, 19,* 350–352. http://dx.doi.org/10.1080/14640746708400114

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1005–1030. http://dx.doi.org/10.1037/0096-1523.22.4.1005

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied, 9,* 23–32. http://dx.doi.org/10.1037/1076-898X.9.1.23

Swallow, K. M., & Jiang, Y. V. (2010). The attentional boost effect: Transient increases in attention to one task enhance performance in a

second task. *Cognition, 115,* 118–132. http://dx.doi.org/10.1016/j .cognition.2009.12.003

Swallow, K. M., & Jiang, Y. V. (2011). The role of timing in the attentional boost effect. *Attention, Perception, & Psychophysics, 73,* 389–404. http://dx.doi.org/10.3758/s13414-010-0045-y

Swallow, K. M., & Jiang, Y. V. (2012). Goal-relevant events need not be rare to boost memory for concurrent images. *Attention, Perception, & Psychophysics, 74,* 70–82. http://dx.doi.org/10.3758/s13414-011-0227-2

Swallow, K. M., & Jiang, Y. V. (2013). Attentional load and attentional boost: A review of data and theory. *Frontiers in Psychology, 4,* 274. http://dx.doi.org/10.3389/fpsyg.2013.00274

Swallow, K. M., & Jiang, Y. V. (2014a). Perceptual load and attentional boost: A study of their interaction. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 1034–1045. http://dx.doi.org/10.1037/a0035312

Swallow, K. M., & Jiang, Y. V. (2014b). The attentional boost effect really is a boost: Evidence from a new baseline. *Attention, Perception, & Psychophysics, 76,* 1298–1307. http://dx.doi.org/10.3758/s13414-014-0677-4

Swallow, K. M., Makovski, T., & Jiang, Y. V. (2012). Selection of events in time enhances activity throughout early visual cortex. *Journal of Neurophysiology, 108,* 3239–3252. http://dx.doi.org/10.1152/jn.00472 .2012

Treisman, A., & Davies, A. (1973). Divided attention to ear and eye. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 101–117). San Diego, CA: Academic Press.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3,* 159–177. http://dx.doi.org/ 10.1080/14639220210123806

Wolfe, J. M., Horowitz, T. S., & Michod, K. O. (2007). Is visual attention required for robust picture memory? *Vision Research, 47,* 955–964. http://dx.doi.org/10.1016/j.visres.2006.11.025

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24,* 295–340. http://dx.doi.org/ 10.1016/0010-0285(92)90010-Y