2012

Interaction of Long- and Short-Term Memory on Selective Attention during Timbre Discrimination

Eva Garcia
CUNY City College

Interaction of Long- and Short-Term Memory on Selective Attention during Timbre Discrimination

A Thesis

Submitted to the Faculty

Of

The City College of New York

By

Eva Garcia

In Partial Fulfillment of the

Requirements for the Degree

Of

Masters of Arts

January 2011

Abstract

The present study investigated the effects of short- and long-term memory on processes of selective attention during timbre discrimination. Pitch served as the distractor dimension, held constant on standard trials and deviating from the standard frequency on distractor trials. Short-term memory was operationalized as levels of covariate context: Within a block of trials, pitch deviants ($p=.28$) were either absent (baseline condition), varied orthogonally (filtering condition) or systematically (positive and perfect conditions) with timbre values. Long-term memory was operationalized as levels of psychophysical context: Within a block of trials, the range of pitch change was psychophysically equated with timbre change in the low-imbalance condition, tilted slightly in favor of the distractor dimension in the medium imbalance condition, or tilted strongly in favor of the distractor dimension in the high imbalance condition. We found that the effects of imbalance (long-term memory) on subjects' performance (accuracy and reaction time) were mediated by the degree of covariate context (short-term memory): The potency of the imbalance manipulation (low=best, high=worst) on distractor disruption (deviant minus standard) grew as the correlation between timbre and pitch increased from 0.0 (filtering), to .72 (positive) to 1.0 (perfect). The results suggest an intimate relationship between attention and memory, with attention acting to suppress or resolve both short- and the long-term influences of distractor activation on target processing.

Introduction

Selective Attention

Selective attention is the ability to focus on task-relevant aspects of the environment and ignore task-irrelevant or distracting aspects. Selective attention is what allows humans to successfully hold a conversation in a crowded room when, for example, the many sights and sounds compete at once for one's attention. Without selective attention, goal-related behavior would be much more difficult to accomplish as one would need to process all incoming stimuli simultaneously.

Integrality

Though the effort to attend selectively is ubiquitous, the outcome of the selection process is frequently imperfect. Consider the cocktail party effect, first described by E. Colin Cherry (1953), who investigated the ability to attend to a single auditory source among multiple sources. Neville Moray (1959) pursued this phenomenon in the laboratory using dichotic listening tasks, with distinct auditory streams presented to each ear. Participants were asked to shadow, or immediately verbalize, what they just heard in the attended ear. Moray found that participants were unable at the end of the task to recall ordinary words presented in the un-shadowed ear. Nevertheless, one third of the participants were able to recall salient information like one's name, suggesting to Moray that selective attention is imperfect.

In many selective attention tasks, one property of a stimulus is made task-relevant whereas another is made task-irrelevant. In Moray's (1959) experiment, for example, spatial location (left or right ear) was relevant and semantic content was irrelevant. When studying pairs of dimensions belonging to one stimulus, such as the pitch and loudness of a sound, attention failures indicate whether the dimensions are integral or separable (Garner, 1974). Two dimensions are *integral* when one member of the dimensional pairing (e.g., pitch) cannot be attended to without the influence of the other member (e.g., loudness). In the case of pitch and loudness, if one were to instruct participants to attend to the pitch of a sound and found that accuracy and/or RT to pitch deteriorates when loudness varies (compared to when loudness remains constant across trials), one would say that pitch and loudness are integral dimensions. With integral dimensions, variations in the unattended dimension interferes with the perceiver's ability to attend to the target dimension. Pitch and timbre, pitch and loudness, and pitch and duration are examples of auditory integral dimensions (Melara & Marks, 1990; Schroger & Wolff, 1998). *Separable dimensions*, by contrast, are those in which one stimulus feature can be attended to selectively without the influence of the other. If, for example, researchers had found that pitch discrimination did not slow despite random changes in distractor values, then pitch and loudness would be classified as separable dimensions. An example of separable dimensions is circle size and diameter orientation (Garner & Felfoldy, 1970).

Garner's speeded classification paradigm often is used to assess whether dimensions are integral or separable. In this paradigm, stimulus values from two dimensions are manipulated across trials. One dimension, or channel, is to be attended (target channel), whereas the other is to be ignored (distractor channel). Within any particular block of trials the target channel must contain a minimum of two values, which subjects are asked to discriminate between, whereas the distractor channel must contain a minimum of one value. An auditory stream consisting of loud versus soft target values (indexed by decibels) at a constant pitch of 500 Hz would be an example of the simplest Garner-type condition, namely, the *baseline condition*. Here, participants are asked to discriminate between the two values of loudness while ignoring pitch.

The baseline condition is compared with conditions in which the two dimensions are varied relative to each other across trials. In the *filtering condition,* two dimensions are varied randomly (e.g., zero correlation between pitch and loudness). Thus in a stimulus set of two target values and two distractor values, four stimuli would be present in the filtering condition, with target value paired with each distractor values an equal number of times. In this way, distractor values are not at all predictive of target values. Finally, in the *correlated condition*, the two dimensions covary. A condition which has a correlation of 1.0 between pitch and loudness values, for example, indicates that each pitch target is always matched with a certain loudness value making the distractor dimension, in this case loudness, highly predictive of the target. If participants perform poorly in the filtering condition relative to baseline they are said to suffer *Garner interference,* a measure of selective attention failure. Garner interference means that the participant was unable to attend to the target dimension without impairment from the irrelevant variation along the distractor dimension and indicates that the dimensions being studied are integral. If the participants perform well in the correlated condition relative to baseline they are said to show *redundancy gain*. Integral dimensions typically reveal both Garner interference and redundancy gain in speeded classification.

Melara and Marks (1990) used the speeded classification paradigm to investigate selective attention of the auditory dimensions of pitch and timbre (sound quality). Two stimulus values were used to measure pitch and two stimulus values were used to measure timbre. Using different combinations of these values, participants were asked to complete ten tasks. Four of these tasks were baseline tasks: Subjects were asked to discriminate between two pitch values when one of the timbre values remained constant during the two pitch discrimination baseline tasks and two timbre values were used while one pitch value was used in the distractor channel during the timbre discrimination tasks. Melara and Marks were careful to match the speed and accuracy of pitch and timbre discrimination at baseline to ensure that they were balanced in discriminability. Two further conditions were filtering, one requiring pitch discrimination in which timbre could take on one of its two values randomly on each trial, the other requiring timbre discrimination, in which pitch could take on one of its two values randomly on each trial.

The remaining four conditions were correlated conditions. Each pitch value was matched with just one timbre value. Two types of correlated conditions were created. The *positively correlated condition* consisted of all congruent trials while the *negatively correlated condition* consisted of incongruent trials only. Congruence is an intrinsic attribute of correspondence between specific stimulus values often found in cognitive tasks. For example, in the well-known Stroop (1935) task, trials in which a color word is printed in a corresponding color are congruent, whereas trials in which word and color mismatch are incongruent. Participants tend to respond more quickly and accurately on congruent than on incongruent trials. Melara and Marks (1990) had no *a priori* basis for determining congruence, so assigned the labels arbitrarily: Congruent sounds were those that had a “twangy” quality (defined by a relatively short duty cycle) and were relatively high in pitch, or had a “hollow” quality (longer duty cycle) and were relatively low in pitch; conversely, incongruent trials were either twangy-low or hollow-high tones. In the positively correlated conditions only the two congruent combinations were included (with participants asked to focus on either the pitch or timbre dimension), whereas in the negatively correlated conditions only the two incongruent combinations were included.

Melara and Marks’ (1990) baseline RTs were comparable (345 ms) for both pitch and timbre values. Garner interference of 63 ms during timbre discrimination (pitch irrelevant) and 32 ms during pitch discrimination (timbre irrelevant) demonstrated an inability to attend selectively to either pitch or timbre when the other dimension varied orthogonally. Melara and Marks also found a redundancy gain, or improvement from baseline, of 7 and 18 ms in their positively correlated timbre and pitch tasks, respectively, and an improvement of 37 and 25 ms during negatively correlated timbre and pitch tasks, respectively. Thus, for both dimensions, participants performed significantly better during negatively correlated tasks, but only marginally so during positively correlated tasks, compared with baseline, suggesting to Melara and Marks that what they labeled congruent was really incongruent and vice versa. This notion is only strengthened by the fact that Melara and Marks found that incongruent trials were classified 15 ms faster than congruent trials across all conditions. These findings suggest that twangy-low pitched and hollow-high pitched sounds show a processing advantage over the other two pitch-timbre combinations. The upshot of Melara and Marks’ study is that the auditory dimensions of pitch and timbre meet Garner’s (1974) criterion for integral dimensions.

Memory and Selective Attention

Long-term memory can be defined as “the persistence of information over the long term, from hours, to days, to years”, whereas s*hort-term memory* refers to “the retention of information over seconds to minutes” (Gazzaniga, Ivry, & Mangun, 2002, pp. G-9). Both long- and short-term information may be active in *working memory,* “the transient representation of task-relevant information.” Working memory guides behavior in the present, and thus has been called “the blackboard of the mind” (Gazzaniga, Ivry, & Mangun, 2002, pp. G-10). Information held in working memory can increase or decrease the efficiency of attentional processing (Awh & Vogel, 2006; Cowan, 1995; Downing, 2000; Melara and Nairne, 1991; Postle, Brush & Nick, 2004). In fact, when performing attention and working memory tasks, distributed, overlapping brain networks are activated, including the mid-dorsolateral prefrontal cortex (PFC), the ventrolateral PFC, the parietal cortex, the anterior cingulate cortex, and the temporal cortex (Banich et al., 2000a, 2000b, 2001; Bledowski et al., 2004; Fan et al., 2003; MacDonald et al., 2000; Milham et al., 2001, 2003). Not surprisingly, then, recent theoretical models have considered explicitly how working memory modulates attentional processing (Awh et al., 2000; Cabeza et al., 2003; Lavie, 2005; Melara and Algom, 2003; Melara et al., 2005).

Short-term Memory and Selective Attention

A recent theory of selective attention called tectonic theory (Melara & Algom, 2003) holds that attentional processes are mediated by both short- and long-term memory mechanisms, which themselves are engaged by different environmental contexts as participants perform a cognitive task. For example, Melara and Algom suggested that information from a *covariate context* is extracted using short-term memory processes. They defined covariate context as the “correlation between stimulus values along two dimensions affecting the dimensional uncertainty of the stimulus” (Melara & Algom, 2003, pp. 439). Short-term memory, then, serves to compute the degree to which target and distractor information relates to each other within a given environment (e.g., a specific block of trials). The output of the calculation itself relates to one’s degree of confidence in reaching a decision about the current stimulus. In Melara and Marks’ (1990) study, for example, uncertainty was lower in each of the correlated conditions, because only two stimulus combinations were possible within a block of trials, than in the filtering conditions, in which on any trial four different stimulus combinations were possible. Melara and Algom claimed that information about the correlation between target and distractor values is held in short-term memory.

Sabri, Melara, and Algom (2001) demonstrated the behavioral effects of covariate context on performance during six experiments using a modified Stroop (1935) task. Covariate context was manipulated by changing the ratio of congruent and incongruent stimuli, i.e., the proportion of corresponding color-word and print-color combinations. Sabri et al.'s (2001) results showed that performance improved as covariate context increased regardless of changes made in three other context types (i.e., psychophysical, set size, and production contexts). This finding was in accord with tectonic theory's claim that performance improves as covariate context is strengthened. Because covariate context is extrinsically linked to short-term memory processes, as it is based on the proportion of recently presented corresponding stimulus values, one can conclude that short-term memory processes greatly influence behavioral responses in selective attention tasks.

Caclin, McAdams, Smith and Giard (2008) recently reported physiological evidence for the involvement of short-term memory processes on selective attention. They investigated integrality among three distinct timbre dimensions: attack time (ATT), spectral center of gravity (SCG), and spectral fine structure (EHA). ATT is a temporal attribute of timbre and refers to how quickly the amplitude of the sound rises (Sethares, 2005). SCG is a spatial attribute of sound commonly interpreted to represent the "brightness" of a sound with brightness increasing as the number of high frequencies increase. EHA is also a spectral property of sound that refers to the degree of attenuation of even harmonics in a tone relative to its odd harmonics (Seago, Holland

& Mulholland, 2010). Subjects were asked to discriminate between EHA values when either ATT or SCG was held constant in the distractor dimension.

Caclin et al. (2008) utilized the Garner paradigm to test interactions among ATT, SCG and EHA. They used a baseline and filtering condition as well as two correlated conditions with covariate contexts of 1.0 and –1.0. They found that participants performed worse on filtering tasks compared with baseline and were better on correlated tasks compared with filtering conditions, in accord with predictions of tectonic theory. (Subjects' baseline and correlated RTs were comparable across conditions). Caclin et al. also examined event-related potentials (ERPs) to each classification dimension in each condition in the P3 latency range, a range associated with working memory (Caclin et al. 2008) and short-term memory (Gross, Metz & Ullsperger, 1992; Gomer, Spicuzza & O'Donnell, 1979). The authors found reduced frontal negativity and enhanced posterior positivity between 250 and 500 ms after stimulus onset in correlated conditions compared with baseline conditions during both ATT/EHA and SCG/EHA classification. In contrast, they found an enhanced frontal negativity and a reduced posterior positivity in filtering relative to baseline. The results provide evidence of the physiological difference between conditions having distinct covariate contexts.

Long-term memory and selective attention

Melara and Algom (2003) also highlighted the role of long-term memory in extracting information from what they refer to as the *psychophysical context*, "the physical separation of values along dimensions affecting the imbalance between dimensions" (p. 439). The psychophysical separation between a 2000 Hz tone and a 1500 Hz tone, for example, is greater than that felt between a 2000 Hz tone and a 1950 Hz tone. Melara and Algom posited that the ease of separation is based on prior experience held in long-term memory. Practicing observers in distinguishing tiny stimulus differences, for example, can improve stimulus discriminability, as measured either behaviorally or physiologically (e.g., Tong, Rao, & Melara, 2009). According to tectonic theory, practice speeds access of stimulus representations in long-term memory. Long-term memory representations to target values that are distinct psychophysically are accessed relatively quickly and efficiently.

Physiological studies highlight the cognitive effects of experience on perception. A common paradigm used to study these effects is the passive oddball tasks in which a stream of stimuli – including frequent standard sounds and infrequent (oddball) deviant sounds – are presented while the observer is asked to attend to something else, such as reading a book. The mismatch negativity (MMN) component is a negative enhancement to the ERP waveform elicited by the oddball 100 to 250 ms after stimulus onset (Schroger & Wolff, 1998). Koelsch, Schroger & Tervaniemi (1999) found that the MMN is present in musicians but not musical novices during passive oddball tasks with very similar frequency harmonics. The authors also administered an active oddball task, asking participants to detect the oddball tones when they appeared. The behavioral results matched the physiological results: Musical experts remarkably outperformed musical novices (83% vs. 18% accuracy).

Tong, Melara and Rao (2009) used the oddball task to demonstrate effects of perceptual training on physiological and behavioral indices. A single training session was conducted consisting of seven conditions in which the deviant tone progressively become similar to the standard tone: 30, 26, 22, 18, 14, 10 and finally 8 Hz away from the standard tone of 1000 Hz. Progression from training condition to the next occurred as long as participants detected oddballs with at least 90% accuracy; otherwise the task was repeated. Behavioral and electrophysiological measures were taken before training, one week after training, and again nine weeks after training. Tong et al. (2009) found that training not only improved participants' sensitivity (d') and RTs to the oddball, but also the magnitude of the MMN, P2, and P3 ERP waveforms. Moreover, Tong et al. (2009) found very strong links between behavioral and electrophysiological responses. A correlation of -0.92 was reported between reaction time and the peak amplitude of the P2 ERP component (during active discrimination), suggesting that P2 reflected "the speed with which perceptual representations are accessed" (Tong et al., 2009, p. 84). This conclusion, coupled with the fact that P2 enhancement occurred during both passive and active tasks, supports the notion that P2 elicitation reflects automatic access to perceptual representations in long-term memory.

Long-term memory itself serves to compute the relative accessibility or imbalance of stimulus representations between, say, target and distractor values. Whenever task-relevant and task-irrelevant values access long-term memory representations at different rates they are said to be imbalanced. When target and distractor values access long-term memory representation at comparable rates, they are in balance. Relative accessibility is gauged in the Garner paradigm by observing participants' average correct RTs during baseline conditions. In Melara and Marks' (1990) study, for example, care was taken to ensure that pitch and timbre were in balance, that is, baseline pitch discriminability equaled baseline timbre discriminability. Stimulus values can be imbalanced in favor of the target dimension or the distractor dimension. According to Melara and Algom, the greater the imbalance favors discriminability of the distractor dimension, the more difficult it will be for observers to attend selectively to the target dimension, particularly when dimensions are integral.

Working Memory and Selective Attention

The products of both short-term memory and long-term memory regularly enter working memory, where selection processes can act upon them. One suggested functional role of working memory is in maintaining an attentional bias (Banich et al., 2000a, 2000b; Desimone and Duncan, 1995) or processing priority (de Fockert et al., 2001) that guides attentional selection. In the Stroop (1935) task, for example, certain PFC activations reliably signal the task requirements (Brass and von Cramon, 2004a; Derrfuss et al., 2005; for a review, see Brass et al., 2005) or task-relevant information (Banich et al., 2000a, 2000b, 2001) presumably held in working memory. As the load on working memory increases in an attention task, task-irrelevant information tends to be processed more extensively, manifested as increased activation in stimulus-specific areas of sensory cortex and resulting in larger behavioral interference from distractors (Banich et al., 2001; Lavie et al., 2004). One interpretation is that high memory load obscures processing priorities, allowing task-irrelevant information to undermine target recognition (de Fockert et al., 2001).

Chen and Melara (2011) recently proposed a formal model that integrates information from short- and long-term memory into working memory. The model is a variant of that proposed by Melara and Algom (2003) in which actual neural activity from scalp recordings was used to probe inhibitory control of distractors in working memory. The authors conducted five computer simulations of data reported in Melara, Chen, and Wang (2005); each simulation varied or eliminated specific mathematical parameters. In the original study, Melara, Chen, and Wang presented targets and distractors asynchronously in a modified Garner paradigm; participants were asked to discriminate between one of three target tones and to withhold a response during the presentation of three distractor tones, thereby permitting an analysis of how the memory of recent distractors affected the processing of current targets. Across conditions, Melara, Chen, and Wang manipulated imbalance to progressively favor distractors. They found that the greater the imbalance, the worse the behavioral performance and the weaker the magnitude of a slow EEG positivity occurring 400 ms to 700 ms after distractor onset. These findings highlight the role of distractor memories in modulating current target processing, each target appearing alone in this paradigm, without distractors concurrently present. Melara, Chen, and Wang therefore concluded that representations in working memory of the degree of distractor change effectively undermined maintenance of representations of target relevance.

Importantly, Melara, Chen, and Wang's (2005) behavioral results parallel those obtained when distractors are perceptually present (i.e., synchronous presentation of targets and distractors; e.g., Algom, Dekel, & Pansky, 1996; Melara & Mounts, 1993; Sabri, Melara, & Algom, 2001). Chen and Melara (2011) argued that working memory representations of distractors, in this case derived initially from long-term memory, act similarly to perceptual representations in modulating attentional processing. In their simulations, Chen and Melara found that slow-wave EEG activity to distractors was as good as free mathematical parameters in indexing the inhibitory processing to the long-term memory representations now residing in working memory.

Chen and Melara (2011) assumed that trial-to-trial variation in the distractor slow wave arises in part from physiological noise in short-term memory that reflects momentary fluctuations in inhibitory processing, and thereby contributes to trial-to-trial variation in external behavioral performance. They evaluated this assumption by comparing model simulations that predict behavioral variation using EEG-noise with those using completely random (e.g., Gaussian) noise. They reasoned that if EEG-noise simulations proved at least comparable to

Gaussian-noise simulations, it would demonstrate that electrophysiological activity could be incorporated into the dynamic operation of the model without an external random variable, thereby explicitly connecting physiological processing to specific predictions of behavioral performance. They found that electroencephalographic noise was superior to Gaussian noise in predicting trial-to-trial variability in behavior. The results suggest that inhibitory processes of the attentional system are mediated by short-term and long-term representations currently residing in working memory. The strength of activation of these representations is affected by environmental context, including covariate context in the case of short-term memory and psychophysical context in the case of long-term memory.

Distraction Paradigms

As we have seen, investigators have separately used the Garner paradigm and the oddball paradigm to probe aspects of selective attention processing. Schroger and Wolff (1998) developed a procedure that combined characteristics of the two paradigms as a method for investigating exogenous mechanisms of attention. In their task, participants discriminated between values of one dimension (e.g., timbre), while ignoring a second dimension (e.g., pitch), as in the traditional Garner paradigm. However, on the preponderance of trials, called standard trials, the irrelevant dimension was held at one value and only infrequently, on deviant trials, shifted to another value (e.g., lower pitch). Participants typically show a slowdown in speed to discriminate the relevant dimension when the irrelevant dimension is presented with its infrequent value, relative to when it is presented with its standard value. The slowdown is independent of actual physical value of the deviant because participants respond relatively quickly to a stimulus with this value when it occurs frequently.

The sudden change in distractor values is usually unpredictable, making the Schroger and Wolff (1998) paradigm – hereafter called the *Garner oddball task* – useful in studying exogenous orienting and reorienting to the relevant dimension. Indeed, Schroger and his colleagues have identified an ERP component called RON (reorienting negativity), a slow negative deflection elicited by the deviant distractors, that is thought to represent processes of reorientation. The Garner oddball task also enables researchers to probe the effects of covariate and psychophysical context both between and within conditions. For example, Schroger and Wolff investigated the effects of relative discriminability between pitch and duration dimensions. They varied the amount of change between standard distractor values and deviant distractor values by creating three filtering conditions: Standards and deviants differed by 50, 200, and 500 Hz in their low-, medium- and high-imbalance conditions, respectively, with duration (short or long) as the target. The investigators found that hit rates to deviants in the medium-imbalance condition were significantly worse than those in the low- and high-imbalance conditions (cf. Sabri et al., 2001).

Liu (2009) used the Garner oddball task to investigate the integrality of timbre and pitch. In pilot research, Liu found that timbre values of 20% and 40% duty cycle could be discriminated from one another at baseline (no deviants) at the same speed and accuracy as frequency values of 490 Hz and 510 Hz could be discriminated from one another. (A thorough description of duty cycle is given in the Methods section). Thus, a 20% timbre difference was balanced with a 20 Hz frequency difference. Liu (2009) then used these values to create two filtering oddball tasks, one in which pitch deviants were presented during timbre classification, the other in which timbre deviants were presented during pitch classification. He found that for both pitch and timbre classification participants were significantly slower in filtering (deviants present) than at baseline (deviants absent), i.e., Garner interference across dimensions. These results are in line with those found to integral dimensions in the traditional Garner paradigm.

Liu (2009) also was interested in how quickly participants were able to recover from the distraction caused by a deviant shift in the irrelevant dimension. He examined the speed and accuracy to the deviant stimulus, the standard stimulus that preceded a deviant, and the standard stimulus that followed a deviant separately. He found that participants were slowest and least accurate to the deviant, fastest and most accurate to the standard that preceded it, and intermediate to the standard that followed it. The results suggested that participants required at least one additional trial to recover and reorient their attention to task-relevant information from the distraction caused by a deviant presentation shift. We adopt similar measures of distraction and reorientation in the present experiment, which was conceived as an extension of Liu's study.

The Present Study

The present study used the Garner oddball paradigm to manipulate covariate context and psychophysical context in a factorial manner. The aim was to investigate the effects of short-term memory (covariate) and long-term memory (psychophysical) manipulations on exogenous orienting processes in attention. To create these manipulations it was necessary to modify the paradigm used by Liu (2009). The stimulus set used in his study included only two distractor values on each dimension, namely, the standard value and the deviant value. We increased the distractor stimulus set to three values in the current study: one standard value and two deviant values (of equal probability). This modification enabled us to distinguish between congruent trials – those deviant trials in which the distractor value corresponded to the target value – and incongruent trials – those deviant trials in which the distractor value did not correspond to the target value. In this way, we were able to create three levels of covariate context: (1) 0.0 correlation (filtering), in which 50% of the deviant trials were congruent; (2) .72 correlation (positive), in which 71% of the deviant trials were congruent; and (3) 1.0 correlation (perfect), in which 100% of the deviant trials were congruent.

We also created three levels of psychophysical context, using Liu's pilot results as our point of departure. In the low-imbalance condition, the degree of deviant change (20 Hz) matched psychophysically the degree of target change (20% duty cycle). In the medium-imbalance condition we enhanced the deviant change slightly (77 Hz) relative to the target change (20%). In the high-imbalance condition, deviant change (115 Hz) was significantly greater than target change (20%).

The study included several distinct measures of attention. By including a baseline task in which no deviant change occurred we were able to investigate condition-level effects, such as Garner interference in filtering and redundancy gain in correlated conditions. By including a mixture of congruent and incongruent trials in the filtering and positive conditions, we were able to measure congruity effects, i.e., the difference in performance between trial types. Finally, by including deviant trials in each non-baseline condition, we were able to measure the degree of exogenous reorienting to deviants and the degree of recovery from distraction to standards that followed a deviant.

In manipulating covariate context and psychophysical context parametrically we were able to examine whether each context separately affects each of our attention measures. These are revealed in ANOVA as main effects of each context. However, our primary theoretical objective was to investigate whether short-term and long-term memory representations mediate selective attention in a common mental workspace, such as working memory. In manipulating covariate context and psychophysical context factorially we were able to examine whether the two independent variables interact in their effects on selective attention. These are revealed in ANOVA as statistical interactions between the two contexts. By utilizing a paradigm that yields three different behavioral measures of selection failure – condition-level effects, congruity effects, and reorienting/recovery effects – the current study provides an especially sensitive test of this hypothesis.

Methods

Participants

Thirty-nine participants from The City College of New York participated in this study. Seven were unable to complete the task and five were excluded from the analysis because of experimenter errors in administering the task order. Data from 27 participants (16 male, mean age = 21.52, SD = 3.65) were included in the final analysis. All participants obtained a minimum of 80% accuracy in each condition. Participants volunteered or received course credit. All participants signed informed consent forms previously approved by The City College of New York's Institutional Review Board.

Stimuli and Procedure

Tones were created on Adobe Audition using a 48 kHz sampling rate with a 16-bit resolution. Tones were 69 decibels, calibrated on the A scale of a Quest Technologies Model 210 sound-level meter, with a duration of 100 ms (10 ms rise and fall times). Inter-stimulus intervals varied from 800 to 1200 ms in rectangular distribution. Presentation software (Neurobehavioral Systems) was used to present stimuli over Sennheiser HD280 pro headphones.

Timbre was defined by the attribute duty cycle of a pulse wave: the ratio between the pulse period – the period of time the wave is “on” - and pulse cycle – the duration of each wave repetition (see Figure 1a and 1b). For example, a duty cycle of 50% indicates that the pulse period is half that of the pulse cycle. In the present study, timbre values of 20% and 40% were used. The sound quality of short and long duty cycles used in this experiment were described to participants as “twangy/rough” and “smooth”, respectively.

Participants were administered 12 different conditions each consisting of 101 trials (see Figure 2 for an illustration of the task). Conditions were grouped into three sets: low imbalance, medium imbalance, and high imbalance. The sets were counterbalanced across participants. An identical baseline condition was included in each set. Here, only the target dimension changed from trial to trial (i.e., either 20% or 40% duty cycle), with auditory frequency remaining constant (500 Hz). Each timbre value appeared equally often in each condition.

Level of imbalance was defined by the frequency separation of the two pitch deviants relative to the timbre values used in a task. Frequency separation increased logarithmically across levels of imbalance. The frequency separation was 40 Hz in the low-imbalance condition (480 Hz vs. 520 Hz), 77 Hz in the medium-imbalance condition (463 Hz vs. 540 Hz), and 115 Hz in the high-imbalance condition (445 Hz vs. 560 Hz; see Table 1for all stimulus values and corresponding presentation rates).

Two deviant tones were used in all non-baseline conditions to provide information about all possible congruent (twangy-higher pitch and smooth-lower pitch) and incongruent (smooth-higher pitch tones) pairings regardless of the relative distractor value (higher or lower pitch). Congruent trials were defined arbitrarily as tones with relatively high frequency values (520 Hz in the low-imbalance condition, 540 Hz in the medium-imbalance condition, or 560 Hz in the high-imbalance condition) paired with short (20%) duty cycle or relatively low frequency tones (480 Hz in the low-imbalance condition, 4630 Hz in the medium-imbalance condition, or 445 Hz in the high-imbalance condition) paired with long (40%) duty cycle. Conversely, incongruent trials were defined as pairings of low frequency/short duty cycle or high frequency/long duty cycle.

Deviant values (p=0.28) were present in each (non-baseline) distractor condition. At least two standards occurred before each deviant. Specifically, two standards were presented before the deviant tone 16 times throughout each condition, three standards were presented before the deviant eight times and four standards were presented before the deviant four times in each condition. Each of these sequences appeared randomly within a condition. Each condition ended with a standard tone.

Each set of imbalance levels contained three distractor conditions (in addition to baseline): filtering, positively correlated, and perfectly correlated. Order of conditions was counterbalanced. The filtering condition (covariate context = 0.0) contained an equal number of timbre-pitch combinations within a given level of imbalance. For example, four deviant trials were used in the low-imbalance filtering condition: 480 Hz/20%, 520 Hz/20%, 480 Hz/40%, and 520 Hz/40%. Each deviant appeared an equal number of times (p = 0.25) in the task.

In the positively correlated condition (covariate context = 0.72) pitch distractors were paired on most (72%) trials with congruent timbre targets. Thus, each positively correlated condition contained 20 congruent trials and 8 incongruent trials. In the perfectly correlated condition (covariate context = 1.0) all deviant trials were congruent, i.e., twangy-high pitch and smooth-low pitch.

Participants were tested individually in a quiet room. They practiced baseline and filtering conditions before the experiment proper began. Practice conditions differed from experimental conditions in that subjects were provided feedback on the computer monitor ("correct!" or "incorrect!") after each response. They were instructed to pay attention to the sound quality of the tone (described as "twangy/rough" vs. "smooth") and to ignore any changes in pitch. Timbre discriminations were made by right or left mouse key; key assignment was counterbalanced across participants. Participants were asked to perform the task as quickly and accurately as possible. Any response occurring two seconds after stimulus onset was disregarded from the analyses. The entire experiment lasted approximately one hour.

Results

Interaction between Psychophysical and Covariate Contexts: Working memory processes

Table 2 contains a summary of performance in each condition. Speed and accuracy correlated -0.93, indicating the absence of a speed-accuracy tradeoff. Separate repeated measures analyses of variance (ANOVAs) were conducted on accuracy and RT, with stimulus Type (3 levels: standard presented before the deviant, deviant, standard presented after deviant), Psychophysical context (3 levels: low imbalance, medium imbalance, high imbalance), and Covariate context (3 levels: filtering, positive, perfect) as within-subject factors.

There was a significant main effect of Psychophysical context on overall accuracy, $F(2, 52) = 10.17$, $p < .01$. There also was a significant main effect of Psychophysical context on overall RT, $F[2, 52] = 4.61$, $p < .01$. Newman-Keuls post hoc analyses revealed that, for both measures, performance was significantly worst in the high-imbalance condition and best in the low-imbalance condition. There also was a marginal improvement in overall accuracy as

Covariate context increased, *F*(2, 52) = 2.63, *p* = .08, but not in RT, *F*[2, 52] = 0.53, *ns*. Most important, an interaction was found in overall accuracy between the psychophysical and covariate contexts, F[4, 104] = 5.98, *p* < .01, but not RT, *F*[*4*, 104] = .99, *ns*. There was also a three-way interaction in overall accuracy among stimulus Type, Psychophysical context, and Covariate context, *F*[8, 208] = 6.56, *p* < .01, but not RT, *F*[8, 208] = .92, *ns*. As can be seen in Figure 3, the interaction between psychophysical and covariate contexts was restricted to performance to the deviant stimulus. Here, participants showed improved accuracy to deviants as covariate context increased in the low- and medium-imbalance conditions, but showed progressively worse accuracy to deviants as covariate context increased in the high-imbalance condition. There were no significant interactions involving overall RT.

Distraction and Reorienting Effects

A main effect of stimulus Type was found for both accuracy (*F*[2, 52] = 142.96, *p* < .01) and RT (*F*[2, 52] = 98.29, *p* < .01). A Newman-Keuls post-hoc analysis of accuracy revealed that participants were significantly less accurate during deviant trials compared with accuracy to standards presented before or after deviants. A similar analysis of RTs revealed significant differences among all three stimulus types: Participants were slowest to deviants, intermediate to standards immediately after the deviant, and fastest to standards immediately before the deviant.

To explore more precisely the effect of the deviant stimulus on distraction, follow-up analyses were performed in which the accuracy or RT to the deviant was subtracted from the accuracy or RT to standards presented before the deviant. As shown in Figures 5A and 5B, the effects of distraction on these accuracy and RT difference scores were mediated by psychophysical context: Participants suffered progressively greater distraction as imbalance increased, *F*[2, 52] = 24.15, $p < .01$ (accuracy) and *F*[2, 52] = 11.84, $p < .01$ (RT). However, there was no effect of Covariate context on either accuracy, *F*[2, 52] = 1.10, *ns*, or RT, *F*[2, 52] = .41, *ns*. Nevertheless, we found an interaction between Psychophysical context and Covariate context on the difference scores in accuracy, *F*[4, 104] = 7.74, $p < .01$, but not RT, *F*[4, 104] = .84, *ns*. The nature of this interaction can be seen in Figure 5A: The disruptive effects of imbalance were strongest when covariate context was greatest (1.0) and weakest when covariate context was absent (0.0). More specifically, the difference in accuracy to deviants and standards grew with covariate context in the high-imbalance condition, but actually shrank with covariate context in the low-imbalance condition. A similar pattern also is evident in RTs (see Figure 5B), though the effect is not statistically significant.

To control for the potentially confounding effects of the physical stimulus on reorienting (deviants were physically different from standards), further follow-up analyses were performed in which the dependent variable was the difference in accuracy or RT between the standard before the deviant and the standard after the deviant. There was a significant main effect of Psychophysical context on the RT difference scores, *F*[2, 52] = 4.02, $p < .02$, but not on accuracy, *F*[2, 52] = .40, *ns*. Reorienting was worst in the high-imbalance condition and best in the low-imbalance condition. There were no other significant main effects or interactions.

Garner Effects

Another perspective on selective attention failure is to compare performance in baseline tasks, in which distracting deviants are absent, with performance in non-baseline tasks (filtering, positive, and perfect), in which deviants are present. We performed an ANOVA of difference scores, in which accuracy or RT in the baseline task was subtracted from that in the filtering, positively correlated, or perfectly correlated tasks. Only performance to standards was included in these analyses to equate the physical stimulus across tasks. The ANOVAs used a repeated measures design, with Psychophysical context and Covariate context as within-subject factors. There was a main effect of Psychophysical context on RT, $F[2, 52] = 4.38$, $p < .02$, but not accuracy, $F[2, 52] = 1.31$, *ns*. A Newman-Keuls post hoc analysis revealed significantly less impairment of standard trial classification during low Imbalance conditions compared to medium and high imbalance blocks. As one can see in Figure 7b, participants were slowed disproportionately in medium- and high-imbalance conditions for filtering, positively correlated, and perfectly correlated tasks relative to baseline. A marginal interaction was found in accuracy between Psychophysical context and Covariate context, $F[4, 104] = 2.27$, $p < .07$, but not RT, F [4, 104] = .99, ns.

Congruence effects

We examined intrinsic associations between pitch and timbre values by performing a congruity analysis. The analysis was restricted to deviant stimuli in the filtering and positively correlated conditions. Here, we defined congruent trials as those in which deviants were twangy/high pitch or smooth/low pitch, and incongruent trials as those in which deviants were twangy/low pitch or smooth/high pitch. A repeated measures ANOVA was performed on accuracy or RTs with Congruence (2 levels: congruent, incongruent), Psychophysical context and Covariate context (2 levels: filtering, positive) as within-subject factors. A significant main effect of Congruence was found for accuracy [$F(1, 26) = 6.39$, $p < .02$] but not RT [$F(1, 26) = 1.72$, $p < .20$] such that congruent trials were classified more accurately than incongruent trials. There was a significant interaction between Psychophysical context and Congruence found for both accuracy [$F(2, 52) = 26.09$, $p < .01$] and RT [$F(2, 52) = 6.37$, $p < .01$]. Participants performed better to congruent stimuli when imbalance was low or medium, but better to incongruent stimuli when imbalance was high (see Figures 8a and 8b). A marginal Covariate context x Congruence interaction was present in accuracy [$F(1, 26) = 3.32$, $p < .08$] but not RT [$F(1, 26) = 1.52$, *ns*]. Similarly, a marginal three-way interaction was found for accuracy [$F(2, 52) = 2.72$, $p < .08$] but not RT [$F(2, 52) = .68$, *ns*].

Discussion

The present study investigated the effects of short- and long-term memory on processes of selective attention during timbre discrimination. A significant interaction was found between psychophysical (long-term memory) and covariate (short-term memory) contexts in the analyses of overall accuracy and RT, distraction and Garner effects but not in the analyses of congruity and RON effects. The results with overall performance have two implications. First, memory of distractors, stored in short-term memory, greatly affects the ability of observers to control attentional processes. Second, the two memory systems do not act independently on attention. Specifically, the effects of redundancy (covariate context) are modulated by the degree of long-term memory change (psychophysical context). When dimensional imbalance was low, such that long-term memory representational changes in the distractor dimension were small, observers were able to reap the benefits of distractor redundancy and discriminate between target values more efficiently as covariate context increased from 0 to .72 to 1.0. However, when dimensional imbalance was high, and long-term representational changes in the distractor dimension were large, observers were misled by redundancy and hence showed an increased loss as covariate context increased. The interaction between psychophysical and covariate contexts demonstrates that output from short-term memory and long-term memory mechanisms combine at some point during cognitive processing where selective attention acts to excite target-relevant information and inhibit target-irrelevant information. Working memory may serve as a repository in which this interaction occurs.

Distractibility and Reorienting Effects

We investigated the integrality, distractibility, and correspondence properties of timbre and pitch dimensions. Our Garner oddball task showed predictable impairment in performance during deviant trials. Following Liu (2009), the present study found differences in RTs between standards presented before the deviants and standards presented after the deviants. Unlike Liu, these effects were not found in accuracy. Additionally, in our study as imbalance increased participants became progressively slowed to standards after deviants relative to standards before deviants, indicating difficulty in recovering from distraction. As seen in Melara and Marks (1990), a congruence effect was found between timbre and pitch dimensions. However, the direction of congruency depended on the level of psychophysical imbalance. Specifically, congruent trials during low- and medium-imbalance conditions were discriminated more efficiently than incongruent trials, whereas the opposite was found during high-imbalance conditions.

Participants were fastest to standards before a deviant compared with standards after a deviant, and to all standards compared to deviant trials. These findings are in accord with Liu's (2009) results. Nevertheless, our findings show smaller reorienting effects (standard before vs. standard after) than Liu, who obtained a statistically significant 27% decrease in accuracy and a 17% increase in RT between standards before and after deviants. In contrast, we found less than

1% decrease in accuracy (not significant) and 6% increase in RT. One possible explanation for the different outcomes focuses on differences in deviant probability between experiments. Liu (2009) used a deviant presentation rate of 0.1, whereas in the present study we used a deviant presentation rate of 0.28. The higher deviant probability in our study meant that deviants were more common, and hence less distracting, than in Liu's (2009) experiment. This explanation is supported by the fact that performance in tasks that were matched between Liu and our experiments – specifically, his filtering task and our low-imbalance filtering task – nonetheless yielded very different behavioral outcomes: 68% accuracy vs. 88% accuracy, respectively.

Another factor that may have influenced our results involves classification uncertainty. In Liu's (2009) study, participants were asked in each condition to attend either to a different timbre value or to a different auditory dimension (pitch vs. timbre). By contrast, our experiment asked participants to attend to the same timbre values across all conditions. It is conceivable that in Liu's (2009) study having to recall the current tasks' goals may have slowed participants' reorienting process once distracted. In our study, the constancy of target relevant information may have served to reduce participants' classification uncertainty, perhaps by maintaining an attention template, leading to relatively good performance to standards after deviants.

Congruence

We found that twangy/high pitched and hollow/low pitched tones were responded to relatively faster and more accurately in low- and medium-imbalance conditions, but not in the high-imbalance condition. This suggests that what participants regarded as congruent or incongruent was mediated by psychophysical context. Melara and Marks (1990), using the traditional Garner paradigm, found twangy/low pitched and hollow/high pitched tones showed relatively good performance. As Melara and Marks used balanced values, we can compare their target and distractor values to the values used in our low-imbalance condition. In this regard, the two experiments produced divergent results. Only performance in our high-imbalance condition mimicked Melara and Marks.

It should be noted that Melara and Marks (1990) used a different pitch range (900 and 920 Hz) and dissimilar timbre values (.1878 and .3128 duty cycle) compared with our study. Differences in absolute stimulus values make it difficult to compare directly timbre and pitch congruence effects for two reasons: (1) timbre, unlike loudness or pitch, is not a dimension inhering positive (high, loud) and negative poles (low, soft). Thus, it is difficult for researchers to associate timbre values with corresponding values on other dimensions; (2) unlike other dimensions, timbre and pitch interact psychophysically in very intricate and complex ways. Specifically, duty cycle is a combination of harmonics that combine to make rectangular-shaped waves. The fundamental pitch of a tone (e.g., 480 or 520 in our low-imbalance filtering condition) also defines the harmonics that make up that tone's timbre value and ultimately the way we perceive that timbre value. Changing the frequency of a tone (e.g., from 900 to 500 Hz) may alter our phenomenological experience of that tone's timbre, further complicating a researcher's ability to match pitch and timbre values.

The origins of congruence effects are likely based on the way we process sound. Caclin et al. (2008), for example, found that congruent stimuli not only yielded shorter reaction times but distinct ERPs and topographies. Caclin et al. (2008) observed interactions between temporal (ATT) and spatial (SCG/EHA) dimensions of timbre. Congruent trials in ATT/EHA conditions displayed a more pronounced positivity at temporal and posterior sites at around 70 ms. In SCG/EHA conditions, subjects displayed more pronounced positivities for congruent trials at posterior sites at about 170 ms. Congruent stimuli also displayed more positive potential fields during SCG/EHA conditions at 280 to 400 ms latencies. Caclin et al. (2008) claimed that latency differences were likely due to differences in the type of timbre properties used between these conditions. A divergence in congruent versus incongruent trials was predicted to take place sooner during ATT/EHA conditions because they represent different aspects (temporal and spectral, respectively) of timbre, whereas SCG/EHA conditions used similar (spectral) dimensions. Caclin et al. (2008) believe that the early congruency effects seen in their study, as compared with traditional Stroop (1935) studies, are due to their use of perceptual (early cognitive processing) versus semantic (late cognitive processing) nature of their stimuli. Because the EEG topography found in their study was similar in both ATT/EHA and SCG/EHA groups, Caclin et al. (2008) suggest that neural substrates in posterior areas of the brain are involved in extracting congruence properties. Further research should be conducted to assess whether these patterns are held during discrimination of timbre and pitch dimensions.

Long- and Short-term Memory Interaction

The interaction found between short-term and long-term memory processes indicate that these two mechanisms combine at some point during cognitive processing. We think a repository account best explains our findings. Here, short- and long-term memory processes are deposited into working memory where task-relevant and task irrelevant information are sorted. Working memory is thus a repository for all stimulus information (e.g., stimulus correlations, target and deviant values) in a given environment (e.g., condition). One role of selective attention is to sift through working memory by increasing excitability to task-relevant information and increasing inhibition to task-irrelevant information. Support for this claim comes a comparison of singlecell recordings of the prefrontal cortex (PFC), which show increased PFC activity maintained during the interval (Miller, Erickson & Desimone, 1996; Rainer, Asaad and Miller's, 1998), a sign of working memory activity, and neuroimaging recordings of PFC, which show increased PFC activity to distractors in Stroop tasks (Banich et al., 2000; de Fockert et al., 2004), a sign of increased inhibition to improve selection.

Alternative Theoretical Explanations

Aside from tectonic theory, other influential attention theories may help explain the current findings. In their biased competition model, Desimone and Duncan (1995) theorize that subjects maintain an attention template in working memory to bias processing in favor of task-relevant features during attention tasks. They claim that stimuli that the best match this attention template will be further processed to assess relevancy. Thus, unlike tectonic theory, Desimone and Duncan (1995) assert that distractors that share similar features to target stimuli will be selected via attention mechanisms for further processing causing more interference than distractors that are more distinct from target features.

Several features of our findings are inconsistent with Desimone and Duncan's (1995) theory. We found that performance worsened when the psychophysical distance between target and distractor values increased. On their account, performance should improve as target and distractor values become more distinguishable. The fact that degree of distractor change is more predictive of selective attention failure than degree of target-distractor similarity stands in direct contrast to Desimone and Duncan's predictions (see also Melara et al., 2005).

One reason why Desimone and Duncan's (1995) predictions are not held in the present study may be that their theory focuses on selective-set paradigms versus paradigms like ours called filtering paradigms. Selective-set paradigms are those in which observers are asked to locate one particular stimulus (usually a visual stimulus) often surrounded by irrelevant stimuli (e.g., Where's Waldo). Filtering paradigms, on the other hand, ask observers to discriminate certain features of a given stimulus within an environment. Chen and Melara (2011) argued that a higher premium is placed on inhibitory processes during filtering tasks compared with selective-set paradigms, since observers must actively highlight target features and inhibit distractor features during filtering paradigms, processes that are unnecessary when all stimuli are presented at once. During selective-set paradigms, mechanisms of comparison are those which observers must rely on the most. These factors may help explain why Desimone and Duncan's (1995) theory is less relevant to the present findings.

Nevertheless, Duncan and Desimone (1995) may provide some insight into the congruence effects found in the present study. Congruent trials are those whose task-relevant and task-irrelevant features share common features. In our study, participants performed better during these trials than during trials in which task-relevant and task-irrelevant values were more distinguishable, at least during low- and medium-imbalance conditions. These findings argue against Desimone and Duncan's (1995) claims that more similar target-distractor values cause more interference. Since a Garner effect was found, we can only assume that improved performance means that features in the distractor dimension helped participants discriminate between target values by sharing common properties. However, our high-imbalance condition revealed the opposite trend; participants performed better during incongruent trials compared with congruent trials. Here, similarity led to more interference during conditions of high-imbalance, a conclusion consistent with Desimone and Duncan's thesis. It may be that when inhibition processes are greatly impaired (e.g., during conditions of high imbalance), stimulus values that share similar features with target values become too difficult to separate perceptually and eventually lead to performance impairment.

By combining claims from both of tectonic and biased competition theory, it may be supposed that observed maintain a template of task-relevant information in working memory. However, unlike selective-set paradigms, they also maintain a rough sketch of task-irrelevant information when making a response. These working memory representations are enhanced during deviant presentation due to the activation of the long-term memory representations they elicit and the perceptual input derived from the environment. The efficient activation of deviant representations during high-imbalance conditions underscores the similarity of features between deviant and current target value during congruent trials. Activation on congruent trials of both long-term memory and working memory representations of distractor values may cause high confusability, a term Desimone and Duncan (1995) use to describe what occurs during similar target-distractor stimuli in selective-set displays. The greater saliency of distractor deviant values in a high-imbalance context may make more accessible and harder to suppress the long-term memory representations of these values. Low and moderate levels of imbalance may not as efficiently trigger long-term memory representations, making it easier to suppress these activations. The initial activation of the latter representations, however, may serve to help target selection as indexed by higher accuracy rates during both low- and medium imbalance conditions.

Our study could also be examined within the perspective of the load theory of attention (Lavie et al., 2004). Lavie and her colleagues claim that if perceptual task demands of target discrimination are high, distractors will not be efficiently perceived, and hence will not greatly affect performance. Normally, perceptual load is manipulated between conditions by increasing the difficulty of target discrimination. As all target values in our study were the same across conditions, we did not directly manipulate perceptual load. Instead, we manipulated differences along the distractor dimension. Nevertheless, we found that the greater the change in the distractor dimension relative to the target dimension (high imbalance), the worse was selective attention. Thus, difficult perceptual load along the distractor dimension, rather than along the target dimension, best explains good selective attention performance in the current study.

Conclusion

Our study aimed to observe the effects of pitch deviants on timbre discrimination. We found that more efficient access of long-term memory representations led to greater impairment in performance during deviant trials and to standards immediately after deviants. Short-term memory processes enhanced performance when long-term memory representations were weak to moderate, but impaired performance during conditions that strongly triggered long-term representations. Additionally, we found a reversal in congruity mediated by psychophysical context.

References

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

Awh, E., Vogel, E.K. & Oh, S. (2006): Interactions between attention and working memory. *Neuroscience* 139, 201–208.

Banich, M., Milham, M., Cohen, N. J., Wszalek, T., Kramer, A., Liang, Z.-P., Gullet, D., Shah, C., & Brown, C. (2000). Prefrontal regions play a predominant role in imposing an attentional ‘‘set’’: Evidence from fMRI. *Cognitive Brain Research*, 10, 1–9.

Banich, M., Milham, M., Atchley, R., Cohen, N.J., Wszalek, T., Kramer, A., Liang, Z.-P., Gullet, D., Shah, C., & Brown, C. (2000). FMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain system in attentional selection. *Journal of Cognitive Neuroscience,*12, 988–1000.

Banich, M. T., Milham, M. P., Jacobson, B. L., Webb, A., Wszalek, T., Cohen, N. J., & Kramer, A. F. (2001). Vision: From neurons to cognition. *Progress in Brain Research,* 134, 459-470.

Bledowski, C., Prvulovic, D., Goebel, R., Zanella, F.E. & Linden, D.E. (2004). Attentional systems in target and distractor processing: a combined ERP and fMRI study. *NeuroImage* 22, 530 –540.

Brass, M., & von Cramon, D. Y. (2004). Decomposing components of task preparation with functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, 16, 609–620.

Brass, M., Derrfuss, J., Forstmann, B., & von Cramon, D. Y. (2005). The role of the inferior frontal junction area in cognitive control. *Trends in Cognitive Sciences*, 9, 314–316.

Cabeza, R., Dolcos, F., Graham, R. & Nyberg L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage*, 16,*317-330.*

Caclin, A., McAdams, S., Smith, B. K. & Giard, M.H. (2008). Interactive processing of timbre dimensions: an exploration with event-related potentials. *Journal of Cognitive Neuroscience*, 20(1), 49–64.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America*, 25, 975-979.

Cowan, N. (1995). Attention and memory: An integrated framework. *Oxford Psychology Series*, 26, 1-4

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention, *Science*, 291, 1803-1806.

Derrfuss, J., Brass, M., Neumann, J., & von Cramon, D.Y. (2005) Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. *Human Brain Mapping*, 25, 22-34

Desimone, R. & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.

Downing, P.E. (2000). Interactions between visual working memory and selective attention. *Psychological Science*, 11(6), 467-473

Fan, J., Flombaum, J.I., McCandliss, B.D., Thomas, K.M. & Posner, M.I.(2003). Cognitive and brain consequences of conflict. *NeuroImage* 18, 42–57.

Garner, W. R. (1974) *The Processing of Information and Structure.* Potomac, Md.: Lawrence Erlbaum Associates.

Garner, W., & Felfoldy, G. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, *1*(3), 225-241.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2002). *Cognitive neuroscience: The biology of the mind* (2nd ed.). New York: Norton.

Gomer, F. E., Spicuzza, R. J., and O'Donnell, R. D. (1976). Evoked Potential Correlates of Visual Item Recognition During Memory-Scaling Tasks, *Physiological Psychology*, 4, 61-65.

Gross, B., Metz, A. M. & Ullsperger, P. (1992) The P300 Component of the Event-related Brain Potential in a Short-term Memory Paradigm. *Exp Angew Psychology*, 39(1), 56-67.

Lavie, N. (2005). Distracted and confused? selective attention under load. *Trends in Cognitive. Science*. 9, 75–82.

Lavie, N., Hirst, A., de Fockert, J., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, 133, 339-354.

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior attentive and pre-attentive auditory processing in musicians. *NeuroReport*, *10*, 1309–1313.

MacDonald, A. W., Cohen, J. D., Stenger, V. A. & Carter, C. S. (2000). Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control. *Science*, 288(5472), 1835-1838

Melara, R. D., & Nairne,J. S. (1991). On the nature of interactions between the past and the present. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 17, 1124-1135

Melara, R. D. & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, 110(3), 422-471.

Melara, R.D., Chen, S., & Wang, H. (2005) Inhibiting change: effects of memory on auditory selective attention. *Cognitive Brain Research.* 25, 431–442

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, 11, 56-60.

Melara, R., & Marks, L. (1990). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics*, *48*(2), 169-178.

Melara, R., & Marks, L. (1990). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 398-414.

Melara, R. D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition, 21*, 627–645.

Milham, M.P., Banich, M.T., Webb, A., Barad, V., Cohen, N.J., Wszalek, T. & Kramer, A.F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, 12, 467– 473.

Milham, M.P., Banich, M.T. & Barad, V. (2003). Competition for priority in processing increases prefrontal cortex's involvement in top-down control: an event-related fMRI study of the stroop task. *Cognitive Brain Research,* 17, 212-222.

Postle, B. R., Brush, L. N. & Nick, A. M. (2004). Prefrontal cortex and the mediation of proactive interference in working memory. *Cognitive, Affective & Behavioral Neuroscience,* 4(4), 600-608.

Sabri, M, Melara, R.D. & Algom, D. (2001). A confluence of contexts: asymmetric versus global failures of selective attention to Stroop dimensions. *Journal of Experimental Psychology and Human Perception and Performance*, 27, 515–537

Schröger, E., & Wolff, C. (1998). Behavioral and electrophysiological effects of task-irrelevant sound change: A new distraction paradigm. *Cognitive Brain Research*, *7*(1), 71-87.

Seago, A. Holland, S. & Mulholland, P. (2010). A novel user interface for musical timbre design. *Proceedings of the 128th Convention of the Audio Engineering Society.* London, UK.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Experimental Psychology* 18, 643–662

Tong, Y., Melara, R. D. & Rao A. (2009). P2 enhancement from auditory discrimination training is associated with improved reaction times. *Brain Research.* 1297, 80–88.

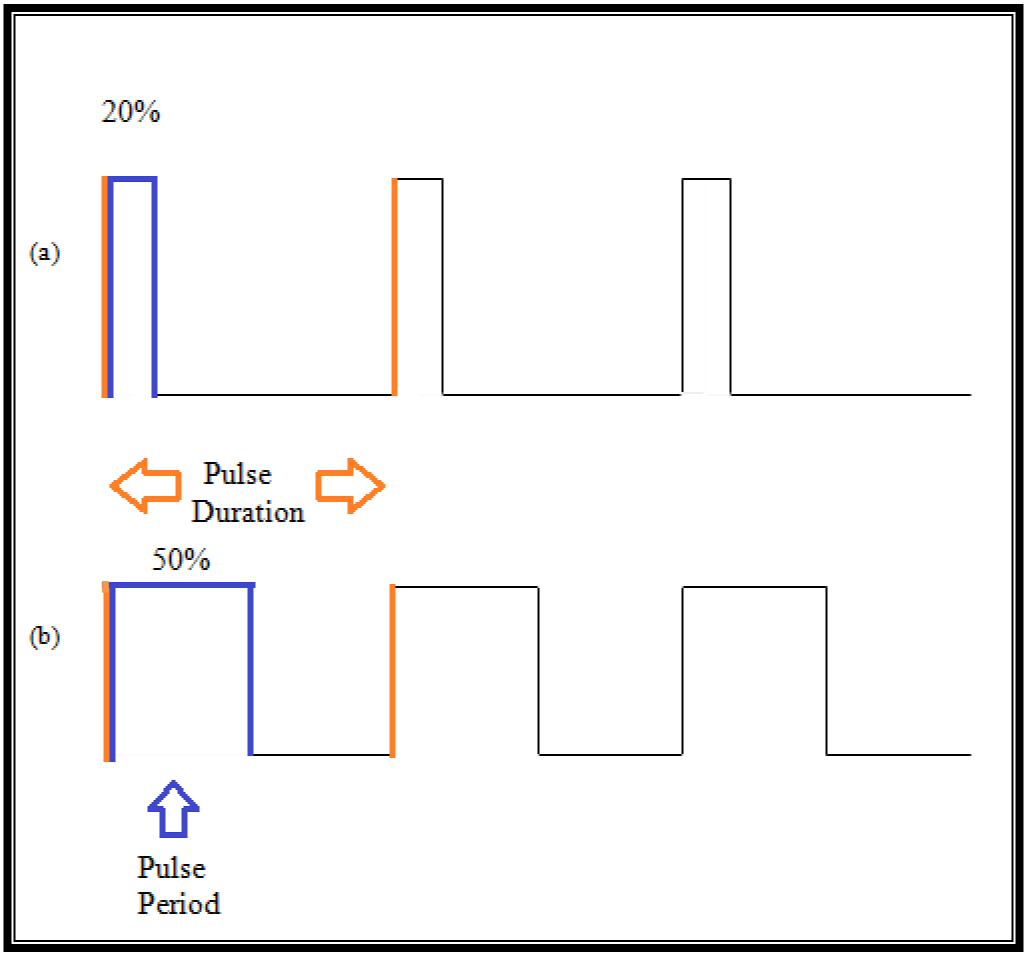


Figure 1. Illustration of (a) 20% and (b) 50% duty cycle. 20% and 40% duty cycle values were used in this experiment. (Adapted from Bar, 2001)

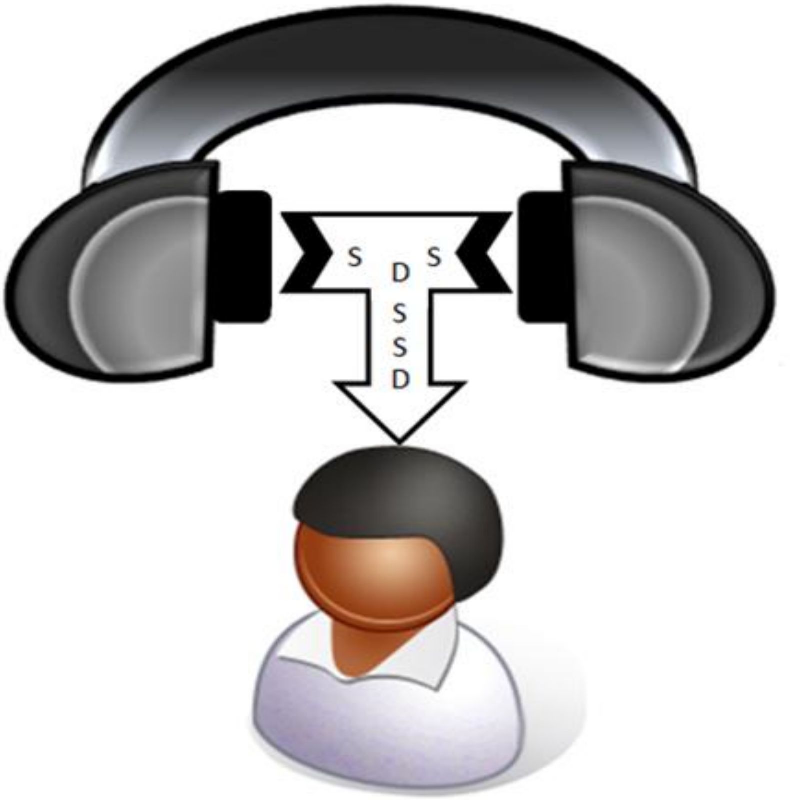


Figure 2. Schematic representation of the auditory task used in this study. “S” represents Standard tones, “D” represents Deviant tones.

Condition	Standards				Deviants							
	Stimuli	p.	Stimuli	p.	Stimuli	p.	Stimuli	p.	Stimuli	p.	Stimuli	p.
Baseline	20%-500 Hz	0.5	40%-500 Hz	0.5	-	-	-	-	-	-	-	-
Low-Filtering	20%-500 Hz	0.36	40%-500 Hz	0.36	20%-480 Hz	0.07	40%-520	0.07	40%-480 Hz	0.07	20%-520 Hz	0.07
Low-Positive	20%-500 Hz	0.36	40%-500 Hz	0.36	20%-480 Hz	0.04	40%-520 Hz	0.04	40%-480 Hz	0.1	20%-520 Hz	0.1
Low-Perfect	20%-500 Hz	0.36	40%-500 Hz	0.36	-	-	-	-	40%-480 Hz	0.14	20%-520 Hz	0.14
Medium-Filtering	20%-500 Hz	0.36	40%-500 Hz	0.36	20%-463 Hz	0.07	40%-540	0.07	40%-463 Hz	0.07	20%-540 Hz	0.07
Medium-Positive	20%-500 Hz	0.36	40%-500 Hz	0.36	20%-463 Hz	0.04	40%-540 Hz	0.04	40%-463 Hz	0.1	20%-540 Hz	0.1
Medium-Perfect	20%-500 Hz	0.36	40%-500 Hz	0.36	-	-	-	-	40%-463 Hz	0.14	20%-540 Hz	0.14
High-Filtering	20%-500 Hz	0.36	40%-500 Hz	0.36	20%-445 Hz	0.07	40%-560 Hz	0.07	40%-445 Hz	0.07	40%-560 Hz	0.07
High-Positive	20%-500 Hz	0.36	40%-500 Hz	0.36	20%-445 Hz	0.04	40%-560 Hz	0.04	40%-445 Hz	0.1	40%-560 Hz	0.1
High-Perfect	20%-500 Hz	0.36	40%-500 Hz	0.36	-	-	-	-	40%-445 Hz	0.14	40%-560 Hz	0.14

Table 1. This table includes all stimuli used in this study and their corresponding proportion (p.) per condition.

Task Type		Mean Accuracy	Standard Error	Mean Reaction Time	Standard Error
Low	Baseline	96.51631831	0.559436848	484.7556711	15.15734288
	Filtering	92.4825816	0.748396971	525.4111369	15.16525785
	Positive	93.3626696	0.798057706	534.7046342	18.11044127
	Perfect	95.48954896	0.614342105	519.4063675	14.15340917
Medium	Baseline	96.88302165	0.689370485	472.2954671	13.93293373
	Filtering	91.0779967	0.881404705	543.9791801	18.01605929
	Positive	92.00586726	0.69948299	546.5591439	16.67305204
	Perfect	93.06930694	0.961853837	552.49747	20.96781971
High	Baseline	97.46974699	0.464812452	477.472781	14.41110669
	Filtering	91.89475614	0.687765605	549.0166051	15.5410839
	Positive	90.24569125	0.850093482	552.136642	14.7870591
	Perfect	90.02566924	0.953324197	558.042555	18.068536

Table 2. Mean accuracy and reaction times across conditions with corresponding standard error values.

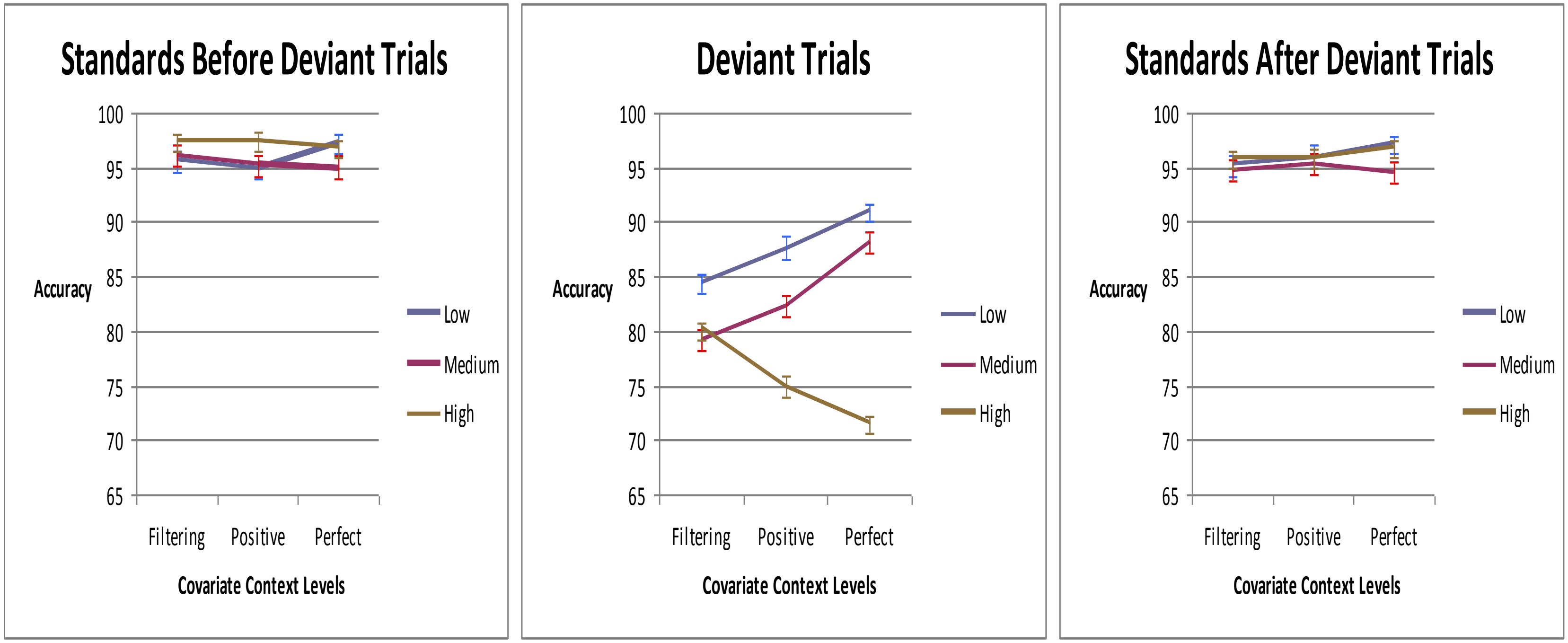
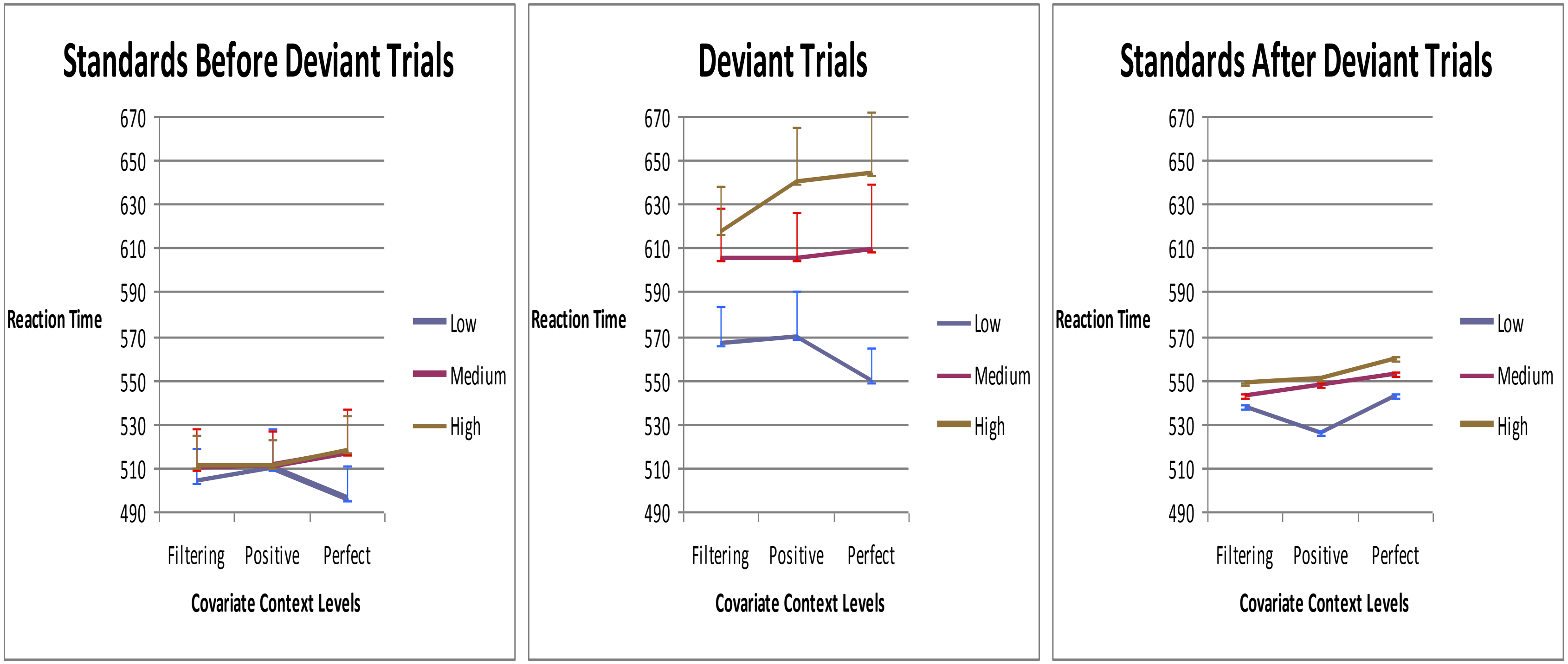


Figure 3. Average Accuracy Rates of Standards Before Deviants, Deviant and Standards After Deviants across all conditions in the present study.



Standards Before Deviant Trials

Deviant Trials

Standards After Deviant Trials

Figure 4. Average Reaction Time of Standards before Deviants, Deviant and Standards After Deviants across all conditions in the present study.

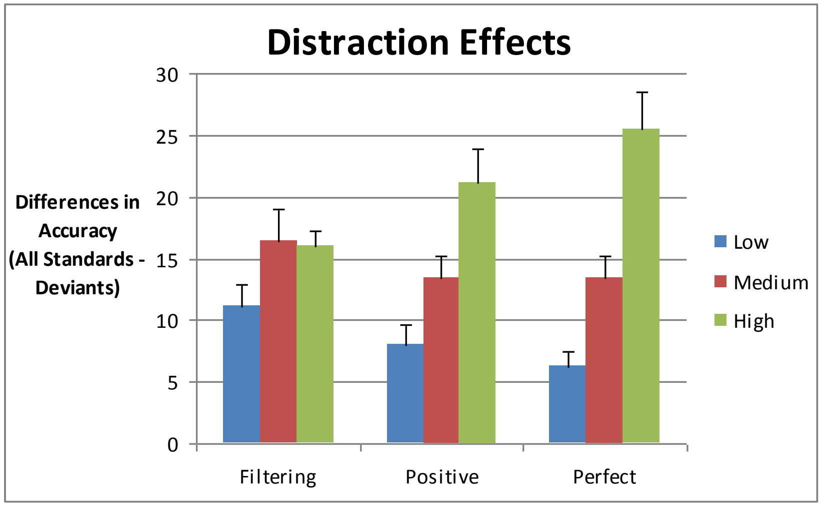


Figure 5a. Differences in accuracy rates between standard and deviant trials across conditions.

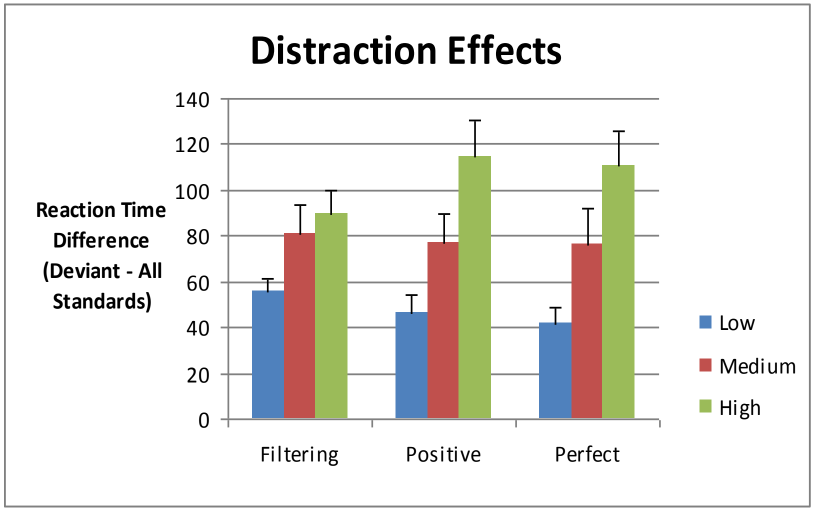


Figure 5b. Differences in accuracy rates between standard and deviant trials across conditions.

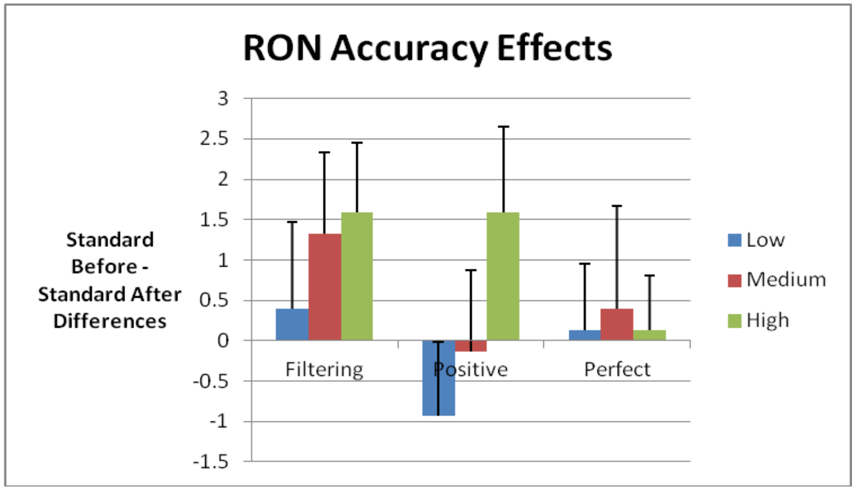


Figure 6a. Accuracy comparison of standards presented before deviants with those presented after deviants

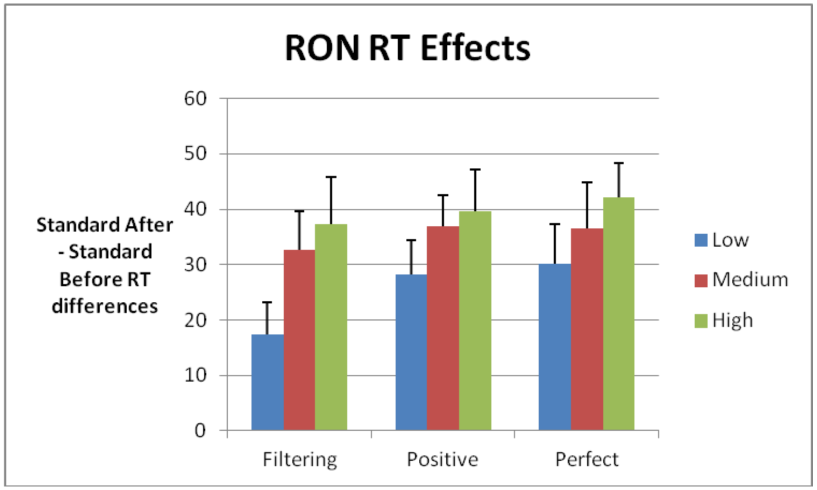
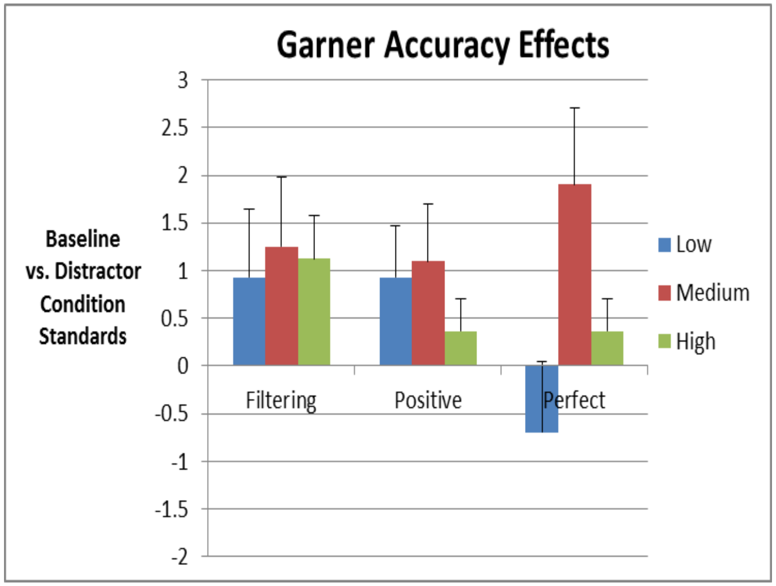


Figure 6b. Reaction time comparison of standards presented after deviants with those presented before deviants



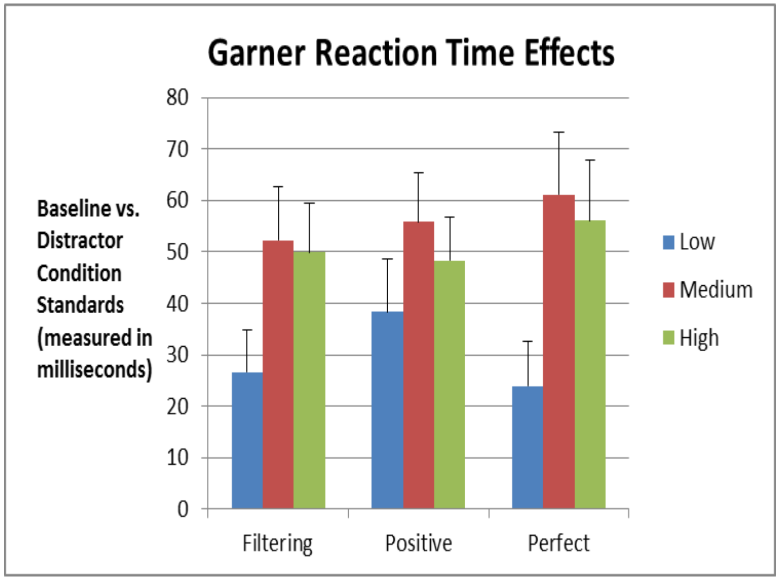


Figure 7a. Accuracy of standards across conditions compared to Baseline standards

Figure 7b. Reaction Times of standards across conditions compared to Baseline

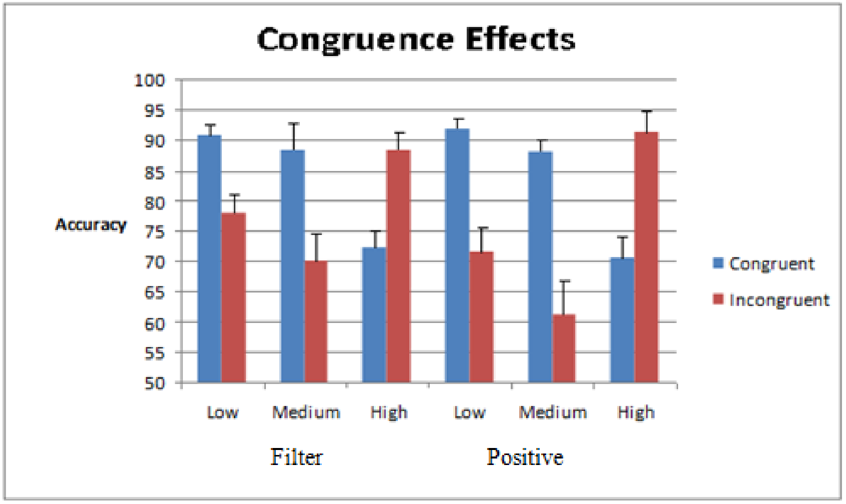


Figure 8. Mean Accuracy of Congruent and Incongruent Trials Across Filtering and Positive Conditions.

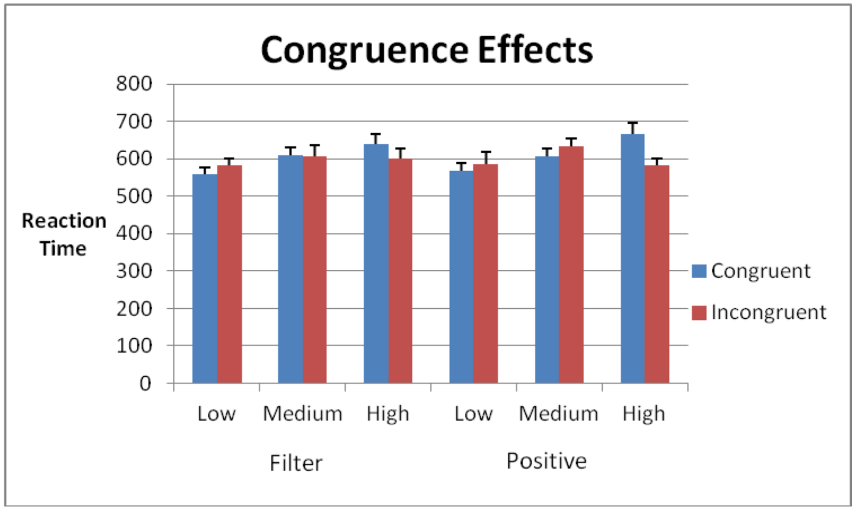


Figure 8. Mean Reaction Time of Congruent and Incongruent Trials Across Filtering and Positive Conditions.

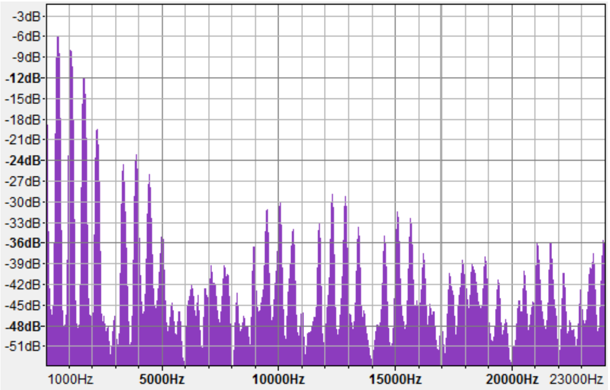


Figure 9a. Fourier Analysis of High-Imbalance relatively higher deviant trial (560 HZ 20% duty cycle)

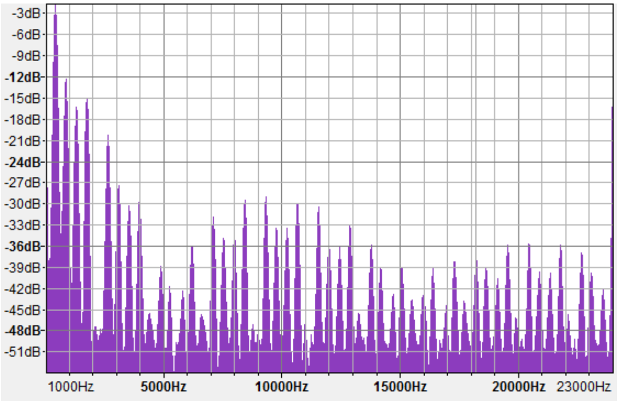


Figure 9b. Fourier Analysis of High-Imbalance relatively lower deviant trial (445 Hz 40% duty cycle)