Fall 2015

# Detecting changes in auditory events

Rachael B. Peck
*James Madison University*

Detecting changes in auditory events

Rachael B. Peck

A thesis submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Arts

Psychological Sciences

November 2015

---

FACULTY COMMITTEE:

Committee Chair : Michael D. Hall, Ph.D.

Committee Members/ Readers:

Jeff Dyche, Ph.D.

Vivien (Kit Ying) Chan, Ph.D.

# Acknowledgements

I would like to take this opportunity to express my deepest gratitude to my thesis committee, Dr. Michael D. Hall, Dr. Jeff Dyche, and Dr. Vivien Chan, for their willingness to assist and advice throughout the writing process. Most of all, I am indebted to Michael D. Hall, my advisor, for his expertise, wisdom and patience and for pushing me beyond where I thought I could go.

Table of Contents

List of Figures

Abstract

Change deafness is defined as the failure to detect the source of an above-threshold change in an auditory scene (Gregg & Samuel, 2008; Hall, Peck, Gaston, & Dickerson, 2015). A new paradigm demonstrated the phenomenon under analogous conditions to its visual counterpart (Hall, et al., 2015). This investigation examined the use of the paradigm through two experiments which involved the same four simultaneously presented events. Experiment 1 distributed events across a virtual 120º on the azimuth while the target event oscillated across a 60º space for each trial. Listeners were instructed to identify the target. Target rate of change was manipulated across four different velocities (80º/*s*, 40º/*s*, 24º/*s*, 8º/*s*). Results confirmed that all conditions differed in error rates from an isolated control task. The 8º/*s* condition displayed the highest error rates, providing strong evidence of change deafness, whereas error rates in the 80º/*s*, 40º/*s*, and 24º/*s* conditions did not significantly differ, providing inconclusive evidence. Response times did not vary across conditions. Experiment 2 compared findings to a filter manipulation and evaluated change deafness by comparing flickered (1*s* and 3*s* initial presentation) and continuously changing target events, which oscillated between wide- and narrow-band filters. All conditions resulted in error rates that did not vary from the control task. The continuous condition had increased response times, providing explicit evidence of change deafness. Rapid response times in flicker conditions indicated the elimination of change deafness. The 3*s* presentation time in one flicker condition further reduced response times, demonstrating the impact of encoding. Experiments support the assessed paradigm as an appropriate method of analyzing change deafness.

Detecting changes in auditory events

Human beings live in a dynamically changing world, and while it may be known that a change in the environment is occurring, humans often fail to detect the source of that change. These situations have been predominantly studied in the visual modality, however, comparable demonstrations in other modalities have recently been investigated, including in audition (Eramudugolla, Irvine, McAnally, Martin, Mattingley, 2005; Gregg & Samuel, 2008; Hall, Peck, Gaston, & Dickerson, 2015) and with tactile changes (Gallace, Auvray, Tan, & Spence, 2006). While the current investigation focuses on factors that affect corresponding auditory conditions, it is still necessary to review the visual change detection (e.g., change blindness) findings that the auditory change detection research is rooted in.

Change blindness occurs when the source of relatively large above threshold changes in a retinal display go undetected (Rensink, 2002), and there are several ways in which the phenomenon has been demonstrated. Traditionally, experiments involve a visual image composed of several objects (e.g. a dinner scene) that switches between complete and modified versions of itself. Each presentation is interrupted by a brief gray screen (typically lasting 250-500*ms*) which serves to mask the iconic memory of the previous image. The modified version of the image usually has one object (e.g., a chair) change locations, disappear entirely (e.g., Blackmore, Brelstaff, Nelson, & Troscianko, 1995; Rensink, O'Regan, & Clark, 1997), or change very gradually over time (e.g., Simons, Franconeri, Reimer, 2000). In a well-known study conducted by Rensink, et al. (1997), participants identified the source of a change within an image that flickered between an original (e.g., an airplane) and modified image (e.g., the same airplane, now missing an engine). Participants took an unusually long time (averaging nearly one minute) to detect the change despite the fact that they were instructed to find it and were actively searching for it.

This surprising occurrence has been generally argued to reflect the need for feedback between specific cortical pathways (Beck, Rees, Frith, & Lavie, 2001; Kanwisher, 2001). Visual pathways are separated into two parallel tracts which process information regarding the identity of objects and their spatial locations. The ventral (“what”) and the dorsal (“where”) pathways originate in the occipital lobe and terminate in the occipitotemporal and occipitoparietal lobes, respectively. The dorsal parietal regions are involved in the spatial orienting of visual attention, while the ventral temporal stream is associated with the activation of category-selective areas (Beck, et al., 2001).

Successful visual change detection activates both the dorsal and ventral pathways whereas only the ventral pathway is activated when the change goes undetected. Consequently, the dorsal pathway is predominantly associated with visual awareness while the isolated ventral pathway could be insufficient to produce the attention necessary for successful change detection (Beck, et al, 2001; Lumer, Friston, & Rees, 1998; Rees, et al., 2002; Robertson, Treisman, Friedman-Hill, & Grabowecky, 1997; Ungerleider & Haxby, 1994). The dorsal pathway will not be activated if transient motion at the point of the change is disrupted. For example, the previously mentioned flicker paradigm creates an instant change in a visual image which eliminates the transient motion. This results in the delocalization of the target object, making the change more difficult to detect (Beck, et al, 2001; Rensink, 1997; Kanwisher, 2001).

Due to the nature of change blindness, it is sometimes confused with another phenomenon known as inattentional blindness. In controlled experiments, inattentional blindness is the failure to detect an above-threshold visual change because one's attention is diverted due to experimental instructions. There is also evidence of "real-world" inattentional blindness, resulting from individuals attending to a specific event (e.g., giving directions to a confederate), and thereby not detecting a larger change (a change in the confederate) (see Simons & Levin, 1998).

In a typical inattentional blindness study, an experimenter will instruct participants to attend to a specific distracter, and while the participant's attention is focused elsewhere, an above-threshold change occurs. For example, Simons & Chabris' (1999) classic study involves a video of several people (wearing black or white shirts) passing a basketball, and participants are instructed to count the number of passes that the white-shirted team makes. After several seconds, the video displays a man in a dark gorilla costume walking between the basketball players and beating his chest. Surprisingly, 46% of people failed to detect this clearly above-threshold change, even though "the gorilla" is in the video for nine seconds. (For other accounts of inattentional blindness see Levin & Simons, 1997; Mack & Rock, 1998; Mack, Tang, Tuma, & Kahn, 1992; Simons & Levin, 1997; Sinnett, Costa, Soto-Faraco, 2006.)

While change blindness and inattentional blindness demonstrations both focus on undetected above-threshold changes, the methodology between the two is markedly different. Change blindness studies explicitly instruct participants to search for a change in the scene, while inattentional blindness studies do not inform the participants whatsoever of an upcoming change. Rather, they explicitly require that attention be directed to a subset of objects or events within the visual scene. While attentional constraints also apply to change blindness, participants in these studies are not instructed on how to direct their attention. Given that attention is of limited capacity (see Beatty, & Pollack, 1967; Kahneman, 1970; Kahneman, 1973; Kahneman), inattentional blindness studies reduce or even eliminate the attentional resources available which would otherwise be devoted to remaining events within the scene. The two phenomena reflect different demonstrations of overlapping, yet somewhat different processes due to the amount of attention devoted to detecting a change, and they should not be referred to interchangeably.

**Change Deafness vs. Inattentional Deafness**

This failure to detect seemingly obvious changes is surprising, however, it is not unique to vision. In traditional change *deafness* demonstrations, an auditory array is presented to participants for a brief amount of time. The array is then replaced by a momentary inter-stimulus interval before a modified version of the array is presented (i.e., an event is now added or missing from the array). Similar to change blindness, participants are aware that a change will occur and are actively searching for it. While there is significantly less literature on change deafness than change blindness, the past decade has shown a dramatic increase in the number of investigations conducted. These investigations include demonstrations with environmental sounds (Eramudugolla, et al., 2005; Gregg & Samuel, 2008; Hall, et al., 2015), as well as with music (Agres and Krumhansl, 2008).

Just as change blindness and inattentional blindness have been described interchangeably, change deafness and inattentional deafness suffer from the same misconception. Despite the different task conditions and attentional focus involved in each demonstration, researchers continue to label both phenomena as “change deafness”. One such demonstration involved participants who were instructed to listen to a recording of an individual speaking and repeat the words spoken (Vitevitch, 2003). These words differed in lexical difficulty and demanded the listener’s attention. While participants were repeating the words, the original speaker was sometimes substituted for another speaker halfway through the list. Over 40% of participants failed to detect this change. While this study successfully demonstrated listeners’ failure to detect a very large auditory change, it did so by explicitly instructing participants to attend to a separate task which distracted them from the changed event (for other demonstrations see Cherry, 1957; Fenn, et al., 2011; Neuhoff, Schott, & Kropf, 2014; Neuhoff, Wayand, Ndiaye, & Berkow, 2015).

It is clear that both change deafness and inattentional deafness demonstrations apply attentional constraints; however, in change deafness studies, participants are not constrained on how to direct their attention. As a result, the current investigation defines change deafness as the inability to detect the source of an above threshold auditory change (while attempting to determine the change) and inattentional deafness as the inability to detect an above threshold auditory change due to experimental instructions directing one’s attention elsewhere.

## Further Difficulties with Change Deafness

A second problem in change deafness is that many experimenters have not identified specific factors that separate change deafness from other lower level processing issues. Currently, many reports of the phenomenon are confounded with encoding failures, and it is impossible to determine which experiments provide evidence of change deafness and which require a more parsimonious explanation. Prior to attributing a detection failure to change deafness, experimenters must maintain that listeners have properly encoded the stimulus array by systematically eliminating any ambiguity.

There have been several recent attempts to investigate the encoding issue in change detection experiments (Gregg & Samuel, 2008; Hall, et al., 2015; McAnally, Martin, Eramudugolla, Stuart, Irvine, Mattingley, 2010; Pavani & Turatto, 2008; Snyder & Gregg, 2011;). Eramudugolla, et al. (2005) provided participants with five seconds to listen to a stimulus array prior to any change being made, thereby affording participants ample time to encode the events presented. The study utilized a method known as the one-shot paradigm, as is typical of most studies claiming to demonstrate change deafness. The one-shot paradigm permits each array to be presented only once, and the latter array presented to participants is either identical to the first or was modified (e.g., one event removed from the array or two events changed locations). In Eramudugolla’s (2005) study, Array-1 might be comprised of the following stimuli: birds chirping, a cello playing, Gregorian chant, and a trumpet solo played simultaneously whereas Array-2 might be comprised of only birds chirping, a cello playing, and Gregorian chant. Participants made a verbal judgment as to whether the scenes were identical or different. Results indicated that participants were highly successful in detecting changes in arrays comprised of four or six events (e.g., ceiling level performance for the majority of arrays), suggesting little to no evidence of change deafness. This is likely due to the extended access participants had with the array, allowing for time to properly encode the stimuli.

Other researchers attempted to determine whether stimulus arrays (comprised of their own variety of events) have been encoded by specifically asking event identification questions. For example, after a stimulus array was presented, Gregg & Samuel (2008) randomly asked participants either a change detection question ("Were the arrays identical or different?") or an event encoding question ("Was the dog barking present in the array?"). Their study found very low error rates for encoding and very high error rates for change detection. Consequently they asserted that the study's change detection failures represented change deafness.

The study involved two separate tasks which likely elicited different strategies from the participants. The change detection ("same-different") task could involve either the identification of separate events or the detection of broad spectral changes. The event encoding task, on the other hand, explicitly required only the identification of events. It is probable that listeners dedicated their attention to only one of the two tasks (i.e., attending only to event identification rather than listening for a change). To account for this confound, Gregg & Samuel (2008) separated the two tasks into a between-subjects manipulation where half the participants were asked each question. This second experiment still did not eliminate the separate task strategies, and therefore, it was not possible to determine whether the successful encoding co-occurred with the change detection failures.

There are several factors which impact the likelihood that stimuli will be properly encoded. First, the perceptual saliency of the each presented stimulus is crucial to recognizing it as an isolated event. The saliency of these events should result in exogenous orienting, in which a listener's attention is automatically directed to the location of the event (Spence & Driver, 1994). Stimuli should also be spectrally dissimilar to eliminate the possibility of masking influences. It will be substantially more difficult to detect changes between the events which share a degree of sensory or perceptual overlap (e.g., see Dickerson and Gaston, 2014). Only events which demand the listener's attention and that can be individually identified will be properly encoded, therefore, a failure to detect changes within subtle, ignored, or masked stimuli could be argued not to be change deafness, but rather might simply indicate an encoding failure.

A second issue which could influence the ability of participants to encode a stimulus array is the number of events presented in that array. It is possible that certain change detection failures could be predicted solely by an excessively complicated scene, and it is likely that the listener's attentional limits have simply been exceeded. For example, Eramudugolla, et al. (2005) found that participants were able to encode a small number of events (four or six); however, when too many events were presented (i.e., eight), performance suffered. Therefore, the failure to detect changes in the eight-event trials did not reflect change deafness, but rather, revealed an encoding limitation.

A final issue that impacts encoding is the paradigm utilized in presenting the stimulus array. There are several negative performance-based consequences that result from using the one-shot paradigm. First, the one-shot paradigm allows the presentation of only two arrays. The arrays reflect a single change, and a failure to detect that change could occur for a number of reasons, only one of which is change deafness. A second issue is that the one-shot paradigm typically involves presenting participants with entirely different scenes on each trial. Greater stimulus variability increases uncertainty and results in a higher probability of informational masking (i.e., an increase in threshold due to stimulus uncertainty) and may negatively impact listener performance (Dickerson & Gaston, 2014; Durlach, Mason, Kidd, Arbogast, Colburn, Shinn-Cunninghamm, 2003; Leek, Brown, & Dorman, 1991). Both Eramudugolla (2005) and Gregg & Samuel (2008) provided participants with altered scene sizes (ranging from four to eight events per scene) which were randomly chosen from a large library of events (e.g., selecting from 11 and 18 events for Eramudugolla, et al., 2005 and Gregg & Samuel, 2008, respectively). Scene compositions varied with each trial, and the large number of possible events likely generated uncertainty regarding which events would be presented.

These two concerns are largely the result of not relying on analogous methods to change *blindness*, which has traditionally employed the flicker paradigm (though the one-shot technique has been used; see Levin & Simons, 1997). The flicker paradigm decreases uncertainty and information masking by allowing participants sufficient time with the stimuli in order to properly encode the array (though potentially only after several iterations of the array). Furthermore, the flicker paradigm promotes the use of both response time and error rates as dependent measures. This not only provides information regarding the probability of errors but also allows for a determination of how long it takes successful listeners to detect the change. While it is not impossible to measure response time in the one-shot paradigm (both Backer & Alain, 2012 and Constantino, Pinggera, Paranamana, Kashino, & Chait, 2012 used this measure), it is more applicable to the flicker paradigm.

It is important to note that despite the large library of possible stimuli and the use of the one-shot paradigm, Eramudugolla, et al. (2005) found little to no evidence of change deafness when participants were presented with four to six events and ample time to fully encode (five seconds). Gregg and Samuel (2008), on the other hand, created a more challenging task by granting participants only one second to encode each scene, regardless of scene size, resulting in evidence of poorer change detection. The high levels of uncertainty and limited time to properly encode likely resulted in the failure to acquire and compare information concerning event identity.

One recent study has attempted to address these encoding and methodological issues. Hall, et al. (2015) used the flicker paradigm in order to assess change deafness in an analogous fashion to change blindness while examining the phenomena under conditions that minimized encoding concerns. This was accomplished by presenting the same four stimuli throughout the entire experiment which essentially removed concerns about sufficient event encoding because the listeners knew precisely what events would be presented. Furthermore, uncertainty was minimized because listeners knew that there would be a change in each trial (as opposed to the same-different task).

The study manipulated the four stimuli to determine the influence of types of change (abrupt versus continuous) on change detection. Three conditions utilized the flicker paradigm in which the stimulus array was presented for a one second long presentation, 500*ms* of white noise or silences (two different conditions), followed by another one second presentation that was a modified version of the first presentation (i.e., one event changed positions). These original and modified arrays were presented in alternation to the listener for over 20 seconds. The third flicker condition was identical to the white noise condition, with the exception of an extended initial presentation (three seconds) for the first stimulus array. This condition was utilized to increase the probability that events were fully encoded prior to any change occurring. A final condition did not utilize the flicker paradigm, but rather employed continuous changes. The array was presented for an uninterrupted 15*s*, while the target event oscillated gradually between two positions at an above threshold velocity.

All three flicker conditions resulted in significantly shorter response times and lower error rates than the continuous condition. Furthermore, response times were substantially shorter in the condition that provided the extended encoding time. In fact, participants responded within 1.5 cycles of presentation while error rates remained around 10% (through error rates were functionally less when considering the study's a priori error rates for reporting the location of isolated events). These results for abruptly changing stimuli indicate the relative ease at which participants completed the task once uncertainty was eliminated and sufficient encoding time was provided.

The manipulation with the continuously changing stimulus, on the other hand, provided a possible demonstration of change deafness. Interestingly, mean response times were as high as 9.6*s* and mean error rates remained around 27%. This is especially surprising as the changes that occurred were above threshold for detection. However, what is of greatest interest is that the continuously changing condition provided the strongest evidence of change deafness, whereas under visual conditions, this should have resulted in little to no evidence of change blindness (Yantis, 1992).

The authors offered two potential explanations for the results of the continuous manipulation.  First, the magnitude of change between any two perceptually adjacent positions in time is smaller than the change in the flicker condition. As Grantham (1997) discusses through his "snapshot hypothesis", rather than motion being detected directly, the rate of change of a moving auditory target is inferred by local time comparisons of the target's different spatial positions. These local time comparisons will henceforth be referred to as "snapshots" throughout the paper. As such, participants should be able to detect the change only if they obtain snapshots of the stimulus just before and after a substantial change, and the detection of abruptly changing stimulus arrays should be easier to detect due increased spatial distance between stimulus presentations.

A second possible explanation for change deafness in the continuous condition stems from the rate of change utilized in the experiment. If the rate of change is slow enough, the obtained snapshots could be near-identical, and listeners would be unable to discriminate between them. The two conditions (flicker and continuous) did not reflect the same rate of target movement; the continuous condition completed one full cycle (from the starting location to distant location and back) in five seconds, whereas the flicker conditions utilized three seconds. While the continuous rate of change was surprisingly fast (i.e. 24º/*s*) and should have been detectable, it is possible, though unlikely, that the poor performance was simply the result of distinctly slower changes than the other conditions being assessed. These issues regarding the rate of change for continuous manipulations will be addressed in the current investigation.

## Goals of the Current Investigation

Two experiments were conducted in order to evaluate parameters that might predict the occurrence of change deafness. Both experiments served as conceptual replications and extensions of the Hall, et al. (2015) study. Experiment 1 sought to explore the possibility that a listener's ability to detect a continuous, spatially moving stimulus is influenced by the rate of change for that stimulus. Change detection should be based upon the auditory snapshots obtained by the listener as they perceptually sample the stimulus, and it is believed that these snapshots should be directly impacted by the rate of change. Faster moving events should result in greater spatial distance between subsequent snapshots, whereas the snapshots obtained for slower moving events should be positioned closer together. Experiment 1 explicitly manipulated the rate of change for continuously moving events at four possible velocities. Two of these rates of change were used to directly compare the velocity in Hall, et al.'s (2015) continuous manipulation with the velocity equated for their flicker manipulation. In doing so, the current investigation will provide further insight into possible explanations for the increased response time of their continuous manipulation.

Experiment 2 addressed questions regarding what dimension is the most appropriate analogue between visual and auditory processing. It conceptually replicated the Hall, et al. (2015) study with frequency-based changes rather than spatial movements in order to assess auditory changes in the modality's dominant dimension. While it was expected that the current investigation would vary quantitatively from the previous investigation, the two studies were hypothesized to produce conceptually similar results.

## Experiment 1

The first experiment sought to assess whether rate of change is critical to the observance of change deafness. To investigate this, the rate of spatial change in a continuously oscillating target stimulus was manipulated in order to determine a point where response times and error rates significantly changed. The auditory system may utilize either a snapshot procedure to detect rapidly moving events or a motion-sensitive mechanism (essentially a "multi-snapshot" mechanism) for movements of slower velocity (Grantham, 1986; Grantham, 1997). Slowly moving stimuli are perceptually sampled more frequently due to the "velocity effect," in which the minimal audible movement angle (MAMA)—or the just noticeable difference (jnd) on the azimuth—increases approximately linearly with velocity (Chandler & Grantham, 1992; Perrott & Musicant, 1977). Quickly moving stimuli, on the other hand, may only be perceptually sampled at the onset and offsets of movement, rather than being continuously monitored. As a result, the MAMA is quite large, and depending upon the precise velocity, it can be upwards of 10º (see Chandler & Grantham, 1992; Grantham, 1986; Harris & Sergeant, 1971).

The current investigation manipulated the continuous spatial movement of one auditory stimulus (the target event) in an array of four events, and participants were asked to identify the target. The target oscillated at one of four velocities (8º/*s*, 24º/*s*, 40º/*s*, and 80º/*s*). It was predicted that there would be less evidence of change deafness for faster moving targets (i.e., at 40º/*s* and 80º/*s*) due to the listener collecting and comparing discrete adjacent snapshots separated by a substantial spatial distance. This increased distance between adjacent snapshots was predicted to result in better detection and would thus result in faster response times or fewer errors. In contrast, the slower moving events (i.e., 24º/*s* and 8º/*s)* would be perceptually sampled at more proximate positions and result in less spatial distance between adjacent snapshots. This was predicted to result in higher error rates and/or slower response times, displaying evidence consistent with change deafness. The 24º/*s* condition mimicked the rate of change utilized in the Hall, et al. (2015) study and served as a control condition. It was expected to consistently show moderate levels of change deafness, as was displayed in the previous study. The speed which would induce maximized evidence of change deafness was not known, therefore, the slowest rate of change possible that still resulted in one full cycle of movement was used (i.e., 8º/*s*).

**Method**

**Participants.** All participants, across both studies, were students from James Madison University undergraduate psychology courses, with the exception of the author. Participation in the experiment served as partial fulfillment of course requirements. Listeners were required to be between the age of 18 and 40 to reduce the possibility of the impact of presbycusis. Also, all were required to understand spoken and written English, as the instructions were provided in English. Participants were free of any known hearing deficits at the time of participation, as indicated by self-report.

In Experiment 1, participants were asked to identify a target sound based on a spatial manipulation. As a result, an a priori performance criterion was established that required listeners to reliably state which sound (the first or the second) was presented at a specified region in acoustic space. This isolated control task had stringent requirements since the events were positioned at distances that were above threshold for detection and the differences in position were designed to be easily detected by listeners. Therefore, data analyses were restricted to the 16 participants (of 17 total) who were able to meet the a priori criteria with a minimum of 80 percent accuracy.

**Stimuli.** Certain characteristics were consistent across stimuli in both experiments. Initial waveforms representing potential target and distracter events (i.e., prior to any additional stimulus manipulations) had been provided by the Army Research Laboratory (ARL) in Aberdeen, Maryland, who collected approximately 20 one-second .wav environmental sounds with a 44.1 kHz sampling rate (16-bit resolution) from the public website www.freesound.org (Music Technology Research Group). ARL evaluated perceptual similarity between all pairs of stimuli within the entire set.

Of the 20 sounds evaluated, four events were deemed to be the most distinguishable and were used throughout the experiments. They contained uniquely identifiable centroids that varied from one another on average by 861 Hz (ranging from 1,831-4,414 Hz). The sounds consisted of the following: (1) a bicycle bell (2) a dog barking (3) a helicopter, and (4) footsteps. One stimulus (the helicopter) initially displayed high amounts of within-stimulus frequency variability. Therefore, using Ableton Live (2014), a specified flat region of the stimulus was extracted and time-stretched from .983*s* to one second in order to match the lengths of the other three stimuli. All stimuli were equated for root mean square amplitude. At the time of presentation, stimuli were submitted to an anti-aliasing, four-pole, Butterworth/Bessel low-pass filter with an 11,000 Hz cut-off frequency and were delivered in a single-walled sound-attenuated chamber over Sennheiser HD 25-SP II headphones. Peak intensities did not exceed 80 dB(A).

All synthesized arrays were sequenced and rendered in Ableton Live 9.1.7 (2015) and Panorama 5 (Wave Arts, Inc., 2012). Panorama 5 is a virtual acoustics processor audio plug-in that is based upon Knowles Electronic Manikin for Acoustic Research (KEMAR) measurements (Gardner & Martin, 1995) and permits assignment of stimuli in a perceived spatial position and distance through manipulated localization cues of interaural phase and level as a function of frequency. Panorama 5 was used to process target movements as well. While starting locations and movements were not the precise positions (and are based on KEMAR rather than human measurements), they appeared to correspond closely to the intended values and intended movements which were successfully distinguished in Hall, et al. (2015). Arrays for Experiment 1 were composed of the four sounds played simultaneously and randomly distributed across four intended starting positions of -60º, -20º, +20º, and +60º around the azimuth at an intended distance of 20*m* in a polar coordinate system.

Figure 1 displays the starting positions and movements utilized in array presentations throughout Experiment 1. Each array was characterized by the movement of one sound while the remaining three sounds were static. Stimulus movements were manipulated by four rates of change, and the one randomly chosen target stimulus oscillated continuously at one of these rates. In one condition, the target stimulus moved one full cycle (from starting to distant position and back again) in 5*s* (i.e., at a rate of 24º/*s*). This rate of change was identical to the continuous condition used in Hall, et al. (2015) and served as a control condition. Other conditions included the target stimulus' completion of one full cycle in either 1.5*s* (i.e., moving 80º/*s*), 3*s* (40º/*s*), or 15*s* (8º/*s*), and will henceforth be distinguished by these rates of change as labels in the text. The target stimulus was intended to oscillate across 60º in all conditions (-60º to 0º; -20º to +40º; +20º to -40º; +60º to 0º), which well exceeds maximal thresholds (approximately 10º along the azimuth) despite position or velocity (Mills, 1957; Grantham, 1986). Furthermore, the resulting maximum distance travelled by each target stimulus was still 20º distant from adjacent events. There were 24 randomized arrays with equal probability of each target event, starting position, and movement. Each sequenced array presentation lasted for a total of 15*s*.

**Procedure**. The principle task of Experiment 1 was consistent with the Hall, et al. (2015) study previously discussed. Experiment 1, as with Experiment 2, utilized the E-Prime v2.0 (SP1; Psychology Software Tools, Inc., 2012) experiment development platform for control of timing, presentation of stimuli, and collection of responses. It was conducted in a blocked fashion such that each condition was delivered separately. Blocks of trials were counterbalanced through a Balanced Latin Square. A brief rest break was provided after each block of trials for both experiments.

Following consent, participants first completed the isolated control task. This permitted the listeners to become familiarized with the stimuli while additionally allowing experimenters to determine whether listeners were able to accurately localize stimuli at the relative target positions manipulated in the experiments. In this assessment, listeners identified which of two examples of the same potential target sound was presented farther to the right (*1* = first sound, *2* = second sound), reflecting the same potential positions during array presentations. All four events were individually presented in random order for each spatial position (i.e., comparison of -60º v. 0º, 0º v. -60º, -20º v. +40º, +40º v. -20º, +20º v. -40º, -40º v. +20º, +60º v. 0º, and 0º v. 60º). Each of the four events were presented in each of the eight relative positions five times for a total of 160 trials.

The isolated control task ended with a brief rest break and was followed by the first block of trials. On a given trial, listeners were presented with a randomly selected stimulus array with one target event oscillating between two positions. The listener was instructed to identify which of the four events constituted the moving target event. Participants were instructed (both verbally and through displayed instructions) to respond immediately should the trial end, and if no response was collected, the subsequent trial would begin after two seconds. Listeners identified the target stimulus (*1* = dog, *2* = helicopter, *3* = footsteps, *4* = bell) by pressing the corresponding button on a serial response box. Response times were measured in milliseconds. An inter-trial interval of 500 *ms* was used after each response and prior to the start of subsequent trials. Each block of trials consisted of stimulus presentations in which each event was presented in each position and utilized each movement. This resulted in 24 trials per block of trials. There were four blocks (one per rate condition) throughout the experiment. The entire experiment, from consent to debriefing, lasted approximately 60 minutes.

It was hypothesized that the fastest velocity (80º/*s*) would produce the least evidence of change deafness due to the larger spatial shifts between adjacent time intervals. Predictions for the 40º/*s* were multi-faceted. If the rate of change alone is critical to the observance of change deafness then the 40º/*s* condition should result in little evidence of change deafness due to the spatial changes between adjacent time intervals. In contrast, if the continuous method of presentation is critical to the observance of change deafness, then the 40º/*s* condition should produce some evidence of change deafness, as indicated by high error rates and long response times. The question of which factor—rate or manipulation—is critical to change deafness was left as an open question for the research to evaluate.

A secondary motivation for the 40º/*s* condition was to use it as a comparison against the original flicker condition in the Hall, et al. (2015) study. The rates of change between the two studies were equal in completing one full cycle (starting location to distant location and back) in three seconds. Therefore, the results of the 40º/*s* condition would provide evidence as to which factor, rate of change or method of presentation, was critical to the observance of change deafness. If the continuous and flicker conditions were approximately equal in response times and error rates, then it would be likely that rate of change is of critical importance to change deafness. If, on the other hand, the continuous condition accrued longer response times and/or higher error rates, it would be clear that the continuous method of presentation itself is of critical importance.

In contrast to the 8º/*s* and 40º/*s* conditions, the relative change in position between participants' adjacent snapshots in the slower conditions (i.e., 24º/*s* and 8º/*s*) was predicted to result in less disparity between adjacent perceptual samples. As a result, the 8º/*s* condition was predicted to display the highest probability of error and longest response times, providing the greatest evidence of change deafness. The 24º/*s* condition was identical to the continuous condition used in the Hall, et al. (2015) experiment and was predicted to elicit slightly more discrete snapshots than the 8º/*s* condition, resulting in better change detection. As a result, it was predicted to show lower error rates and faster response times than the 8º/*s* condition, while indicating some evidence of change deafness by producing much higher error rates and slower response times than both the 40º/s and 80º/s conditions.

**Results and Discussion**

Median response times (for correct answers) and error rates from the individual listeners were calculated separately for each rate condition. The decision to restrict analyzed response times to correct responses was based upon the fact that response times for incorrect answers have been demonstrated to be slower than correct answers (Ratcliff

& Rouder, 1998). Response times for correct answers that exceeded three standard deviations above and below of the grand mean for each participant were eliminated.

Finally, a participant could only detect a difference in the array after the target event's change in position had at least exceeded the minimum audible movement angle (MAMA). The MAMA increases approximately linearly with velocity (Chandler & Grantham, 1992; Grantham, 1985; Grantham, 1997; Perrott & Musicant, 1977), so each condition required a different amount of time before listeners could detect the change, with the faster conditions requiring the longest time. Therefore, to determine how much of an observed difference might actually be the result of change deafness, the time necessary to cross the MAMA needed to be removed.

The amount of time required for the target event to move 12° was extracted from each listener's median response time for each condition. Specifically, 0.15*s*, 0.30*s*, 0.50*s*, and 1.50*s* was respectively removed from the 80º/*s*, 40º/*s*, 24º/*s*, 8º/*s* conditions. This calculation was based off of Grantham's (1997) results which showed a 4.8º MAMA for 20º/*s* and a 7.8º MAMA for 60º/*s*. A MAMA of 12º served as a close estimation for an 80º/*s* velocity. The MAMA for the 80º/*s* condition was removed from all conditions in order to reduce the likelihood of a type I error, though this adjustment actually minimized the chance of demonstrating change deafness in slower conditions.

Several measures were consistent throughout both experiments. First, a measure of sensitivity (*d*') was utilized to determine how well participants were able to discriminate between the presented sounds in the isolated control tasks. A dependent samples t-test was also utilized to compare the error rates for the control task to the mean error rates of each condition for both experiments. Finally, Bonferonni pair-wise comparisons of means were used in all subsequently discussed analyses to decrease the rate of type I error (Armstrong, 2014). Likewise, Greenhouse-Geisser corrections were applied when sphericity assumptions were violated.

A series of corresponding, yet separate, analyses were conducted for both mean response times and mean probabilities of error. To assess the relationship between rate of change and the mean response times/error rates, as well as to evaluate the impact of each event-type and starting position, the response time/error rates for each event and position were analyzed. Response times/error rates were submitted to two separate 4x4 repeated measures ANOVAs with condition (80º/s, 40º/s, 24º/s, and 8º/s) and either event (bell, dog, helicopter, footsteps) or starting positions (-60°, -20°, +20°, +60°) as the factor. An additional analysis allowed a determination of whether a block order effect influenced the participants' failure to detect changes. The mean response times/error rates for order of presentation were submitted to a 4x4 repeated measures ANOVA with order of presentation (first, second, third, fourth) and condition as the factors.

Mean sensitivity was very high across participants for the isolated control task, *d'* = 3.04 (*SE* = 0.121), reflecting the participants' ability to discriminate between the different positions presented. A 4x4 repeated-measures ANOVA was conducted to determine whether the type of event (bell, dog, helicopter, footsteps) or target movement (-60º to 0º; -20º to +40º; +20º to -40º; +60º to 0º) impacted listener sensitivity. There were no significant main effects for either factor, $F(3, 45) = 1.915$, $p = .141$, $\eta_p^2 = .113$ (event) and $F < 1$ (movement). There was a marginal interaction between factors, $F(9, 135) = 1.813$, $p = .071$, $\eta_p^2 = .108$, such that participants were slightly more sensitive to the footsteps in -60º to 0º movement than in -20º to +40º movement. The overall high performance in the control task allows for assessment of potential change deafness and provides an additional comparison condition for the evaluation of error rate.

Figure 3 displays the mean probabilities of errors (panel A) and mean response times (panel B), along with corresponding standard errors within the control and experimental conditions. As can be seen in panel A, the average error rate for the isolated control task produced a mean error probability of $p = .026$. This error probability is representative of baseline performance and likely accounted for a very small percentage of the errors seen in the experimental conditions. This indicated that participants were highly sensitive to the change in position and could distinguish the manipulations of virtual position (for both directions) without difficulty, leading to ceiling or near-ceiling level performance. The failure to maintain this ability in the experimental conditions indicates the added difficulty in deciphering between the four events. While the overall increase in error rates could also potentially reflect masking influences, masking does not account for the differences across conditions. The more likely explanation is that the higher error rates provide initial evidence that a level of change deafness is occurring in each condition.  Mean rates for the isolated control task were visibly significantly reduced from the mean error rates of each experimental condition, $t(15) = 4.69$, $p < .001$; $t(15) = 5.22$, $p < .001$; $t(15) = 5.05$, $p < .001$; $t(15) = 7.80$, $p < .001$, respectively, for the 80º/*s*, 40º/*s*, 24º/*s*, 8º/*s* conditions.

As hypothesized, panel A further displays higher probability of error within the 8º/*s* condition than in the 40º/*s* or 24º/*s.* Surprisingly, the 80º/*s* condition also displayed a slightly higher error probability than the 40º/*s* and 24º/*s* conditions, which directly opposed the anticipated results. The differences between mean error rates contributed to a main effect of condition, $F(3, 45) = 4.322$, $p = .009$, $\eta_p^2 = .224$. Subsequent pair-wise comparisons of means provided additional evidence of these tendencies, revealing that errors in the 8º/*s* condition were significantly higher than in the 40º/*s* condition ($p = .003$) and errors in the 80º/*s* condition were marginally higher than in the 40º/*s* condition ($p = .075$). The 8º/*s* condition explicitly indicates change deafness whereas the 80º/*s*, 40º/*s*, and 24º/*s* are not conclusive. Each differs significantly from the isolated control task which minimally indicates the increased complexity of the experimental conditions, and could possibly also indicate some level of change deafness.

Mean response times (panel B of Figure 3) showed little difference across conditions despite the changes in velocity. The range of observed mean response times across the four conditions was quite small (7.92*s* to 8.85*s*). Statistical analyses confirmed that the average response time did not significantly differ across conditions, $F(3, 45) = 1.96$, $p = .137$, $\eta_p^2 = .140$. It could potentially be argued that the absence of differences was due to the transformation of data. To address this possibility, a similar analysis was conducted on the untransformed data and also failed to produce a significant main effect for condition, $F(3, 45) = 1.558$, $p = .217$, $\eta_p^2 = .115$, indicating that the the transformation of response times did not change the analysis.

The lack of variability of response time, when compared to error rates, provides evidence of a speed-accuracy trade-off in the 8º/*s* and 80º/*s* conditions. Both conditions display an inflated error rate while maintaining the same speed across conditions. This shows that while participants were able to sustain the same response time throughout the experiment, their accuracy suffered in the two extreme conditions (Ratcliff & Rouder, 1998). Some support for this interpretation comes from anecdotal reports from participants. For example, when participants were asked approximately how long they felt they had taken to respond, many indicated as much as 25*s*, despite the fact that each trial was no longer than 15*s*. Participants felt rushed to respond, despite the ample time provided. This potentially explains why response times stabilized across conditions and likely resulted in the increased error rates for the extreme velocities.

There were no interactions or main effects for order of presentation for either response times [$F(3,9) = 1.238$, $p = .352$, $\eta_p^2 = .292$] or error rate ($F < 1$). Similarly, there were no main effects of starting position (response time: $F(3, 45) = 2.058$, $p = .123$, $\eta_p^2 = .146$; error rate: $F < 1$). Target event did impact responses, however. Figure 4 displays the mean error rates (panel A), mean response times (panel B), and corresponding standard errors for each condition, as a function of target event. The figure shows no main effect of event for response time, $F(3, 45) = 1.467$, $p = .238$, $\eta_p^2 = .101$, but does indicate a difference in error rate, which contributed to a main effect of event-type, $F(1.87, 28.02) = 6.62$, $p = .005$, $\eta_p^2 = .306$. Pair-wise comparisons of means confirmed that participants produced significantly fewer errors for the helicopter than both the dog and the bell ($p < .001$ and $p = .013$, respectively). This is likely due to the discrete, quickly changing pattern contained within the dog and bell (i.e., individual “barks” and “rings” were apparent), whereas the helicopter produced a noise-based hum that was less difficult to detect. Even so, any inconsistencies in stimulus difficulty resulted no more than a probability of $p = 0.03$ difference in error rates. These errors rates across targets are exceptionally low, and given the lack of significant interactions, suggest that while target events played a small role in the detectability of specific trials, they did not form the basis for the more general differences displayed in the error rates across conditions.

In summary, Experiment 1 provided clear evidence of change deafness in the 8º/*s* condition through increased error rates, while also displaying a speed-accuracy trade-off which indicated that accuracy likely suffered due to the participants' desire to respond as quickly as possible. Response times for the 80º/*s*, 40º/*s*, and 24º/*s* conditions, when taken alone, appear to be very slow, but it is not conclusive as to whether it represents change deafness or not. Additional comparisons across data sets will be completed in the general discussion.

## Experiment 2

Researchers have suggested that the most appropriate auditory analogue to the visual space dimension might not be auditory space, but rather auditory frequency (Hirsh, 1952; Hirsh, 1959; Kubovy & Pomerantz, 1981). In contrast to the visual system's retinotopic structure, any sound that enters a listener's normally functioning cochlea will be subjected to a spectral analysis that will result in a frequency-to-place transformation (Traunmuller, 1990). The filters which make up the cochlea are tonotopically ordered, and the hair cells responding to the higher frequency levels are located at the base of the cochlea while those responding to lower frequency levels are located towards the apex. This organization is preserved throughout the auditory pathway and is shown in all subdivisions in the brainstem (i.e., the cochlear nucleus, superior olivary complex) as well as the midbrain and cortex (i.e., the central nucleus of the inferior colliculus and medial geniculate nucleus, as well as the primary auditory cortex in the superior temporal lobe) (Saenz & Langers, 2013; Reale & Imig, 1980). Kubovy & Van Valkenburg (2001) further argue that frequency and time are the two dominant dimensions in audition, or “indispensable attributes” (Kubovy, 1981), whereas space and time are dominant to the visual system.

Most change deafness evaluations have utilized space as a way to further segregate auditory events and have typically manipulated the addition, deletion, or switching positions of auditory events in a given array, rather than conducting a frequency-based change (Eramudugolla, et al. (2005); Gregg & Snyder, 2008; McAnally, et al., 2010; Pavani & Turatto, 2008). To my knowledge, a frequency-based manipulation has not been explicitly manipulated. Experiment 2 therefore manipulated this critical auditory dimension in a change detection task in order to permit a comparison between visual and auditory processing using the dominant dimensions from both modalities.

Experiment 2 explored potential change deafness for salient frequency-based changes based upon a filtering manipulation. It served as a conceptual replication of the Hall, et al. (2015) study which manipulated salient spatial changes throughout four conditions (three flicker conditions and one continuous condition) and was targeted at maximizing a participant’s ability to properly encode each stimulus array. The current investigation mimicked three of these manipulations. The first two conditions utilized the flicker paradigm which involved several presentations of an array that alternated between an original (wide-band) and a modified (narrow-band) version of itself. The modified array included a narrowly filtered version of the target event whose timbre had dramatically changed, yet the event’s source was still identifiable. One flicker manipulation presented each array for an equal amount of time (e.g., one second each) to determine whether an abrupt change evoked successful change detection. A second flicker condition included an extended initial presentation time in order to maximize encoding for the array. The final manipulation presented the stimulus array continuously for the entirety of each trial, while the target stimulus gradually oscillated between two timbres (wide- and narrow-band filtered versions). This will determine whether the level of spectral change across local time regions in auditory signal is critical to change detection. In a given trial, the target stimulus gradually changed from its original (wide-band) format to a narrowly filtered version of itself, and it continuously oscillated between these two versions at the same rate of change as the flicker conditions.

It was expected that the participants' responses to the manipulated spectral changes would be quantitatively different from the Hall, et al. (2015) experiment since listeners should be more successful at detecting changes in frequency than in spatial movements. However, it was hypothesized that conceptually, the two experiments would produce a similar pattern of results. Specifically, the continuous condition should result in more evidence of change deafness than the two flicker conditions due to limited time between a given participant's adjacent snapshots of the target stimulus. This limited time should result in the comparison of very similar auditory snapshots and result in longer response times and higher error rates. In contrast, it was predicted that both flicker conditions would produce little to no evidence of change deafness because of the comparison of two highly dissimilar adjacent snapshots, as evidenced by Hall, et al. (2015). While the extended initial presentation of one flicker condition was expected to maximize encoding of the array and therefore decrease response times, error rates between the two flicker conditions were not expected to differ.

**Method**

**Participants.** The requirements for participation were as described in Experiment 1. Specifically, 17 listeners (including the author) from James Madison University participated in the experiment. Because the experimental task required sensitivity to filtering, a preliminary task was used to ensure that participants were indeed sensitive to such manipulations in isolation. The manipulations were designed to be easily detected by the listeners, therefore, data analyses were restricted to the 15 listeners who met the criteria with a minimum of 80 percent accuracy.

**Stimuli.** Parameters for initial stimulus generation were as described in Experiment 1. The same four target/distracter events (a bell, a dog barking, a helicopter, and footsteps) were used throughout the experiment, and stimuli were equated for root mean square amplitude. The centroids will be used as the center frequencies for the respective filtering manipulations in order to maintain an identifiable timbre despite the significant loss of energy which will follow filtering. Centroids for each individual sound were identified using an algorithm provided within a LibXtract Vamp plugin (Jamie Bullock, 2012) for the Sonic Visualiser platform 2.3 (www.sonicvisualiser.org). Stimuli were then normalized for amplitude. Spectral movement was accomplished by narrowing the bandwidth of the second-order bandpass filter through a device created in Max for Live 6 (Cycling '74). Center frequencies were found to be 4,414 Hz (bell), 1,831 Hz (dog), 3,364 Hz (helicopter), and 2,870 Hz (footsteps). Bandwidths across events were equivalent in Mel-frequency as an attempt to match the perceived size of the filtering manipulation despite the differences in spectral centroids. These bandwidths for all wide-band and narrow-band conditions were 1,780 Mels and 367 Mels, respectively. These bandwidths in Hz were 10,106 Hz (wide-band) and 1,681 Hz (narrow-band), 5,000 Hz (wide-band) and 832 Hz (narrow-band), 8,031 Hz (wide-band) and 1,336 Hz (narrow-band), and 7,055 Hz (wide-band) and 1,174 Hz (narrow-band), respectively, for the bell, dog, helicopter, and footsteps.

Stimulus arrays composed of filtered stimuli were rendered in Ableton Live 9.1.7 (2015). They were composed of the four sounds presented simultaneously binaurally so that all were centered on the head. During presentation, all stimuli were initially set to the wide-band settings in order to present the most representative timbre for each event. A single target stimulus oscillated (either continuously or flickered) between the wide- and narrow-band filter versions. Filter changes in the continuous condition were manipulated through automation in Ableton Live. Gain (in dB) was systematically (linearly) adjusted as the bandwidth narrowed to account for any loss in average amplitude. Stimulus presentation for all conditions continued for 15*s*. However, because flicker conditions require inter-stimulus intervals between subsequent array presentations, the total trial time varied between 22*s* and 21*s.*

As discussed, there were three stimulus conditions in Experiment 2, namely, one continuous manipulation and two flicker conditions, and they will henceforth be referred to as the continuous condition, the 1*s* flicker condition, and the 3*s* flicker condition. Figure 1 graphically displays the continuous manipulation. In a given trial for the continuous condition, the target stimulus (i.e., the footsteps in Figure 1) completed a gradual linear change between the wide- and narrow-band filters, while the three non-target stimuli were presented through only the wide-band filter. Each complete cycle (wide- to narrow-band and back) was completed in three seconds. The trial lasted for a total of 15*s*.

Two flicker manipulations also were included to determine whether an abrupt change eliminated evidence of change deafness. Figure 2 graphically displays the conditions and the timing of each manipulation. The 1*s* flicker condition alternated between one second presentations of an original (all events through wide-band filter) and modified version (the target event through the narrow-band filter) of an array which were separated by 500*ms* inter-stimulus intervals and lasted for a total of 21*s*. As can be seen, the 1*s* flicker condition and 3*s* flicker condition were identical, with the exception of the initial array presentation. The 3*s* flicker condition utilized an extended three seconds of presentation in order to determine whether a longer initial exposure would result in better encoding and a reduced response time. Each trial in the 3*s* condition lasted for a total of 22*s*.

All three conditions included four separate arrays (one per target event), resulting in 12 arrays throughout the experiment. The experiment was conducted in a blocked fashion by condition. The order of blocks of trials were counterbalanced through a Latin Square. A brief rest break was provided after each block of trials. The presentation of conditions, including delivery of sound, were as described in Experiment 1.

**Procedure**. Prior to the experimental trials, participants were asked to complete a task in which they identified which of the two stimuli presented was the most heavily (narrow-band) filtered in order to determine whether participants were able to differentiate between the more and less filtered events. Listeners initially completed a brief familiarization task in order to become acquainted with the sound of the filtered events. They were presented with one-second segments of both versions of each event (bell, dog, helicopter, footsteps). Each stimulus event was presented in alternation five times for each filter, resulting in 10 presentations per event. Participants were notified in advance of the correct label and were instructed to listen to the stimuli at their own pace and to not make responses for the familiarization segment.

Following familiarization, listeners were instructed to complete a forced-choice identification task to assess how well they could identify a narrow-band filtered event. Two sounds were presented, and participants were instructed to state which was the most heavily filtered (*1* = first sound, *2* = second sound). Specifically, participants were asked to state which sound presented was “less full” or had “less energy”. All stimuli were randomly presented (in pairs and in ordering with a probability of 0.5 for each filter) 10 times for a total of 80 trials. This isolated control task lasted approximated five minutes.

A brief rest break was provided prior to the experimental conditions. As with Experiment 1, listeners were verbally instructed to identify the changing stimulus event as soon as it was detected and to respond immediately should the trial end. If participants did not respond after the trial ended, the subsequent trial automatically began after two seconds. Remaining procedural information, including the manner in which participants made responses and inter-trial intervals, were as previously described in Experiment 1. There were 32 randomized trials per condition, reflecting eight trials per target event. There were three conditions in the experiment (96 trials total). The entire experiment lasted approximately 40 minutes.

The continuous condition was predicted to produce more evidence of change deafness than the two flicker conditions due to the smaller spectral differences within local time regions, or snapshots, obtained by the listener. Hall, et al. (2015) showed that the increased difficulty in the continuous condition resulted in more errors relative to the isolated control task than with the flicker conditions, and similar results were predicted in the present experiment. Furthermore, the flicker manipulations were expected to produce little to no evidence of change deafness, as revealed by rapid response times and lower error rates. Such low error rates were expected due to the large spectral differences between adjacent snapshots. Due to the results of the Hall, et al. (2015) study, the 3*s* flicker condition was expected to display significantly faster mean response times as the participants were provided with ample time to properly encode each stimulus array, which was made from precisely the same sound events as each other condition. Because the Hall, et al. (2015) study showed no significant difference in accuracy between the flicker conditions, error rates were not expected to substantially change across all flicker conditions.

**Results and Discussion**

Median response times and mean error rates were summarized per condition in the manner previously described for Experiment 1. Specifically, the probability of errors was calculated for the control task as well as for each condition. The mean error rate from the isolated control task was compared to the mean error rates for each condition in order to provide a more reasonable estimate of the probabilities of error that were actually due to change deafness.

In the flicker conditions, listeners could not actually detect a change until the presentation of the second array on a given trial, and the time to initiate that array from the beginning of the trial varied across conditions. As a result, an evaluation of change deafness across these conditions requires timing from the start of the second array stimulus. Therefore, the amount of time prior to the beginning of the second array (1.5*s* and 3.5*s* for the 1*s* flicker and 3*s* flicker, respectively) was extracted from each participant's median response time.

Accounting for the point where participants could first reasonably detect a change in the continuous condition presented a separate challenge. Unlike spatial manipulations where the MAA and MAMA are easily identified, the threshold levels for detecting a bandwidth change in a bandpass filter are unknown. To find a rough estimate of the threshold, a short discrimination task was presented to four students in James Madison University's Auditory Perception Laboratory. Students listened to one second samples of the wide-band bell stimulus (10,106 Hz bandwidth) juxtaposed to samples of the bell stimulus with bandwidths ranging from 6,300 Hz to 4,500 Hz in 200 Hz steps. Participants listened to a randomized set of the ten stimuli and indicated whether the presented stimuli sounded the same or different. This short task was only completed for the event with the highest centroid (i.e., the bell). Stimuli at higher frequencies have greater thresholds for detection (due to the logarithmic organization of the auditory system) and they require more time to pass before a change can be detected. By using the event which requires the greatest threshold, response times for all events would be adjusted by the event which produced the poorest performance, thus decreasing the probability of finding evidence of change deafness. Detection was high for stimuli at 5,900 Hz ($p$ = 1.0), but decreased substantially at 6,100 Hz ($p$ = .25) and 6,300 Hz ($p$ = 0.0). Therefore, the threshold was identified to be between 5,900 Hz and 6,100 Hz, and 5,900 Hz was used as the best (though admittedly poor) "just noticeable difference".

Because it took 0.70*s* to move from a bandwidth of 10,106 Hz to a bandwidth of 5,900 Hz, 0.70*s* was extracted from the median response time for each listener in the continuous condition.

A dependent samples t-test was conducted to compare the error rates of the isolated control task to the mean error rates of the experimental conditions. To assess the relationship between condition and response times as a function of event type, individual median response times for each listener in each condition were submitted to two 3x4 repeated measures ANOVAs with condition (1*s* flicker, 3*s* flicker, continuous) and event (bell, dog, helicopter, and footsteps) as the factors. Like Experiment 1, the possibility of a block order effect also was evaluated by submitting median response times to a 3x3 repeated measures ANOVA with condition (1*s* flicker, 3*s* flicker, continuous) and order of presentation (first, second, third,) as the factors. A corresponding set of analyses also were completed for error rates.

Mean sensitivity was high across participants for the isolated control task, *d'* = 2.60 (*SE* = 0.13), indicating that participants were able to discriminate between the filter manipulations presented in the study. A one-way repeated-measures ANOVA, with event as the factor (bell, dog, helicopter, footsteps), was conducted to determine whether the type of event affected the listeners' sensitivity. There was no significant main effect, $F(9, 42) = 1.332$, $p = .277$, $\eta_p^2 = .087$, reflecting that listeners were similarly effective in responding to each type of event. As with Experiment 1, the high performance level in the present control task provides an added comparison condition for evaluating error rate.

Figure 5 displays the mean probabilities of errors for control and experimental conditions (panel A) and mean response times (panel B), along with corresponding standard errors. Error probabilities were not statistically different for the continuous conditions or 1*s* flicker condition, [$t(14) = .543$, $p = .596$ and $t(14) = .992$, $p = .371$, respectively] and were marginally different from the 3*s* flicker condition, $t(14) = 2.11$, $p = .053$. There were not significant differences in error rate between any of the experimental conditions, $F < 1$. Overall error rates in the experimental/change detection task were extremely low, ranging from an average probability of $p = .033$ for the 3*s* flicker condition to a probability of $p = .054$ for the continuous conditions. In fact, the mean probability of error failed to exceed observed error rates for the isolated control task. This indicates precisely how simple the frequency-based task was for the participants to complete. The low error probability is especially relevant when compared to the continuous conditions in Experiment 1 (ranging from .169 to .279). The reduced probabilities of errors for the present experiment indicates that when the task is easy enough, error rates might not provide any information regarding the occurrence of change deafness or differences across conditions.

Response times varied dramatically across conditions. As can be seen in panel B of Figure 5, the 3*s* flicker condition shows the fastest median response times while the slowest were obtained in the continuous condition. This tendency contributed to a significant main effect of condition, $F(2, 28) = 29.058$, $p < .001$, $\eta_p^2 = .675$. The untransformed response times similarly displayed a main effect, $F(2, 28) = 16.176$, $p < .001$, $\eta_p^2 = .554$. Post-hoc pair-wise comparisons confirmed that response times for the 3*s* flicker condition were significantly faster than for the 1*s* flicker and continuous conditions ($p = .025$ and $p < .001$, respectively), and response times for the 1*s* flicker condition were significantly faster than for the continuous condition ($p < .001$). The significantly slower response times for the continuous condition indicate change deafness. Unlike the continuously moving stimuli, flickering stimuli create extremely noticeable comparisons between adjacent stimulus presentations, allowing participants to successfully detect the source of a change. Moreover, when the initial array is presented for an extended time, it provides the listener with ample time for encoding the stimuli, which results in significantly faster response times.

Further analyses were conducted to determine whether the condition effects observed depended upon the target event. Figure 6 displays the mean error rates (panel A), mean response times (panel B), and corresponding standard errors associated with each target event in each condition. It is readily apparent in panel A that the probability of errors for specific events varied across conditions. This tendency contributed to a marginal interaction of condition and event, $F(2.796, 39.144) = 2.832$, $p = .054$, $\eta_p^2 = .168$. There was no main effect of event, $F(3, 42) = 1.17$, $p = .333$, $\eta_p^2 = .094$. Panel B similarly reveals that median response times differed as a function of the combination of target event and condition, and this tendency likewise contributed to both a significant interaction of event and condition, $F(3.87, 54.14) = 25.268$, $p < .001$; $\eta_p^2 = .643$ and a main effect of event, $F(3, 42) = 65.518$, $p < .001$, $\eta_p^2 = .824$. Post-hoc pair-wise comparisons of means further revealed that the overall response time with the dog was significantly slower than the bell ($p = .044$) and the helicopter ($p = .002$). Response times for the footsteps were significantly faster than the other events ($p < .001$). Furthermore, within the 1s flicker condition, performance with the dog was marginally slower than the helicopter ($p = .099$). In the 3s flicker condition, performance with the helicopter was significantly slower than for the footsteps (*p* =.039). Finally, in the continuous condition, performance with the dog was marginally slower than with the footsteps (*p* =.081).

The performance of participants also improved as they completed the experiment. This was statistically supported through a main effect of presentation order, $F(2.86, 39.01) = 9.425$, $p = .031$, $\eta_p^2 = 825$. Pair-wise comparisons further indicated that the third condition had significantly fewer errors than the first, $p = .026$. However, given the exceptionally low probability of errors throughout the experiment, this impact on performance is inconsequential.

In summary, these results provide support that change deafness can be easily demonstrated through continuous manipulations, while it is eliminated in flicker conditions. Error rates did not vary significantly between any of the conditions. In fact, they were actually lower than the isolated control task because the task was created to be simple enough to eliminate errors. Low error rates indicated that participants did not have difficulty doing the task or encoding important information about the target events. However, the task still provided some evidence of change deafness in the pattern of response times across conditions. The flicker conditions displayed substantially faster response times than the continuous condition, as the repeated static presentations allowed participants to conceive distinctly different adjacent snapshots, providing enough information to quickly determine the target event. Furthermore, when ample time is initially provided for the listener to encode the stimulus array, successful performance is dramatically faster. In contrast, the increased response times for the continuous condition suggests that participants could not compare local snapshots as effectively as in flicker conditions, thus resulting in change deafness.

# General Discussion

## Spatial Manipulation

Experiment 1 began by exploring the possibility that change deafness might vary as a function of the target event's velocity, such that slower velocities could increase change deafness while faster velocities would eliminate it. It was initially anticipated that the faster rates of change would force the immediate comparison of two adjacent perceptual snapshots, allowing for a greater ability to detect the source of the change. However, the opposite situation occurred, and the 80º/*s* condition actually displayed very high error rates. While this may initially seem to provide evidence of change deafness, there is reason to believe that the high error rates were actually the result of a failure to properly encode the moving stimulus. Performance suffered because participants were forced to compare dynamically changing scenes, which resulted in the loss of spatial acuity. For example, in the case of the dog-barking event, each individual "bark" lasted approximately 0.25*s* while simultaneously travelling approximately 24º. This was problematic because at any given point in time the participant could only hear a fraction of the "bark" at a particular location in space. Had participants been able to encode the event, error rates would probably have looked similar to the 40º/*s* and 24º/*s* conditions.

Consistent with the hypothesis, evidence of change deafness was found with the 8º/*s* condition (i.e., slowest velocity), as depicted in the error rates presented in Figure 3, panel A. Grantham (1997) had previously speculated that performance may degrade at very slow rates of change, though he did not explicitly manipulate extremely slow velocities. The high amounts of change deafness observed in the 8º/*s* condition appear to support his speculation and is likely attributable to two causes. The slow rate of change may have made it more difficult for participants to successfully distinguish between adjacent snapshots. This could be the result of either the stimulus moving so slowly that participants' adjacent snapshots were nearly identical or that the subsequent snapshots were achieved too slowly and exceeded the limits of a memory trace. A memory trace is the direct comparison of two sounds and only lasts approximately 750*ms (*Macmillan, Braida, & Goldberg, 1987). If this time is exceeded the individual rely on context memory, which is less sensitive to time (lasting approximately 2-4*s*), to identify locations of events in previous snapshots (Macmillen, et al., 1987; Macmillan, Goldberg, & Braida, 1988). Both of these explanations indicate some form of change deafness.

The 40º/*s* provided evidence of the lowest error rates and shortest response times, whereas the error rates and response times for the 24º/*s* condition were overall somewhat higher, but did not differ substantially from the 40º/*s*. Previous research has shown that there is greater sensitivity to movement in moderate rates of change (specifically between 15º to 20º/*s* as evidenced by Carlile & Best, 2002; Grantham, 1986), and this appears to be true for these two slightly faster rates of change as well. In contrast, the probability of error was substantially higher for very slowly moving events (i.e., 8º/s) and very rapidly moving events (80º/*s*). The point of degradation for fast-moving events in the present study occurred somewhere between 40º/*s* and 80º/*s.* Grantham (1997) found that performance suffered at 60º/*s.* Therefore, the point of performance degradation appears to be between 40º/*s* and 60º/*s*. Future research will have to determine the precise point of performance degradation for both fast- and slow-moving events.

The compilation of two existing hypotheses provides a suitable explanation of these results. Grantham (1985) found that the stimulus' total displacement over time served as a more salient cue than the particular rate of that change, and therefore suggested that individuals infer the direction of movement by comparing endpoints of that moving stimulus (i.e., his snapshot hypothesis). Conflicting evidence, however, showed that individuals actually are able to respond directly to rate of change while also attending to the onset and offsets discussed in Grantham's snapshot hypothesis. It was found that the information arriving between the onset and the offset of a signal contributed to the motion perception through the comparison of multiple smaller snapshots within local time regions (Perrott, Constantino, & Ball, 1993; Perrott & Marlborough, 1989). To account for this, the snapshot hypothesis was modified to explain that the auditory system could employ either the original snapshot mechanism (attending only to endpoints) or the multiple snapshot mechanism (utilizing temporally adjacent snapshots over the course of movement) for faster moving events. In contrast, he found that only the multiple snapshot mechanism is utilized for slower moving events.

Experiment 1 displayed significant differences in error rates across conditions while response times did not vary, and this lack of differences between average response times across conditions makes it difficult to come to any conclusions based on that data alone. A somewhat different understanding can be gained when the response times across conditions are compared to those of the flicker conditions in the Hall, et al. (2015) study. First, the strong similarities in both response times and error rates for the 40º/*s* and 24º/*s* conditions suggests that results displayed in the continuous conditions of the Hall, et al. (2015) study were not due to the specific velocity utilized in their study. Rather, they were likely the result of the manipulation itself. Furthermore, all four conditions of Experiment 1 and the continuous conditions from Hall, et al. (2015) displayed slow response times, while response times in the flicker conditions in Hall, et al. (2015)—specifically the 3s flicker condition—are extremely fast. Thus, at first glance, it appears that only the 8º/*s* condition shows evidence of change deafness. However, when taken in the context of Hall, et al. (2015), and as long as participants in the current investigation were not generally slow responders compared to that study, it appears that the continuous conditions encourage slower completion of the task.

**Frequency-based Manipulation**

Experiment 2 began with a question concerning whether parallel mechanisms for auditory change detection exist for both spatially- and frequency-based changes. If so, frequency-based changes should display similar patterns to spatial manipulations while exhibiting lower response times. Errors did not vary significantly across conditions, and after accounting for the probability of error in the isolated control task, it appears that the errors displayed in the experimental conditions actually represent sensitivity issues within the task itself. Some may debate whether or not the continuous manipulation provided evidence of change deafness at all. There are, however, significant differences across conditions with respect to response time. The continuously changing condition showed substantially slower response times than the flicker conditions, providing clear evidence of change deafness. Furthermore, the 3*s* flicker condition, which allowed for ample encoding time, resulted in significantly faster response times than both the 1*s* flicker and continuous conditions.

A modified version of the snapshot hypothesis could serve as a satisfactory explanation for the results. Just as an individual makes real-time perceptual snapshots of a scene to detect movement, one may separately sample a given scene and compare adjacent samples in order to detect changes in frequency. Alternatively, the snapshot hypothesis could be representative of a more complex single mechanism that is dependent upon both frequency and space. If so, the mechanism would operate in the same manner across two different dimensions, and the detection of a spatial movement would be dependent upon the frequency information of both the target and non-target events in the array. This is speculated to be the case due to the influence that frequency has on the localization of an event. For example, the time needed for higher frequencies to move completely around the head is quite long in comparison to the relative period for that wave. This is especially problematic for interaural phase difference localization cues because it either alters or prevents the proper calculation of a phase difference, and the individual may perceive the sound as coming from separate locations. Interaural level differences will also be affected by frequency, mostly due to the fact that when frequencies are around 1,200 Hz, the human head will create a sound shadow in which frequencies are blocked from reaching the ear opposite from the origin. Because frequency dramatically alters a person's ability to localize, it can be speculated that it will also affect an individual's ability to detect a changing event (Middlebrooks & Green, 1991).

Taken one step further, Grantham (1986) may clarify why response times for flicker conditions were significantly faster than for the continuously changing condition. Humans are more sensitive to auditory changes in stationary events than in dynamically changing events, and the flicker paradigm allows for comparisons across these ideal conditions (Grantham, 1986; Grantham, 1997). Experiment 2 provides support for this sensitivity in a dimension beyond what is typically studied. A second reason why response times for flicker conditions may have been significantly faster is that participants were extremely familiar with the sound of the array, as each trial always began with the same static presentation of events (i.e., all events presented through the wide-band filter). Such familiarity likely resulted in better encoding and was advantageous to faster response times once an event changed. In contrast, one event was constantly changing within the continuous manipulation, and a portion of the difference in response times between the flicker and continuous conditions may have resulted from minor problems encoding the target event due to its dynamic nature. While it is unlikely that these problems severely impacted the listener, it did add an additional level of complexity that was not experienced in the flicker conditions. Future research should evaluate this possibility by extending the time of static presentation prior to any change occurring in the continuous condition while also devising a task which can directly measure the time needed to identify each stimulus. These modifications would equate the static presentation between conditions while providing valuable information regarding the precise time required to encode a particular event.

The difference in response times for the 3*s* flicker condition and the 1*s* flicker condition, for both the present experiment and the Hall, et al. (2015) study, clearly indicate that one second is not enough time to completely encode an array comprised of the same four events. This hold true, even when the same events are consistently used. Therefore, it is likely that studies which involved one- or two-second presentations of an array composed of at least four separate events [i.e., Gregg & Samuel, 2008; Gregg & Samuel, 2009; Pavani & Turatto, 2008; also see McAnally, et al., 2010 (in one condition)] simply did not provide participants with sufficient time to encode.

Furthermore, the variable nature of the stimulus arrays (i.e., selecting from a pool of 11-18 stimuli from trial to trial) further increases the complexity for participants. In addition to the limited time to encode the stimulus array, listeners did not even know what events might make up that array. This complexity further increased uncertainty and the likelihood of informational masking.

The present study, as well as Hall, et al. (2015), made an effort to eliminate these issues and achieved full success preventing aforementioned complications. This, and the fact that the present study utilized frequency-based manipulations, led to all the conditions in Experiment 2 displaying very low response times. These response times were significantly reduced from the length of the corresponding conditions in the Hall, et al. (2015) study (1.19*s* versus 4.45*s*; 1.99*s* versus 6.41*s*; and 3.80*s* versus 9.51*s* for the 3*s* flicker, 1*s* flicker, and continuous conditions respectively). Similarly, the response time for the continuously changing condition in Experiment 2 was four seconds faster than the shortest response time for Experiment 1's continuous (i.e., 3.80*s* versus 7.82*s*) manipulations. This provides further evidence to the ease at which participants detect frequency-based changes compared to spatially-based changes.

The experiment's methodology was constructed to eliminate errors altogether, and the results indicate that the method was successful. Had frequency manipulations been constructed to be more difficult (i.e., less dramatic filter changes), there would have likely been more errors across all conditions. The logic behind the manipulation was to study frequency-based manipulations using the same experimental conditions as Hall, et al. (2015). Because frequency—not space—is audition's dominant dimension, it was unsurprising that the participants were extremely successful in completing the task.

Despite the ease of the task, the pattern of results clearly mimicked that of Hall, et al. (2015), indicating that while change deafness is eliminated in flicker conditions, it is likely to be evident with continuously changing events.

**Broader Implications**

The current investigation extends a relatively new approach to evaluating change deafness. It is possible that previously used methodologies actually increased the chances of a change detection failure and are incapable of easily distinguishing change deafness from other sources of error. It is critical that researchers design experiments in a way that maximizes the probability of change detection so when a detection failure does occur, there is clearly evidence of change deafness.

Evidence of change deafness was restricted to continuous conditions most likely due to the similarity between perceptual samples obtained by the listener. This similarity increased the difficulty in discriminating between dynamically changing events. The continuous manipulations in both Experiment 1 and Experiment 2 showed high error rates and/or long response times in comparison to the low error rates and short response times of the flicker manipulations (Experiment 2 and Hall, et al., 2015). Thus, it is clear that continuously moving stimuli are significantly more difficult to detect and could potentially provide the only means of finding valid evidence of change deafness. This, however, directly conflicts with what is known about why change *blindness* occurs. Continuous changes in the visual domain that evolve at the same rate of change as the flicker conditions should elicit the greatest change detection (Yantis, 1992). The inverse is true for audition. Contrastingly, the flicker paradigm, which is used to sever change detection in the visual domain, significantly increases one’s ability to detect auditory changes. These disparities highlight the fact that change detection mechanisms differ between vision and audition.

Researchers previously identified change deafness based upon an arbitrary cut-off point (i.e., 30% error rates) which is commonly used throughout change blindness literature (Dickerson & Gaston, 2014; Eramudugolla, et al., 2005; Gregg & Samuel, 2008). This reliance on a non-analogous phenomenon has resulted in the absence of well-established parameters for what does and does not constitute evidence of change deafness. Consequentially, a broad array of change detection failures have been inaccurately labeled as change deafness.

Change deafness should be identified by at least one experimental condition displaying either significantly higher error rates from the isolated control task, significantly longer response times from the other conditions, or both. In order to test for change deafness in a way that minimizes encoding concerns, participants should be acquainted with the events presented in the array. Minimally, listeners should complete an introductory familiarization task which introduces them to the sounds used within the study. Each experimental stimulus array should be presented for a long enough period of time to enable participants to properly encode it. While the exact length of time needed for encoding is not known, the current investigation, as well as previous investigations, have found that three to five seconds is optimal for an array composed of four stimuli. This is based off of the limited error probabilities and reduced response times which were clearly displayed within Hall, et al. (2015) and Eramudugolla, et al. (2005), both of whom used an extended presentation time. Also, no more than four to six stimuli should be used in each stimulus array so that participants can realistically encode the events presented.

The easiest way to eliminate encoding concerns in a change detection task is to use the same events throughout the entire experiment. While this is not absolutely necessary, it will reduce uncertainty and the likelihood of informational masking. Stimuli presented should also be spectrally distinct to eliminate masking, as it is considerably more difficult to detect changes between events which share a degree of sensory overlap (Dickerson & Gaston, 2014). To account for this, the present experiment utilized stimuli that were spectrally distinct. Specifically, each stimulus displayed peak intensities at different frequency regions in order to reduce the possibility of stimulus overlap.

Future experiments should follow these parameters in order to increase a participant's likelihood of change detection. If there still persists a condition which significantly differs in performance between other conditions, then this is indicative of change deafness. This requirement provides a well-defined, empirically-based definition of the phenomenon as opposed to an arbitrary cut-off point that had previously been used (Eramudugolla, et al., 2005; Gregg & Samuel, 2008).

In contrast to the parameters previously discussed, change deafness should not be studied through the one-shot paradigm, as it provides participants only one opportunity to listen to the stimulus array prior to a change occurring. This limited access to the array severely affects a listener's ability to encode the events presented and increases the likelihood of a detection failure, further blurring what is and is not change deafness. The continuous and flicker paradigms, on the other hand, work to ameliorate this issue by allowing listeners ample time to properly encode the stimulus array.

Given the results of the present two experiments, it is suspected—though cannot be confirmed—that several studies previously conducted may not demonstrate change deafness to the level that they claim. There are several potential sources of error in these studies, and change deafness is but one among many alternative explanations. It is possible that performance would have improved dramatically had these experiments simply utilized the flicker paradigm to present participants with a limited number of unchanging stimuli. Furthermore, response times likely would have become substantially smaller if the stimulus array was initially presented to listeners for an extended period of time (i.e., three to five seconds). Until follow-up studies are conducted, however, it is impossible to determine what level of error rates are representative of change deafness and what levels require a more parsimonious answer.

This paradigm can be further modified to address attentional versus inattentional deafness concepts raised in the introduction. Currently, an experiment is being developed to investigate this issue. Like Eramudugolla's (2005) investigation, performance will be compared on the same task with and without misdirected attention. In developing this experiment, it would be possible to actually manipulate and determine whether there is a difference between what is discussed as change deafness and inattentional deafness using the same stimuli used in the current investigation.

The principle motivation behind the current investigation was to gain a better understanding of when change deafness would and would not occur. Several aspects of the study's methodology should be improved upon in future research. First, the isolated control task in both experiments was used to determine baseline error rates for the experimental condition. It could be argued, however, that the control experimental tasks were not equivalent in assessing a single task. Specifically, the control task instructed participants to identify which of two isolated events was in a specific location/of a specific quality, whereas the experimental conditions instructed listeners to identify a changing stimulus. Future studies should ensure that the two tasks are equal, such that they prompt the use of similar strategies from the listener. For example, the control task for Experiment 1 could have been improved by asking participants to identify which event (bell, dog, helicopter, or footsteps) was in a specific location, rather than simply using a two-alternative forced choice method.

Next, Experiment 2's frequency-based manipulation served as a conceptual replication of the Hall, et al. (2015) spatial manipulation. While great effort was taken to equate the conditions for time, the magnitude of the spatial- and frequency-based changes between the two studies was not identical. It was not known how much of a frequency change equated to a 60 degree movement, therefore, the development of the filter-based manipulations was designed to create an above-threshold change rather than equating it to space. Future research should use direct scaling as a basis of manipulation, or alternatively, empirically evaluate each dimension prior to creating the experiment. By equating the magnitude of change across the two dimensions, researchers will be able explicitly comparable the results between experiments.

Finally, Experiment 1 is readily compared to Hall, et al. (2015) despite the reliance on different participants. Therefore, the present investigations would have benefited from the addition of a flicker manipulation to serve as a control task. These changes, if employed, will help to clarify the nature of the manipulations and their effect on change deafness by allowing the explicit comparison of a condition that is not expected to produce change deafness with a condition that is expected to.

Change deafness was originally thought to be analogous to change blindness. As demonstrated, this is an improper comparison. While extremely slow rates of visual change consistently display change blindness, evidence of the phenomenon is eliminated with rates of change that are equivalent to the flicker condition. This directly opposes what occurs in audition, where a continuously changing event goes undetected, despite utilizing a comparable speed to the flicker conditions. The current investigation sought to further explore auditory change detection with what is believed to be a more appropriate paradigm. This paradigm is analogous to change blindness, eliminates confounding factors, and ultimately allows for proper identification of change deafness. If these results continue to be supported, they will reveal fundamental differences between auditory and visual attentional mechanisms.

References

Ableton AG (2015). Ableton Live Suite 9.1.7 [Computer software].

Agres, K. R., & Krumhansl, C. L. (2008). Musical change deafness: The inability to detect change in a non-speech auditory domain. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 969-974). Cognitive Science Society Austin, TX.

Backer, K. C., & Alain, C. (2012). Orienting attention to sound object representations attenuates change deafness. *Journal of experimental psychology: human perception and performance*, *38*(6), 1554.

Beck, D. M., Rees, G., Frith, C. D., & Lavie, N. (2001). Neural correlates of change detection and change blindness. *Nature Neuroscience*, *4*(6), 645-650.

Blackmore, S. J., Brelstaff, G., Nelson, K., & Troscianko, T. (1995). Is the richness of our visual world an illusion? Transsaccadic memory for complex scenes. *Perception-London*, *24*, 1075-1075.

Carlile, S., & Best, V. (2002). Discrimination of sound source velocity in human listeners. *The Journal of the Acoustical Society of America*, *111*(2), 1026-1035.

Chandler, D. W., & Grantham, D. W. (1992). Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity. *The Journal of the Acoustical Society of America*, *91*(3), 1624-1636.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with Two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975-979.

Constantino, F. C., Pinggera, L., Paranamana, S., Kashino, M., & Chait, M. (2012).

Detection of appearing and disappearing objects in complex acoustic scenes. *PLoS ONE*, *7*(9). doi:10.1371/journal.pone.0046167

Dickerson, K., & Gaston, J. R. (2014). Did you hear that? The role of stimulus similarity and uncertainty in auditory change deafness. *Frontiers in psychology*, *5*.

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd Jr, G. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *The Journal of the Acoustical Society of America*, *114*(1), 368-379.

Eramudugolla, R., Irvine, D. R., McAnally, K. I., Martin, R. L., & Mattingley, J. B. (2005). Directed attention eliminates 'change deafness' in complex auditory scenes. *Current Biology*, *15*(12), 1108-1113.

Fenn, K. M., Shintel, H., Atkins, A. S., Skipper, J. I., Bond, V. C., & Nusbaum, H. C. (2011). When less is heard than meets the ear: Change deafness in a telephone conversation. *The Quarterly Journal of Experimental Psychology*, *64*(7), 1442 1456.

Gallace, A., Auvray, M., Tan, H. Z., & Spence, C. (2006). When visual transients impair tactile change detection: A novel case of crossmodal change blindness? *Neuroscience Letters*, *398*(3), 280-285.

Grantham, D. W. (1985). Auditory spatial resolution under static and dynamic conditions. *The Journal of the Acoustical Society of America*, *77*(S1), S50-S50.

Grantham, D. W. (1986). Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *The Journal of the Acoustical Society of America*, *79*(6), 1939-1949.

Grantham, D.W. (1995). Spatial hearing and related phenomena. In Moore B.C.J. (Ed.) *Hearing: handbook of perception and cognition* (297–345). New York: Academic Press.

Grantham, D. W. (1997). Auditory motion perception: Snapshots revisited. In Gilkey, R. & Anderson, T. R. (Eds.), *Binaural and spatial hearing in real and virtual Environments* (295-327). Psychology Press.

Gregg, M. K., & Samuel, A. G. (2008). Change deafness and the organizational properties of sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(4), 974.

Gregg, M. K., & Samuel, A. G. (2009). The importance of semantics in auditory representations. *Attention, Perception, & Psychophysics*, *71*(3), 607-619.

Hall, M. D., Peck, R. B., Gaston, J., & Dickerson, K. (2015, Nov). *The relationship between auditory change detection and the encoding of events*. Poster presented at meeting of the Psychonomics Society, Chicago, IL.

Harris, C. S. (1972). Effects of intermittent and continuous noise on serial search performance. *Perceptual and motor skills*, *35*(2), 627-634.

Harris, J. D., & Sergeant, R. L. (1971). Monaural/binaural minimum audible angles for a moving sound source. *Journal of Speech, Language, and Hearing Research*, *14*(3), 618-629.

Hirsh, I. J. (1952). Certain temporal factors in audition. *Science*, *116*(3020), 523-524.

Hirsh, I. J. (1959). Auditory perception of temporal order. *The Journal of the Acoustical Society of America,* 31, 759- 767.

Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition*, *79*(1), 89-113.

Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual deficit during a mental task. *Science*, *157*(3785), 218-219.

Kahneman, D. (1970). Remarks on attention control. *Acta Psychologica*, *33*, 118-131.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kubovy, M. (1981) Concurrent-pitch segregation and the theory of indispensable Attributes. In Kubovy, M., & Pomerantz, J. R. (Eds.), *Perceptual organization* *(*55-99). Erlbaum, Hillsdale, NJ.

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, *80*(1), 97 126.

Leek, M. R., Brown, M. E., & Dorman, M. F. (1991). Informational masking and auditory attention. *Perception & psychophysics*, *50*(3), 205-214.

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, *4*(4), 501-506.

Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, *280*(5371), 1930-1934.

Mack, A., Tang, B., Tuma, R., Kahn, S., & Rock, I. (1992). Perceptual organization and attention. *Cognitive Psychology*, *24*(4), 475-501.

Mack, A., & Rock, I. (1998). Inattentional blindness: Perception without attention. *Visual Attention*, *8*, 55-76.

Macmillan, N. A., Braida, L. D., & Goldberg, R. F. (1987). Central and peripheral processes in the perception of speech and nonspeech sounds. In *The Psychophysics of Speech Perception* (pp. 28-45). Springer Netherlands.

Macmillan, N. A., Goldberg, R. F., & Braida, L. D. (1988). Resolution for speech sounds:

Basic sensitivity and context memory on vowel and consonant continua. *The Journal of the Acoustical Society of America*, *84*(4), 1262-1280.

McAnally, K. I., Martin, R. L., Eramudugolla, R., Stuart, G. W., Irvine, D. R., & Mattingley, J.B. (2010). A dual-process account of auditory change detection. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 994.

Middlebrooks, J. C. & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*. 42, 135-159.

Mills, A. W. (1957). On the minimal audible angle. *The Journal of the Acoustical Society of America*, *30*(4), 237-246.

Moore, B. C., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, *45*(4), 224-240.

Neuhoff, J. G., Schott, S. A., Kropf, A. J., & Neuhoff, E. M. (2014). Familiarity, expertise, and change detection: Change deafness is worse in your native language. *Perception*, *43*(2-3), 219-222.

Neuhoff, J. G., Wayand, J., Ndiaye, M. C., Berkow, A. B., Bertacchi, B. R., & Benton, C. (2015). Slow change deafness. *Attention, Perception, & Psychophysics*, *77*(4), 1189-1199.

Pavani, F., & Turatto, M. (2008). Change perception in complex auditory scenes. *Perception & Psychophysics*, *70*(4), 619-629.

Perrott, D. R., Costantino, B., & Ball, J. (1993). Discrimination of moving events which accelerate or decelerate over the listening interval. *The Journal of the Acoustical Society of America*, *93*(2), 1053-1057.

Perrott, D. R., & Marlborough, K. (1989). Minimum audible movement angle: marking the end points of the path traveled by a moving sound source. *The Journal of the Acoustical Society of America*, *85*(4), 1773-1775.

Perrott, D. R., & Musicant, A. D. (1977). Minimum auditory movement angle: Binaural localization of moving sound sources. *The Journal of the Acoustical Society of America*, *62*(6), 1463-1466.

Perrott, D. R., & Tucker, J. (1988). Minimum audible movement angle as a function of signal frequency and the velocity of the source. *The Journal of the Acoustical Society of America*, *83*(4), 1522-1527.

Psychology Software Tools, Inc. [E-Prime 2.0] (2012). Retrieved from http://www.pstnet.com.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347-356.

Reale, R. A., & Imig, T. J. (1980). Tonotopic organization in auditory cortex of the cat. *Journal of Comparative Neurology*, *192*(2), 265-291.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, *3*(4), 261-270.

Rensink, R. A. (2002). Change detection. *Annual review of psychology*, *53*(1), 245-277.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368-373.

Robertson, L., Treisman, A., Friedman-Hill, S., & Grabowecky, M. (1997). The interaction of spatial and object pathways: evidence from Balint's syndrome. *Journal of Cognitive Neuroscience, 9*(3), 295-317.

Saenz, M. & Langers, D. R. M. (2014). Tonotopic mapping of human auditory cortex. *Hearing Research,* 307, 42-52.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception-London*, *28*(9), 1059-1074.

Simons, D. J., Franconeri, S. L., & Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception-London*, *29*(10), 1143-1154.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*(7), 261 267.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16-20.

Sinnett, S., Costa, A., & Soto-Faraco, S. (2006). Manipulating inattentional blindness within and across sensory modalities. *The Quarterly Journal of Experimental Psychology*, *59*(8), 1425-1442.

Snyder, J. S., & Gregg, M. K. (2011). Memory for sound, with an ear toward hearing in complex auditory scenes. *Attention, Perception, & Psychophysics*, *73*(7), 1993 2007.

Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(3), 555.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*, *88*(1), 97-100.

Ungerleider, L. G., & Haxby, J. V. (1994). ‘What’ and ‘where’ in the human brain. *Current Opinion in Neurobiology*, *4*(2), 157-165.

Vitevitch, M. S. (2003). Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 333.

Wave Arts, Inc. (2012) Panorama 5 [Computer software].

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*(3), 295-340.

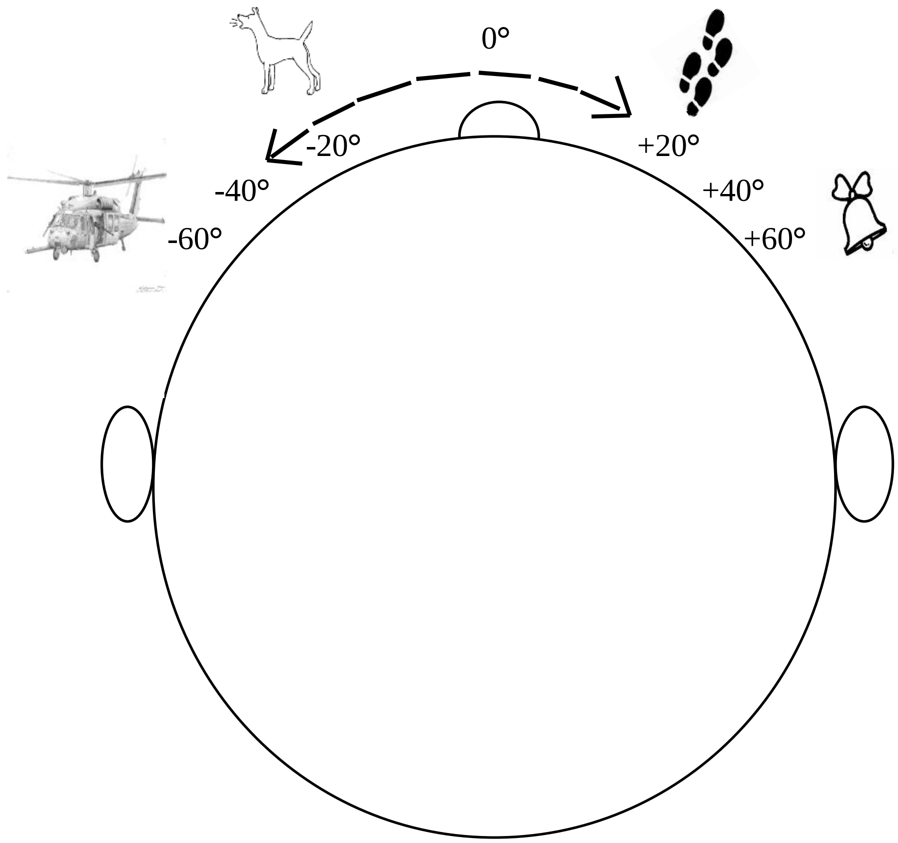


*Figure 1.* Schematic drawing of starting positions and spatial movements for target stimuli in Experiment 1. Stimuli oscillated a fixed distance (e.g., 60º) on the azimuth at a velocity of 80º/*s*, 40º/*s*, 24º/*s* or 8º/*s*.

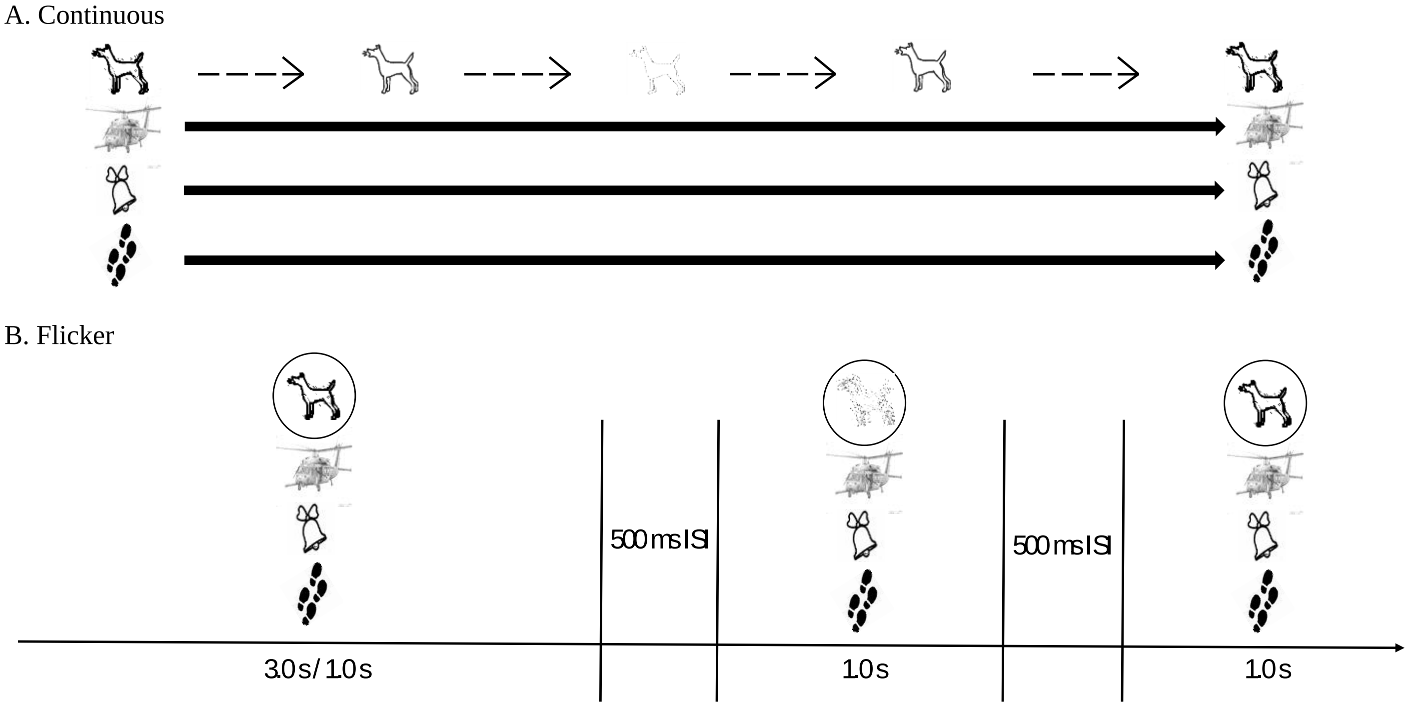


*Figure 2.* Schematic drawing of spectral changes for target stimuli for Experiment 2 including for the continuous manipulation (panel A) and for the 1*s* and 3*s* flicker manipulations (panel B). All stimuli will initially be presented through a wide-band filter, and the target stimulus will oscillate between the wide- and narrow-band filters. The shading indicates frequency-based information is changing within the target stimulus (e.g., the dog).

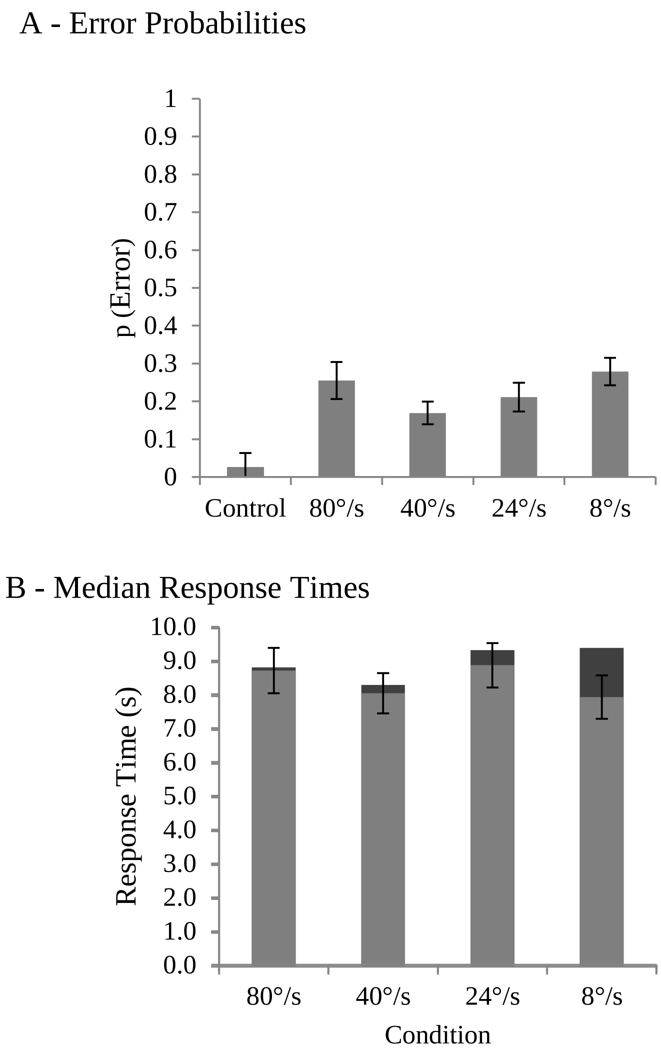


*Figure 3.* Mean probabilities of error (panel A), median response times (panel B), along with corresponding standard errors in Experiment 1 as a function of target velocity. Mean probability of error for the isolated control task is displayed (panel A). The double bar in response times (panel B) indicates the untransformed data (in dark gray).

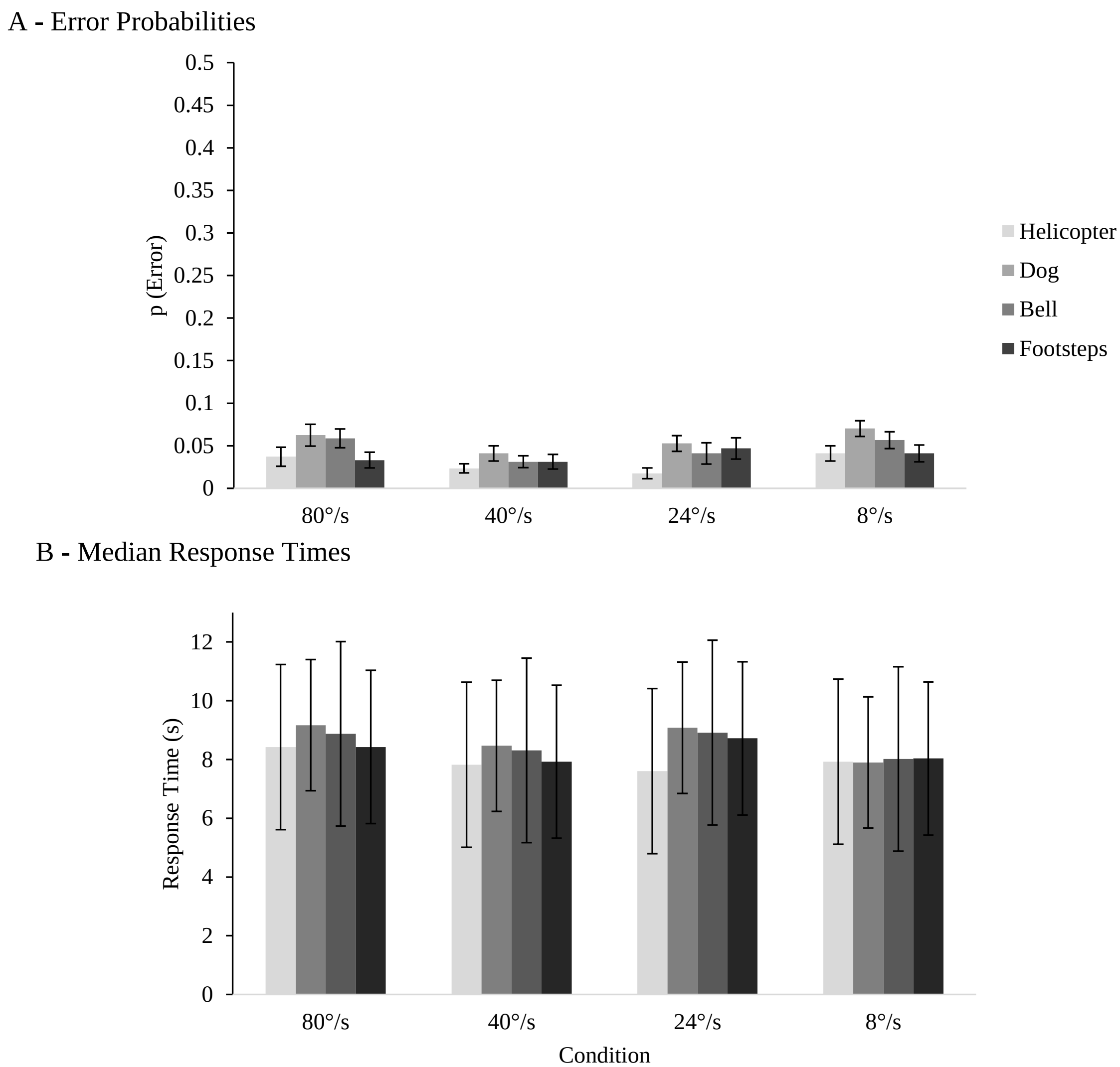


*Figure 4.* Mean probabilities of error (panel A), median response times (panel B), along with corresponding standard errors in Experiment 1 as a function of target velocity and target event.

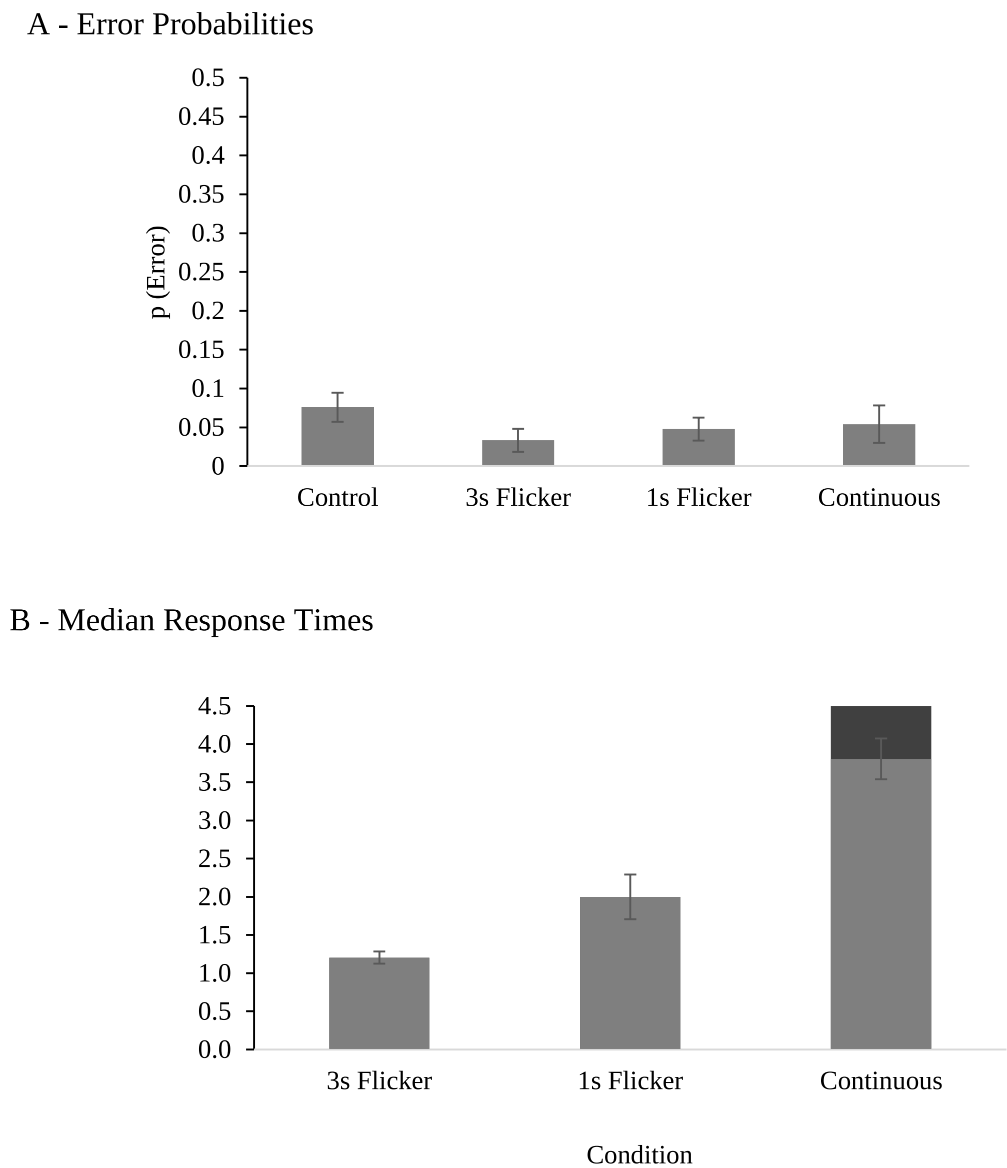


*Figure 5.* Mean probabilities of error (panel A), median response times (panel B), along with corresponding standard errors in Experiment 2. Mean probability of error for the isolated control task is displayed (panel A). The double bar in response times for the continuous condition (panel B, far right) indicates the transformed and untransformed data.

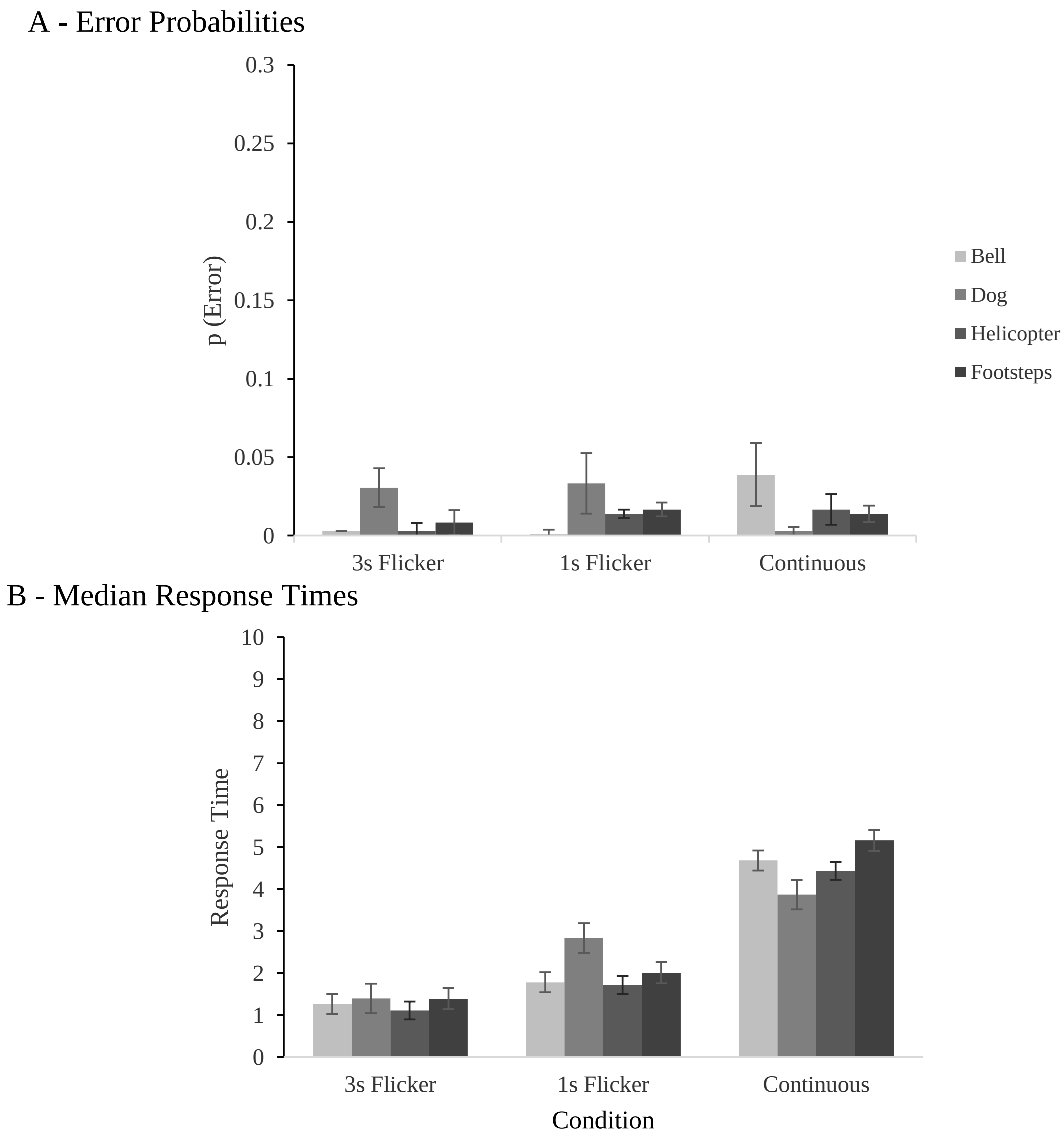


*Figure 6.* Mean probabilities of error (panel A), median response times (panel B), along with corresponding standard errors in Experiment 2 as a function of target event.