12-2013

Effects of Time Varying Background Noise Conditions On Human Perception and Performance

Andrew H. Hathaway
University of Nebraska-Lincoln, ahathaway@unomaha.edu

EFFECTS OF TIME VARYING BACKGROUND NOISE CONDITIONS ON HUMAN PERCEPTION AND PERFORMANCE

by

Andrew Harris Hathaway

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Architectural Engineering

Under the Supervision of Professor Lily M. Wang

Lincoln, Nebraska

December, 2013

EFFECTS OF TIME VARYING BACKGROUND NOISE CONDITIONS ON HUMAN PERCEPTION AND PERFORMANCE

Andrew H. Hathaway, M.S.

University of Nebraska, 2013

Adviser: Lily M. Wang

This thesis was designed to study the effects of changing noise conditions on human perception and performance. In two phases, participants were exposed to a number of noise conditions and their performance on an arithmetic task involving short-term memory was monitored and their subjective perception of noise conditions was collected via questionnaires.

In the first phase, participants were tested while being subjected to RC-29(H) and RC-47(RV) conditions created by broadband noise fluctuating on different time intervals, resembling the changing noise conditions potentially found in modern HVAC systems. These intervals varied from two minutes to ten minutes. Results show a significant relationship, $p<0.05$, between task performance in the form of percentage of correct responses and the noise conditions; as the fluctuation time interval shortened, subjects' performance decreased.

In the second phase, participants were tested while being exposed to four different levels of the noise bursts presented with or without an associated rattle noise, resembling low-level sonic booms potentially produced by newly developed supersonic aircraft as experienced in the built environment. The noise bursts exhibited peak A-weighted sound

pressure levels (L_{Apk}) ranging from 55 to 70 dBA. Few statistically significant relationships were found in relation to task performance; however, statistically significant relationships were seen in most of the subjective perception ratings. Both the 70 dBA and 65 dBA were rated statistically significantly more annoying than the 55 dBA and 60 dBA bursts alone as well as the 55 dBA burst plus rattle, implying that noise bursts at or above 70 dBA, and potentially at or above 65 dBA, with accompanying rattle should be avoided. At lower levels, the addition of rattle in this lab study did not result in much difference. It is suspected that rattle occurring in a person's personal living or work space could be considered to be more annoying. Field studies are suggested for future work.

Acknowledgements

First of all, I must thank the NASA Nebraska Space Grant Consortium for funding this study. This work would not have been possible without their financial support.

Thank you to my family for their continued support and encouragement to pursue a graduate degree in architectural engineering and acoustics. Thank you to all of my friends, especially Rachel Minier, for being around for support, advice, and most importantly distraction. Thanks to all of my bandmates, Alexander Kirtz, Andrea Purdy, Kara Purdy, Nathan Soto, and Justin Vaughan, for providing relief from all of this work through music.

A big thanks to all of my fellow and former classmates here at the University of Nebraska, especially Christopher Ainley for his thesis work to pave the way for my thesis. Matt Blevins, Kristin Hanna, Ellen Peng, and John Lloyd put in many hours assisting with this study, for which I am incredibly grateful.

I must also make sure and thank Paul Sim for his creation of the JavaScript program that was robust enough to be used for the arithmetic test in Christopher Ainley's study as well as this study. Thank you to Dr. Jonathan Rathsam from the NASA Langley Research Center for his guidance and availability to answer question about his work pertaining to this project.

Thank you to my advisory committee, Dr. Siu-Kit (Eddie) Lau and Dr. Clarence Waters, for their guidance through this process. Finally, and most importantly, thank you to Dr. Lily Wang for being my teacher, advisor, and mentor over the past two years.

Table of Contents

List of Tables

Chapter 4

Chapter 6

List of Figures

Chapter 5

Chapter 6

Chapter 1: Introduction

1.1. Introduction to Work

Researchers have been attempting to better understand the impact of noise on humans, particularly on perception and performance. Humans are exposed to numerous potentially annoying and distracting noises throughout the day. This research, in two phases, focuses on two types of noises humans can be exposed to within the built environment.

The design of aircraft that produce low level sonic booms is currently underway, with the hope that these aircraft may one day be used for flight over land. NASA is interested in furthering the study of what effect these low level sonic booms may have on humans on ground, particularly indoors. These booms are impulsive in nature and would be experienced randomly in terms of timing. Recent research has grown regarding these effects of low level sonic booms on the ground (Sullivan et al. 2010, Marshall and Davies 2010, Marshall and Davies 2011, Rathsam et al. 2012) as well as rattle noise often associated with these noises indoors (Loubeau et al. 2013, Rathsam et al. 2013).

Also, modern heating, ventilation, and air conditioning (HVAC) designs can lead to fluctuating noise conditions within the built environment. Some research has been conducted on the effects of fluctuating background noise conditions (Teichner et al. 1963, Moorhouse et al. 2007, Dittrich and Oberfield 2009, Wang and Novak 2010). However, much more work is still needed to fully explore and understand the effects of these types of fluctuating noise conditions.

Human performance has been analyzed in a number of different ways and under other types of noise conditions; one such way is to use an arithmetic task that involves

memorization as an appropriate performance task under different noise conditions (Broadbent 1958). This task was selected for the current research because it has been found that loud bursts of noise can impact performance on this test (Woodhead 1964) and indirectly perception (Ainley 2012).

While previous studies have analyzed the effects of different types of noise bursts on some combination of human performance and perception (Woodhead 1958, Woodhead, Woodhead 1959, Berlyne et al. 1966, Sullivan et al. 2010, Marshall and Davies 2010, Marshall and Davies 2011, Rathsam et al. 2012), there are still some areas that have yet to be fully analyzed. For instance, varying levels of short noise burst stimuli with and without a rattle element could be studied with relationship to the arithmetic task previously mentioned. This study aims to use a finer range of noise burst levels around the cut-off level found in Ainley's tests while comparing rattle and non-rattle bursts. The benefit of this is to help determine whether the addition of rattle noise related to bursts of noise experienced indoors has any effect on performance and the perceived qualities of the noise.

The goal of this research project is to better quantify human reactions in the form of human performance and perception to short bursts of broadband noise with and without a rattle element and also to fluctuation background noise level conditions of different time intervals. Performance was analyzed by the total percentage of correct answers and average response time for each problem in the arithmetic task. Perception was analyzed via responses to subjective questionnaires that cover loudness of noise, changes in the noise, rumble of the noise, annoyance to the noise, and distraction of the noise. The fluctuation phase of this study aims to determine a cut-off time interval for

unacceptable background noise level fluctuations. The main questions involved with the bursts of noise plus rattle phase are determining if rattle accompanying a noise burst leads to a significant detriment in performance or significant difference in ratings of subjective perception, compared to the same burst of noise presented without rattle.

1.2. Outline of Thesis

Chapter 2 discusses previous research pertinent to this study and explains how this study was developed. Chapter 3 presents the methodology including the creation of the sound signals and test sessions, and the statistical analyses used in the level fluctuations phase of this study. Chapter 4 presents and discusses the results of the level fluctuations phase of this study. Chapter 5 presents the methodology including the creation of the sound signals and test sessions, and the statistical analysis used in the bursts of noise plus rattle phase of this study. Chapter 6 presents and discusses the results of the bursts of noise plus rattle phase of this study. Chapter 7 summarizes the results and presents ideas for future work.

Chapter 2: Literature Review

This chapter discusses previous research that led to the motivation for and application to this research. Previous research is separated into subsections involving (1) types of noise stimuli, (2) performance tasks, and (3) subjective perception. The application of the previous studies to this study will also be discussed.

2.1. Types of Noise Stimuli

Researchers have been attempting to better understand the impact of noise on humans, particularly on perception and performance. This impact depends greatly on the type of noise stimulus and on the type of task. Unexpected, or novel, noise stimuli can either facilitate or distract from performance of a visual task depending on both the attention demands of the task and the relationship between noise stimuli and task (SanMiguel et al. 2010).

2.1.1. Continuous Noise Stimuli

Some early research finds constant noise above 90 dB to be detrimental to the performance on a few tasks (Broadbent 1957). The effects of both high and low frequency noise on a reaction task where subjects were required to touch a brass disc with a stylus when a corresponding light was lit were studied. Two versions of recorded noise of actual machinery of approximately equal energy in each 1/3 octave band from 100 to 5000 Hz were used: one filtered above 2000 Hz and one filtered below 2000 Hz. 24 subjects participated in two 25 minute sessions separated by 24 hours – one session in high frequency noise and one session in low frequency noise. The noise was played continuously throughout each session at 80, 90, and 100 dB for the high frequency noise and 83, 93, and 103 dB for the low frequency noise. The three dB increase for the low

frequency noise was to give an approximate equal subjective loudness. The 24 subjects were split into three equal groups that corresponded to a pair of intensity levels. A significant decrease in performance, as measured in the form of errors, was found in the sessions with 93 and 103 dB noises.

Broadbent then went on to study the effects of continuous noise on the performance on an "intellectual" task in the form of a subtraction task involving heavy use of the subjects' working memory (1958). Subjects were again divided into three groups and participated in two sessions 24 hours apart. The first group experienced two sessions of continuous 70 dB noise, the second group experienced a session of 70 dB noise and then a session of 100 dB noise, and the third group first experienced the 100 dB noise and then experienced the 70 dB noise. A significant decrease in performance over time was found in the 100 dB cases for both groups. However, a significant decrease in performance over time was also found in the 70 dB case for the group that experienced it on the second day.

Frankenhaueser and Lundberg studied the effects of continuous white noise on the performance of a different type of arithmetic task (1977). Subjects were divided into three groups and experienced white noise of 56, 72.5, or 85 dBA in the first 75 minute session; these levels were chosen to have subjectively equal intervals between them. All subjects experienced the same 72.5 dBA noise during a second 75 minute session. As also found in Broadbent's test, performance results were significantly lower in the second session for subjects that had experienced a louder noise in the previous session.

2.1.2. Low Frequency Noise Stimuli

Different noise characteristics, such as frequency or temporal content, also affect human performance and perception – not just the overall level of the noise. Low frequency noise has been found to affect work performance and perceived annoyance (Persson Waye et al. 2001). Subjects were evaluated for sensitivity to noise in general via a questionnaire developed by Weinstein in 1978, as well as sensitivity to low frequency noise via their own questionnaire. They performed reaction time tasks, short-term memory tasks, proof-reading tasks, and grammatical reasoning tasks during two sessions: one in a flat frequency reference noise and the other in a digitally processed, low frequency version of the reference noise. Both noises were based off of a recorded ventilation noise and were presented at 40 dBA. Subjects performed worse on the tasks and rated the noise more annoying in the low frequency noise session than the flat frequency noise session, with more pronounced effects on the subjects rated with a high sensitivity to low frequency noise. The average annoyance rating of all subjects of the reference noise was 2 while the low frequency noise was rated 2.5, both on a 4 point scale, while the subjects with a high sensitivity to low frequency noise rated the noises as 2.3 and 3.1 respectively.

2.1.3. Fluctuating Noise Stimuli

Teichner et al. found an increase or a decrease in noise detrimental to performance for a 60 minute memory task session involving a single fluctuation from a high level to low level or low level to high level (1963). The decrease in performance was more pronounced the greater the difference between the lower and higher level became no matter which level was experienced first.

Moorhouse et al. studied level fluctuations in low frequency noise in attempts to define parameters to quantify fluctuating noise compared to steady noise (2007). They defined a noise as fluctuating when the difference between L_{10} and L_{90} is greater than 5 dB and the rate of change for the root mean square fast sound pressure level is greater than 10 dB per second. For low frequency noises defined as fluctuating by these parameters, it was found that their acceptable level was on average 5 dB higher than non-fluctuating low frequency noise.

Dittrich and Oberfield studied perceived loudness and annoyance of 900 millisecond noise stimuli randomly changing level every 100 milliseconds utilizing a two-interval forced-choice task (2009). The sound pressures were drawn from a normal distribution around a center SPL utilizing a standard deviation of 2.5 dB. It was found that perceived annoyance was not just linked to the perceived loudness of the sound. The behavior of the sound in the first 100 – 300 milliseconds was found to have the greatest impact on the annoyance rating of the overall signal, meaning that a primacy effect (the tendency to remember or be influenced by the beginning of a signal than the rest of it) is apparent in the results.

2.1.4. Noise Burst Stimuli

A group of 24 men were tested to see if their performance on a visual task involving matching cards (a Mackworth multimodal test) was hindered by bursts of noise presented during both "busy and slack periods" (Woodhead 1958). The sound stimuli presented was a four second tape recording of an explosive sound with a peak intensity of 100 dB. Since it was a recording of an actual explosion, the signal was not constant

throughout. The spectral content was centered around 300 Hz in the first second, but shifted to center around 2000 Hz in the one second later.

Subjects experienced four different iterations of the sound stimuli presented over four minute sessions: two different versions with two impulses presented at busy periods and two impulses presented at slack periods, one version similar to the first two but with a visual stimuli alerting the subject three seconds before the burst, and one silent control session. The order of presentation of sessions was randomized via a Latin square design.

It was found that the burst of noise caused decrease in performance on the visual track that was apparent for the 30 seconds following the presentation of the burst. There were no statistically significant findings indicating whether the burst being presented during busy or slack times had any effect on performance. The warning light actually proved a hindrance to performance – a trend, but not a statistically significant one. It was probably more of a distraction than an aid to prepare for the upcoming burst.

Woodhead then tested for variation in performance on the Mackworth test relative to the intensity of a low frequency burst (1959). The noise stimuli presented was a 0.95 second recording of a rocket firing with most of the energy below 150 Hz and none above 3000 Hz. The signals had peak intensities of 85, 95, and 115 dB and the presentation order was randomized via a Latin square design.

She again found the decrease in performance to last for 30 seconds following the presentation of the burst of noise. The decrease in performance was significantly greater in the 95 and 115 dB sessions. Comparing her results to previous studies, Woodhead concluded that 90 dB was a critical level.

However, not all research has found bursts of noise to be detrimental to task performance. Berlyne et al. found that bursts of 75 dB white noise caused arousal in subjects leading to better memory retention in a paired-associate memory test (1966). This corroborates results of non-acoustic psychology tests showing heightened arousal to be linked with short-term memory retention such as Walker and Tarte (1963) and Weiner and Walker (1966).

Woodhead utilized a 100 dB version of the previously discussed rocket recording to investigate changes in performance on an adapted version of Broadbent's previously mentioned arithmetic task when the burst was presented in the memorization phase or the calculation phase (1964). The burst caused a statistically significant decrease in performance when presented during the memorization phase (from 81 to 68 percent correct). However, it may have served as an arouser during the calculation phase, leading to a faster calculation time with no decrease in accuracy.

More recently, research has been conducted to see if bursts of noise with intensities on the order of potential low-level sonic booms as experienced indoors have any detrimental effect on perception and performance when presented during the memorization phase of the arithmetic task (Ainley 2012). Ainley utilized 250 millisecond filtered white noise bursts with peak A-weighted sound pressure levels (L_{Apk}) ranging from 47 to 77 dBA presented over a generated ambient background noise of 37 dBA equivalent continuous A-weighted sound pressure level (L_{Aeq}).

Ainley found no significant relationship between task performance and the level of impulse. However, a decreasing trend in percentage of correct answers in impulse-presented questions as the impulse level increases can be seen. Based on subjective

perception ratings, Ainley suggests that bursts of noise with a L_{Apk} of 67 dBA or higher may be considered unacceptable when presented in a background noise level of 37 dBA Leq.

2.1.5. Sonic Boom Stimuli

One particular type of noise burst stimuli that is of interest is the sonic boom. As with other impulsive noise testing, performance results in tests involving sonic booms have spanned from impairment to improvement and depended on the type of task. Bursts of noise tend to cause a startle reflex or an orienting response in humans, and Thackray discussed these two as they relate to sonic booms (1972).

The startle reflex is primarily an involuntary muscle response starting with an eye blink and moves toward the legs. The overall startle can last from 0.3 to 1.5 seconds depending on the intensity of the individual reaction. The involuntary muscle response can be disruptive and impair performance. Habituation has been shown to lessen the effects of the startle reflex. However, the eye blink has not been shown to habituate.

The orienting response tends to occur due to stimuli of lesser intensity than one that would evoke a startle response. The orienting response is characterized by a turning of the head or body toward the source of the stimuli. The shift of attention caused by the orienting response can cause disruption, but it may also be a source of arousal leading to better performance.

Thackray, Touchstone, and Bailey then studied human reactions to simulated sonic booms, as experienced indoors, via a hand-steadiness test and some physiological measurements on twenty male university students (1974). The simulated sonic booms utilized had indoor peak A-weighted sound pressure levels (L_{Apk}) of 74 and 83 dBA. The

subjects performed significantly worse at the hand-steadiness test, indicating a greater startle response, in sessions involving the higher level boom. Responses to the lower level booms fit more with characteristics of the orienting response. They concluded that their tested indoor intensities were right around threshold levels for evoking startle responses strong enough to be measured by their tests. They report that the startle responses measured in this test could be disruptive to tasks involving precise arm and hand work, but they do not believe it to be enough to disrupt performance on other types of tasks. They also note that it is likely that these laboratory responses would differ from "real-world" responses.

Thackray, Touchstone, and Bailey then attempted to find a cut-off level below which subjects would not experience a startle reflex by utilizing the same hand-steadiness test as mentioned above (1975). Subjects were exposed to three sets of two repeated simulated sonic booms with L_{Apk} levels of 65, 71, or 74 dBA as experienced indoors. Results showed that about one-fifth of the subjects experienced an arm-hand startle for the two louder booms, while no subjects responded to the 65 dBA boom. Subjective annoyance responses were also collected, with no significant differences reported between the different burst levels. 60-70% of the subjects reported that they believed that they would be able to adapt to booms of these levels over time.

A second test was carried out to test habituation to higher level simulated booms. Subjects were exposed to 12 booms all at either 72 or 81 dBA as experienced indoors. No significant evidence of habituation of the eye blink response was seen. However, significant habituation effects were seen in the arm-hand response test.

In other research focused on subjective perception to sonic booms, subjects tend to rate booms experienced indoors more harshly than booms experienced outdoors (Johnson and Robinson 1967, Miller 2011). This is believed to be the case because people have different expectations with regard to noises experienced inside a building as they would to noises experienced while outdoors.

2.1.6. Rattle Noise Stimuli

Other important factors affecting human perception and performance, particularly in the study of loud bursts of noise and sonic booms, are extra noises and vibrations in the built environment cause by the burst or boom. One of particular interest is the resulting rattle produced by the transmission of low frequencies through the structural elements of a building, something easily excited by sonic booms (Miller 2011). In a questionnaire asking whether indoor or outdoor perceived sonic booms were found to be the most annoying, 83% indicated indoor perceived booms as more annoying. Many elaborated that this was because of the presence and annoyance of rattle.

In another study detailed in Miller's report, the annoyance due to rattle is investigated. Subjects reported being more annoyed by the rattle of larger structural elements, such as windows and doors, than they were by small objects, such as glasses or wall-hung artwork. The researchers attribute this to subjects believing that the rattling of larger objects could potentially cause greater harm.

Ongoing work studying the perception of rattle noise is currently being carried out at NASA Langley Research Center (Loubeau et al. 2013, Rathsam et al. 2013). In the first of these studies, psychoacoustic tests asking subjects to evaluate the noise based on several factors, including annoyance, were carried out while rattle noise was presented

over headphones both with and without accompanying sonic booms. Annoyance to the different rattles varied, even though they all had the same Perceived Level (PL) values. Rattles generated by larger objects were perceived as more annoying than rattles generated by smaller objects. The combination of sonic boom and rattle noise was shown to sometimes be perceived as more annoying than the sonic boom alone.

Rathsam et al. has utilized the Interior Effects Room at NASA Langley Research Center to develop predictive capabilities of annoyance to booms experienced indoors (2013). The Interior Effects Room provides a more realistic listening environment than headphones as well as low frequency generated tactile vibrations. A broadband background noise was added to the environment measuring 38 dBA. Test subjects were again asked to subjectively rate the presented booms, rattles, and booms plus rattles. Weighted peak acceleration of the vibrations in the floor at the subjects' feet was found to be the best single-predictor of subject's annoyance. Rathsam states that follow up tests are necessary to isolate the effects of acoustic and vibration stimuli.

2.2. Task Performance under Different Noise Conditions

A number of tasks have been used to study the effects of noise on task performance, including, but not limited to: visual (Broadbent 1957, Woodhead 1958), reaction time, short-term memory, proof reading, verbal grammatical reasoning (Persson Waye et al. 2001), typing, and mathematical (Wang and Novak, 2010) tasks.

Broadbent developed an arithmetic task for his study on the effects of continuous noise on task performance, as mentioned in Section 2.1.1. (1958). This study was designed to confirm findings from a previous similar study that utilized a simpler visual

task, that a higher noise level was detrimental to task performance and that it had an effect on performance that lasted into another session not involving noise.

The arithmetic task that he devised involved first presenting subjects with a six-digit number for memorization. The subjects pushed a button to make the six-digit number disappear once they felt that they had adequate time to memorize the number. A four-digit number was then immediately presented, and the subject was asked to subtract this number from the memorized six-digit number. A session consisted of 30 of these subtraction problems, and subjects participated in one session per day for two days in a row. The subjects (18 males) also had a practice session during orientation the day before the first actual session. The number of correct responses was recorded, as well as the observation time of the first number, and calculation time after the second number appeared.

As mentioned in Section 2.1.1, the subjects performed worse in the session with the louder background noise level. They also performed worse during the quieter session if they experienced the loud session the day before but not the other way around. In regards to the observed aftereffect, Broadbent poses a possible explanation that attention may be distributed between multiple sensory channels. If the attention to a particular task has been previously interrupted, it can continue to be interrupted even without the presence of the interrupting stimuli.

Broadbent also points out that this task requires the division of attention between the immediate memory storage and the calculation phase, and that a previous study suggested that noise effects are not apparent when attention is undivided. He concludes

that it should not be assumed that working memory will be affected by noise if it is the only task being performed.

Woodhead later utilized a modified version of Broadbent's arithmetic task to study the effects of noise of 100 dB on human performance (1964). Woodhead modified the task by only allowing subjects 10 seconds to memorize the six-digit number – a time interval based on the average memorization times from Broadbent's results.

The bursts of noise were presented either four seconds into the memorization period or five seconds into the calculation period. Woodhead reported a significant decrease in performance, from 81 to 68 percent correct responses, when the burst was presented during the memorization phase as compared to a session without any bursts of noise. No change in performance was found when the bursts of noise were presented during the calculation phase.

Ainley utilized a modified version of this test to study the effects of noise bursts ranging from 47 – 77 dBA on human performance (2012). The six-digit number was displayed for 10 seconds and then replaced by a four-digit number and a text box. The four-digit number and text box remained on screen until the subject submitted their answer. A 15 second intermission followed before the presentation of the next six-digit number. This was carried out for five 20 minute sessions, with each session preceded by a five minute practice period.

As stated in Section 2.1.4, Ainley found no significant relationship between task performance and the level of noise burst. However, a decreasing trend in percentage of correct answers in impulse-presented questions as the impulse level increases was seen.

Frankenhaueser and Lundberg utilized a different type of arithmetic task to study performance under noise of varying intensities (1977). The performance time on a task, developed by Norinder, requiring addition or subtraction of paired one digit numbers (e.g. 1 + 5 =___, 7 – 4 =___) was monitored over two 75 minute sessions. As detailed in Section 2.1.1, subjects experienced 56, 72.5, or 85 dBA continuous white noise during the first session. All subjects then experienced the same 72.5 dBA noise during the second session. Subject feedback regarding comfort and concentration was taken every 25 minutes. Heart rates and performance were also monitored. As also found in Broadbent's test, performance results were significantly lower in the second session for subjects that had experienced a louder noise in the previous session.

Tafalla and Evans utilized the Norinder arithmetic task while studying the role of effort on task performance under various bursts of noise conditions (1997). The bursts of noises, ranging from three to five seconds in length and presented at random intervals from 25 seconds to a minute, had L_{Apk} measurements of 90 dBA and were made from source recordings of traffic, office machinery, and unintelligible speech. Tafalla and Evans reported that noise only had a significant detrimental effect on performance time when the subject's effort was low. They also found that some psychophysiological indexes of stress increase with noise when the subject's level of effort was high.

The arithmetic task is a performance task of interest to this study because it involves components of a digit span task involving memory and simple mathematics involving reasoning. In the past the task has generally been expressed as involving both short term and working memory.

2.3. Subjective Perception under Different Noise Conditions

Annoyance has been shown as a key factor for many people's subjective perception of noise (Zimmer et al. 2008). Zimmer had subjects rate sounds before, during, and after exposure while performing a digit memorization task. The sounds presented were frequency modulated tones, broadband noise, and speech. For the speech signals, ratings of annoyance were significantly the highest during the task, and higher to a lesser degree after the task compared to before the task. Ratings of the less disruptive sounds remained relatively constant across rating times. A second set of tests was carried out increasing the exposure time to the sounds. Longer exposure times resulted in increased annoyance ratings for all sound signals compared to the shorter exposure time.

Lim et al. discussed the effect of aircraft noise compared to background noise levels in a community on annoyance (2008). It was found that annoyance was rated higher in areas with a lower background noise when exposed to equal levels of aircraft noise. This shows that the difference between stimuli noise and background noise is an important factor.

Annoyance is also a factor regularly considered when studying subjective perception to sonic booms (Sullivan et al. 2010, Rathsam et al. 2012, Loubeau et al. 2013). Other subjective perception of sonic boom research has also studied the subjective factors of loudness and startle (Marshall and Davies 2010, Marshall and Davies 2011).

Wang and Novak surveyed subjective perception ratings of loudness, rumble, distraction, and changes in the noise, along with annoyance, to describe subjective perception of assorted HVAC noise conditions (2010). One noise condition, a recording

of a heat pump cycling on and off every 30 seconds, had particularly interesting subjective perception results. In terms of L_{Aeq}, it was the quietest of all six test signals. However, it had the highest annoyance rating, the second highest subjective loudness rating, and the highest, by a large margin, subjective changes in time rating.

2.4. Applications to This Research

The current research applies the arithmetic task used in previous studies by Broadbent (1958), Woodhead (1964), and Ainley (2012) under fluctuating background noise levels as well as noise bursts of varying intensities both with and without a rattle element. The goal is to study any correlations and statistically significant relationships in task performance and subjective perception of the noise for two phases of testing: phase one, a background that varies between a typical background noise level and an elevated level at various time intervals, and phase 2, different noise burst intensities both with and without a rattle element as experienced over typical background noise conditions.

The results of phase one will be compared to previous studies and utilized to help set guidelines for permissible time intervals of fluctuation noise levels. The results from phase two will help gain insight on the effect of rattle noise as perceived in conjunction with bursts of noise such as low-level sonic booms.

Chapter 3: Fluctuations Methodology

The purpose of this study was to evaluate the performance and perception of humans under fluctuating noise conditions of varying time scale. Subjects completed an arithmetic test under different acoustic conditions for four different test sessions and filled out subjective questionnaires over the test environment at the end of each session. Each session lasted a total of thirty minutes and was comprised of three parts: (1) A five minute practice period, (2) a twenty minute test period, and (3) five minutes for a subjective questionnaire.

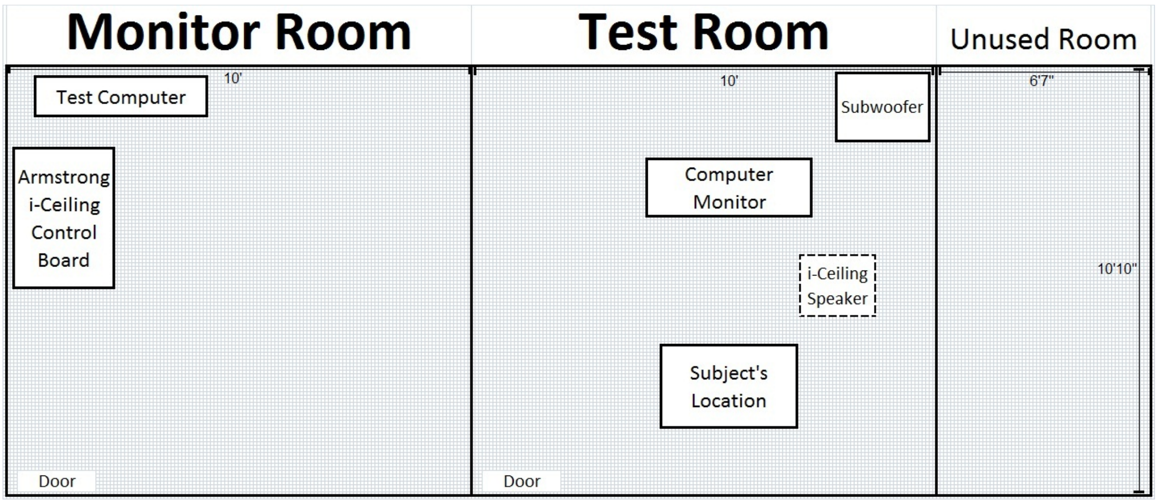
For each test session, subjects experienced a noise environment similar to that produced by a heating ventilation and air conditioning (HVAC) system turning on and off. They were subjected to two levels of background noise, one matching a room criteria rating of RC-29(H) and one matching a room criteria rating of RC-47(RV). The exposure time interval for each level varied across sessions with intervals of two minutes, five minutes, eight minutes, and ten minutes. For example, during the test session with five minute intervals, a subject first experienced five minutes of RC-29(H), then five minutes of RC-47(RV), repeating this pattern for the duration of the twenty minute test period.

3.1. Facilities

3.1.1. Nebraska Test Chamber

All testing was carried out in the old Nebraska Test Chambers at the Peter Kiewit Institute (PKI) on the University of Nebraska campus. The test chambers were acoustically isolated from the nearby rooms via staggered wood stud construction walls with an STC rating of 47. The test room, measuring 10’ x 10’10” x8’, resembled an

office or similar workspace with gypsum board walls, carpet, and acoustical ceiling tiles (ACT). The average mid-frequency reverberation time was measured as 0.25 seconds. The layout of the Nebraska Test Chambers is shown in Figure 3.1.



<u>Fig. 3.1.</u> The layout of the Nebraska Test Chambers showing locations of the subject, test equipment, and loudspeakers used in this study (not to scale). Room height was 8'.

The test room contained a chair with built-in desk, a wireless keyboard to input answers, and a computer monitor to display test questions. The chair was oriented in the room so that a subject's head was 4'8" from the wall shared with the monitor room, 3'6" from the back wall, and 3'6" off the ground. It should be noted that this is the location of the sound level meter used for the measurements mentioned in Section 3.2.2. This is considered the approximate subject head location due to variations in subject height and head movements during testing. Subject's head position was not monitored during testing. The subject sat approximately 4' away from the 23.5" computer monitor. For legibility, all fonts displayed on the monitor were sized to be at least 36 point.

The loudspeakers used to implement the noise conditions were an Armstrong i-ceiling loudspeaker and a JBL Northridge ESeries subwoofer. The i-ceiling loudspeaker resembles an ordinary ACT and was situated next to a diffuser in the ceiling to give the

perception that the sound was coming from the diffuser. The subwoofer, covered in fabric and situated in the corner of the room, provided the low-frequency content for the noise signals. Two loudspeakers, utilized in an unrelated test, were also in the room. They were covered in fabric and the subjects were told to ignore them for the test. A photograph of the interior of the test room is shown in Figure 3.2.

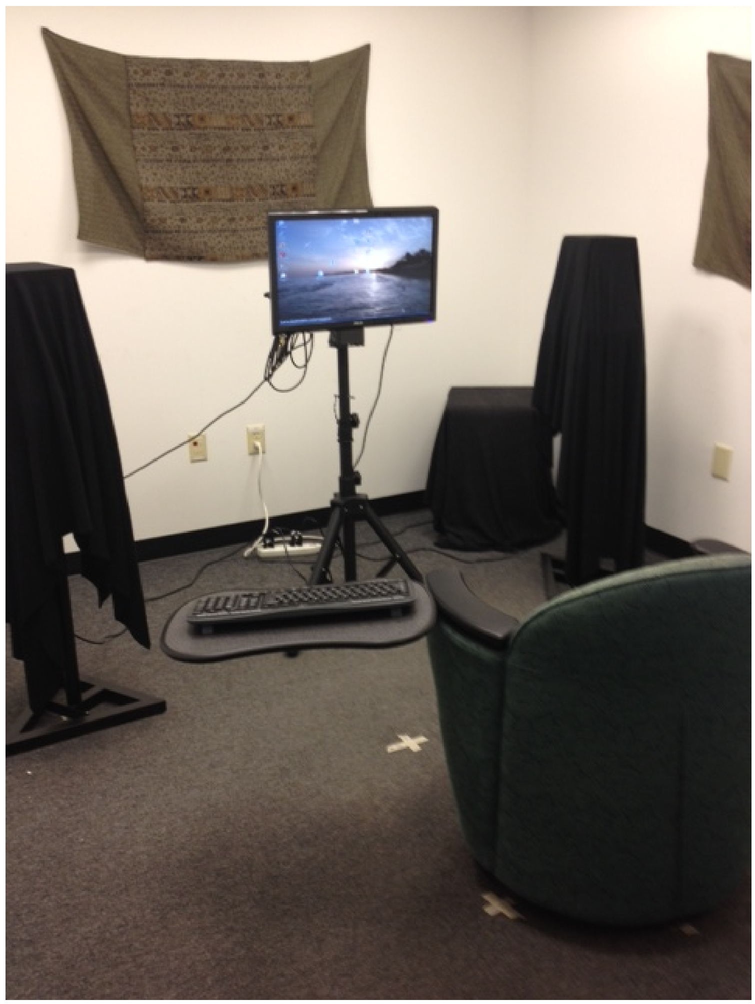


Fig. 3.2. A picture of the interior of the test room.

The test room was bordered by the monitor room to the left and an unused room to the right. The monitor room contained the test computer and the power amplifier for the loudspeakers. It was also the room that the test monitor worked from and in which the subjective questionnaires were filled out.

The test room was controlled for temperature as best as possible. However, it was a non-climate controlled space. A portable air conditioning unit was brought in between sessions when possible. An average of 77.1 ^{o}F was measured across all test sessions.

3.1.2. Sound and Computer Systems

A diagram of the Nebraska Test Chamber testing and sound system is shown in Figure 3.3. The test computer ran the arithmetic test and generated the sound signals. With the computer and loudspeaker controls in the monitor room, the loudspeakers were the only sources of noise in the test room.

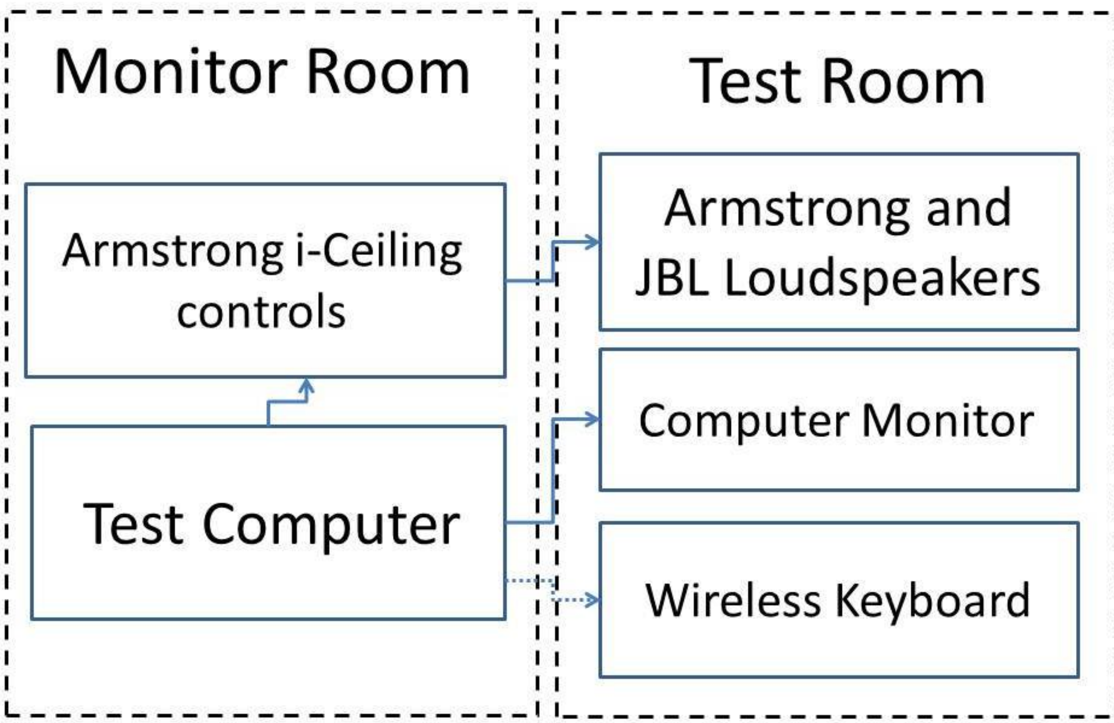


Fig. 3.3. A diagram of the Nebraska Test Chamber system showing both the testing system and the sound system.

3.2. Experimental Methods

This section reviews the experimental methodology and is separated into four subsections: (1) creation of the signals used, (2) recording and measurement procedures used to analyze each signal, (3) procedure involved behind the creation and running of the test sessions, and (4) statistical analysis.

3.2.1. Signal Creation

Eight sound files of two types were created for this study. One type, the broadband background noise, was synthesized to an approximate room criteria rating of RC-30(N). This level was picked to be representative of a quiet workspace. A ten second file was looped and calibrated using the equalizer in CoolEdit until a RC-30(N) was measured in the test chamber while being played back over the JBL subwoofer and the Armstrong i-ceiling. Measurements made after the testing sessions showed that this signal actually measured as RC-29(H) when played back in the room.

The second type, or the elevated background noise level, was synthesized to an approximate room criteria rating of RC-50(V) – selected to be representative of a workspace with a loud HVAC system running. This ten second signal was also calibrated using the equalizer in CoolEdit while being played over the JBL subwoofer and Armstrong i-ceiling in the testing room. Initial measurements found a RC-48(V), while final measurements yielded a RC-47(RV) when played back in the room.

Each test session alternated between a RC-29(H) and RC-47(RV) sound file of equal duration. Each of the four sessions is named after the length of one sound file: “2 minutes”, “5 minutes”, “8 minutes”, and “10 minutes”. The ten second .wav files were extended to create eight extended length files. A two minute, five minute, eight minute,

and ten minute RC-29(H) .wav file was created, as well as RC-47(RV) files of matching durations. For a single test session, a playlist was created with an RC-29(H) file and its matching length RC-47(RV) file and then played back on a loop with WinAmp. A visualization of these four sessions is shown in Figure 3.4.

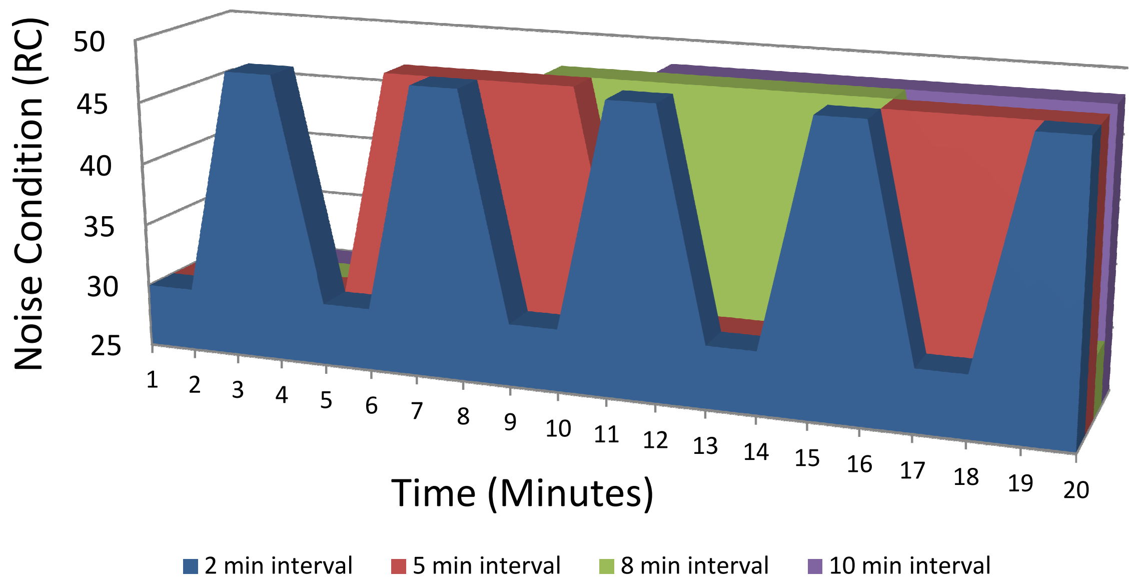


Fig. 3.4. Visualization of the noise conditions presented in the four fluctuation tests.

Looped as is, an abrupt transition would occur between the two sound files. To create a natural sounding transition, an envelope was applied to the RC-47(RV) files using CoolEdit. A spline curve relating to a percentage of attenuation of the original signal was selected for this envelope and was applied to both ends of the sound file. A CoolEdit screenshot of one resulting signal is shown in Figure 3.5. This envelope yielded a more realistic transition resembling an HVAC system turning on and off.

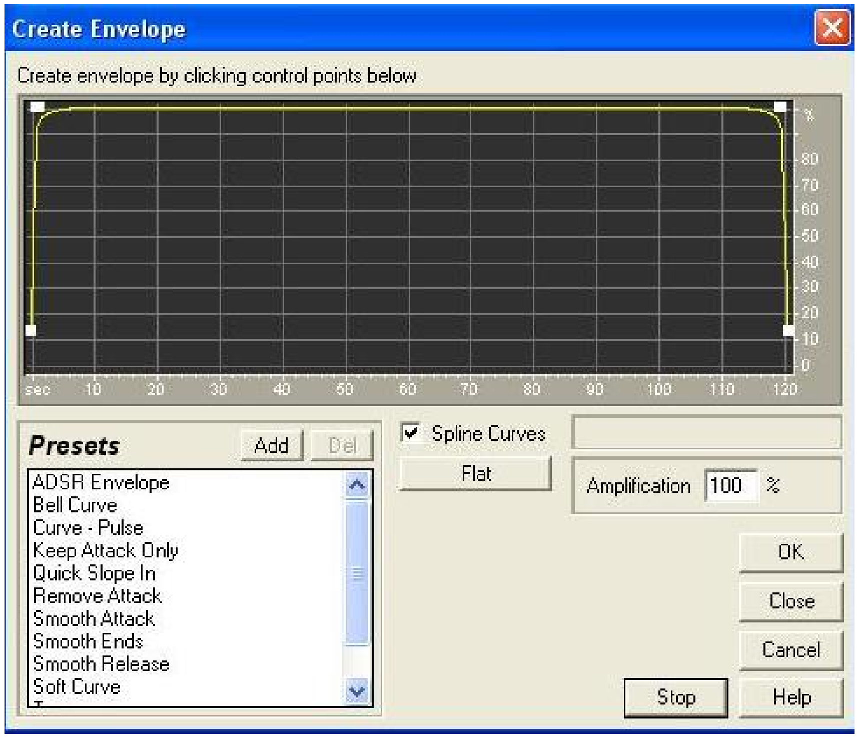


Fig. 3.5. Screenshot of envelope applied using CoolEdit to RC-47(RV) .wav files

3.2.2. Signal Recordings and Measurements

All signals were recorded and measured in the test room at the head position of the subject (mentioned in Section 3.1.1) using a Larson-Davis 824 sound level meter (SLM). The recording and measurement procedures for the signals are reported in the following subsections.

3.2.2.1. Signal Recordings

All signals, as played back in the room over the Armstrong i-ceiling loudspeaker and JBL Northridge subwoofer, were recorded to .wav files using Presonus Studio One recording software for archival purposes. The above mentioned SLM was used as a microphone and connected to a Presonus AudioBox44VSL external sound card. The recording computer was kept in the monitor room as to not add extra noise to the

recording. A diagram of the audio playback system and recording system is shown in Figure 3.6.

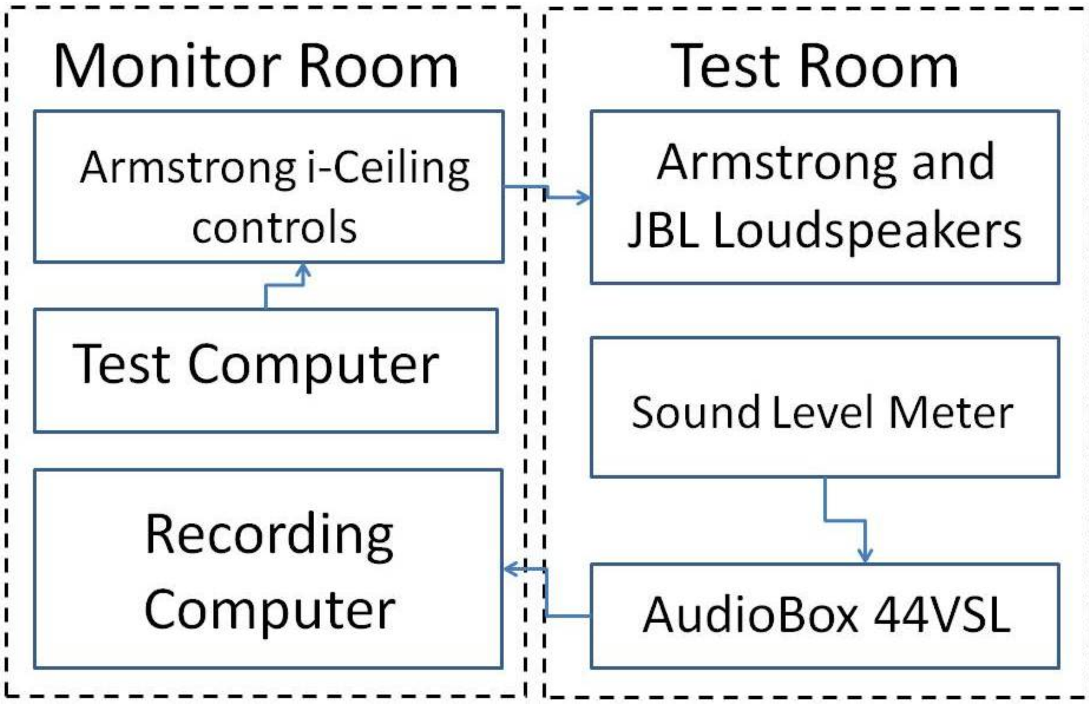


Fig. 3.6. A diagram of the Nebraska Test Chamber system showing the equipment used for recording test signals in the test room.

A single period of both the RC-29(H) and RC-47(RV) for each test session was recorded, including both transitions. A 1 kHz tone generated in CoolEdit was also recorded in the room using the same recording settings over an 8 second period. The 1 kHz tone was measured to be 45 dB at the 1 kHz octave band in the room using the Larson-Davis 824 SLM averaged over a 20 second time interval. This SLM measurement was used to calibrate the other recorded sound files. All recordings used a 44.1 kHz sampling rate.

3.2.2.2. Signal Measurements

Signals were also measured using the Larson-Davis 824 SLM using the settings shown in Figure 3.7. Sequentially played five minute RC-29(H) and RC-47(RV) signals

were measured for a recording period of 10 minutes. A two minute time period of each RC-29(H) and RC-47(RV) was analyzed.

Sound Level Meter / RTA Settings	
Bandwidth:	1/3
Detector:	Fast
Weighting:	Flat
Peak-1 Weighting:	Flat
Second Display:	TWA
Gain:	0
RTA Detector:	Fast
RTA Weighting:	Flat
Filter Range	12.5-20k

Intervals	
Intervals:	Enabled
Interval Time Sync:	No
Interval Save Ln:	Yes
Interval Save Ln Table:	No
Interval Auto Stop:	Yes
Interval Period:	0:10:20
Interval Threshold:	0
Interval Exchange Rate:	3 dB
Interval Spectra Option:	At Max

Ln	
Ln:	Enabled
Ln Start Level:	15 dB
Spectral Ln Option:	Interval
Ln Percentiles	
Ln Percentiles	
L 1.0	
L 10.0	
L 50.0	
L 90.0	
L 95.0	
L 99.0	

Time History	
Time History:	Enabled
Time History Period:	4
Time History Units:	1/32 seconds
Resolution:	0.1 dB

Fig. 3.7. A list of settings used for measurement of signals with a Larson-Davis 824 SLM.

Measurements were made every 125 ms – the shortest measurement interval available on this SLM. The SLM was set to “fast” mode, and 1/3 octave band data was recorded. This data was then exported to Excel for calculations.

3.2.3. Test Session Procedure

This section details the preparation and implementation of testing procedures. It contains three subsections: test session scheduling, test session design and procedure, and recruitment and orientation procedure.

3.2.3.1. Test Session Scheduling

The overall test consisted of an orientation session and nine regular test sessions – all of which were 30 minutes long. The nine test sessions were broken up into two groups: five noise burst sessions and four fluctuating noise sessions. Subjects first experienced all five of the noise burst sessions before moving on to the fluctuation sessions. Subjects were only allowed to participate in one session per day. However, a few exceptions ended up being made due to scheduling issues. The sessions that did occur within the same day were separated by more than four hours. The noise burst sessions were part of a previous study and will not be discussed further (Ainley 2012). The four fluctuation tests were the previously discussed "two minutes", "five minutes", "eight minutes", and "ten minutes" tests.

The test presentation order was determined with a Latin square design to avoid a test order bias. For the Latin squares design, there were four test sessions and 30 subjects. Seven 4x4 squares were used for the first 28 subjects. The order for the last two test subjects was determined with a random order function in Microsoft Excel.

3.2.3.2. Test Session Design and Procedure

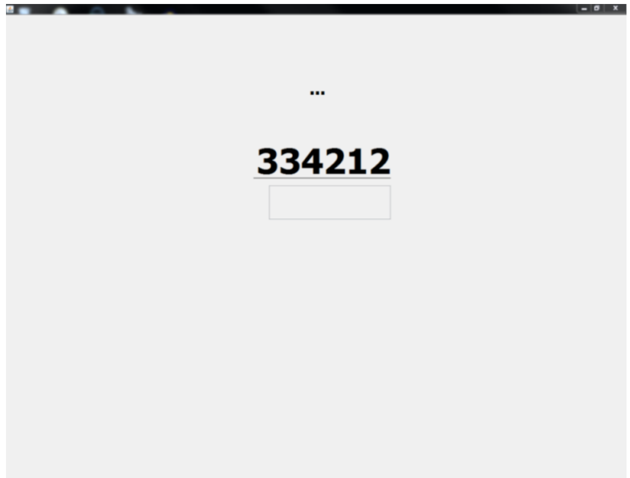
This subsection details the design and procedure used for the arithmetic test. Every test session consists of a five minute practice period, a twenty minute test period, and five minutes allotted to fill out a subjective questionnaire. Each five minute practice period

involved its own unique set of questions. Scores were not recorded during this time, as it was just for the subject to become reacquainted with the arithmetic task. After five minutes, the subjects were notified that the practice session had completed and prompted to begin the main test.

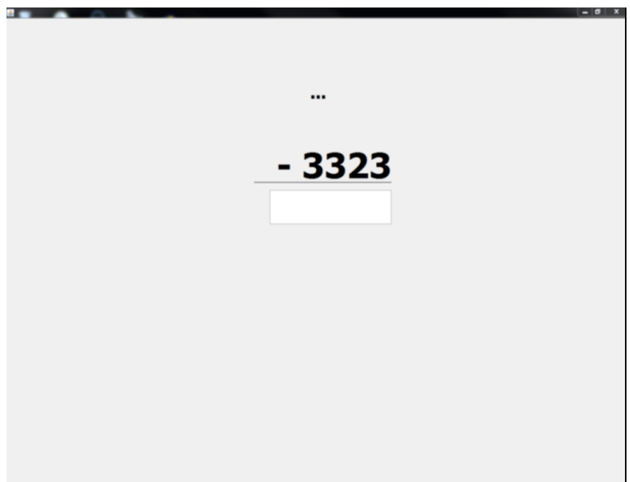
The arithmetic task required the subject to find the difference between a six-digit number and a four-digit using only their working memory. No tools, i.e. a calculator or pen and paper, were allowed in the test room. This arithmetic task was designed based off of previous tests by Broadbent (1958) and Woodhead (1964).

First, the six-digit number was presented on the screen for ten seconds as shown in Figure 3.8(a). Next, the six-digit number was replaced by a four-digit number and a single-row text box as shown in Figure 3.8(b). This remained on the screen until the subject entered and submitted their answer in the text box. Once an answer was submitted, there was a 15 second intermission before the presentation of the next test question.

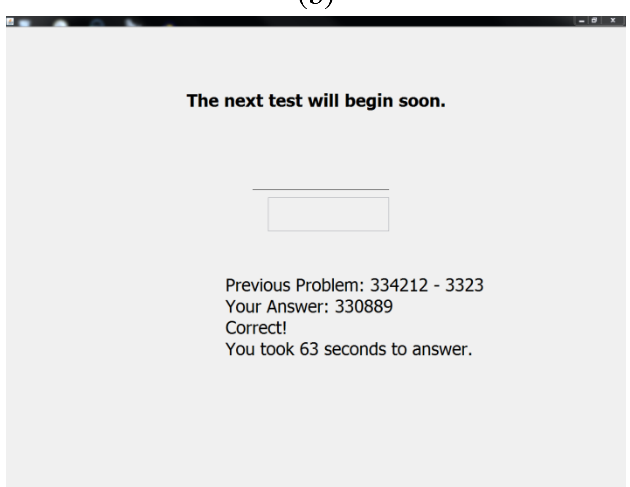
During the practice portion of the test session, subjects were given feedback on their performance on the previous problem during the 15 second intermission as shown in Figure 3.8(c). The feedback included their answer to the previous problem, the correct answer (if different from the subject's answer), and the response time, i.e. the length of time it took the subject to answer the question starting from the presentation of the four-digit number. This feedback was not presented during the main testing period. Instead, the sentence "The next test will begin soon." was displayed until the final three seconds of intermission when the subject was prompted with a "ready, set, go!" message. There were no clocks or other timers present in the test room.



(a)



(b)



(c)

Fig. 3.8. Screenshot from practice portion of arithmetic test program (a) displaying the first number, (b) displaying the second number, and (c) after answer submission with feedback.

Only digits 1, 2, 3, and 4 were presented in the four and six digit numbers. Difficulty was controlled during the test sessions. Difficulty, as applied to this test, is defined as the number of times a subject has to "borrow" during a subtraction problem. In subtraction, borrowing is when the top digit in a column is smaller than the bottom digit in the column. The subject then borrows from the top digit in the column to the left. For the practice session, the difficulty started off as easy (no borrows) and ramped up in difficulty. The difficulty remained constant at three borrows for the main test. Due to the specific requirement of controlling borrows, it was necessary to write all test questions from scratch.

Unique test question sets were created for each practice session and each main session. These question sets were presented with the same noise condition to each subject. Woodhead's 1964 tests found that an average subject could complete around 22 questions in a 20 minute session. Therefore, question sets were made long enough so subjects would not run out of test questions in the fixed amount of time. The five minute practice session sets contained 20 questions, while the 20 minute tests contained 45 questions. No subject was able to complete all questions in a practice or regular session within the allotted time.

At the beginning of each test session, subjects were told: *"Remember that you may experience some environmental fluctuations in temperature, lighting, and noise during today's test. Also, remember for this experiment, we are mainly interested in memory, accuracy, and speed."* Subjects were also reminded to completely shut down any cell phones and encouraged to leave any bags or additional items in the monitor room during testing.

A Java program automatically conducted the arithmetic test. The program displayed the test questions for the required length of time and recorded the subject's response. It also saved time stamps for all events, e.g. presentation of first number, presentation of second number, and when the subject submitted an answer. The program required the test monitor to upload .txt files of the arithmetic problems, select a location folder for and name the output data .csv worksheet.

Test questions were written in the form of a .txt file that could be imported into the test program. Each test problem was written on a single row with a comma separating the six-digit and four-digit numbers. An example text file can be seen in Figure 3.9. A unique text file was required for each five minute practice, each twenty minute main test portion of each session, and two for the orientation session.

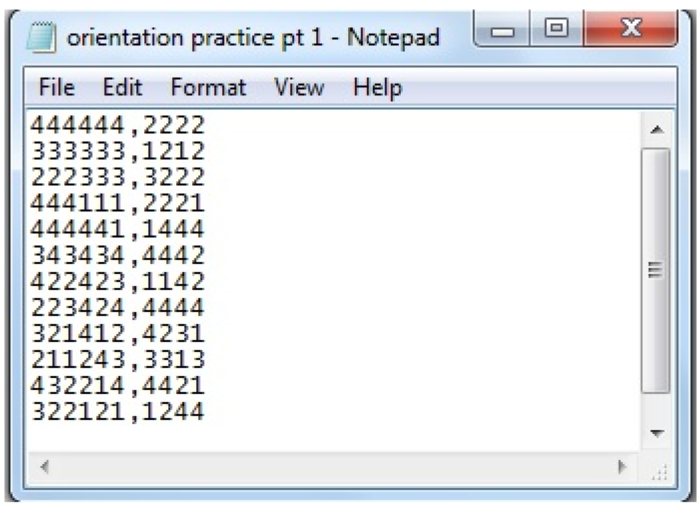


Fig. 3.9. Screenshot of orientation practice questions text file.

The final five minutes of each test session were allotted for the subjects to fill out a subjective questionnaire about their experience in the room on that day (Figure 3.10). Space was also provided for the subjects to add any additional comments about that specific session.

During their final test session, subjects were additionally asked to complete a noise sensitivity questionnaire – taken from the reduced version of the Noise Sensitivity Questionnaire (NoiSeQ) developed by Schutte et al. (2007). This questionnaire is shown in Figure 3.11. Total noise sensitivity for each participant was calculated based on the information provided on the questionnaire. Again, subjects were allowed to add any additional comments, this time about the overall testing experience, on the backside of this questionnaire.

Questionnaire

Please rate the conditions in this room based on the following attributes. Please do not skip any questions.

Please base your ratings only on the session today. Do not base your ratings on any previous sessions.

1	Air freshness	1 fresh air	2	3	4	5	6	7 stale air
2	Temperature comfort	1 too cold	2	3	4	5	6	7 too hot
3	Air movement	1 circulating	2	3	4	5	6	7 stuffy
4	Loudness of noise	1 very quiet	2	3	4	5	6	7 very loud
5	Changes in noise over time	1 not changing	2	3	4	5	6	7 changing a lot
6	Rumble of noise	1 not rumbly	2	3	4	5	6	7 very rumbly
7	Annoyance to noise	1 not annoying	2	3	4	5	6	7 very annoying
8	Noise distractions	1 not a problem	2	3	4	5	6	7 very distracting
9	How bright lights are	1 not too bright	2	3	4	5	6	7 too much light
10	Comfort of work station	1 comfortable	2	3	4	5	6	7 uncomfortable

Fig. 3.10. A copy of the subjective questionnaire that participants completed at the conclusion of each test session.

Final Questionnaire

Please provide your gender and age.

Gender: Male Female

Age: ______________

Please rate each statement in order. Please do not skip any questions. If possible, imagine yourself in each situation and response accordingly without spending too much time considering if you agree or disagree with a given statement. We are looking for your personal opinions. There are no correct or incorrect responses.

		Strongly agree	Slightly agree	Slightly disagree	Strongly disagree
1	I need an absolutely quiet environment to get a good night's sleep.	1	2	3	4
2	I need quiet surroundings to be able to work on new tasks.	1	2	3	4
3	When I am at home, I habituate to noise quickly.	1	2	3	4
4	I become very agitated if I can hear someone talking while I am trying to fall asleep.	1	2	3	4
5	I am very sensitive to neighborhood noise.	1	2	3	4
6	When people around me are noisy I don't get on with my work.	1	2	3	4
7	I am sensitive to noise.	1	2	3	4
8	My performance is much worse in noisy places.	1	2	3	4
9	I do not feel well rested if there has been a lot of noise the night before.	1	2	3	4
10	It would not bother me to live in a noisy street.	1	2	3	4
11	For a quiet place to live I would accept other disadvantages.	1	2	3	4
12	I need peace and quiet to do difficult work.	1	2	3	4
13	I can fall asleep even when it is noisy.	1	2	3	4

Fig. 3.11. A copy of the noise sensitivity questionnaire that participants completed at the conclusion of their final test session.

3.2.3.3. Recruitment and Orientation Procedure

Subjects were recruited by fliers posted on the University of Nebraska – Omaha campus.

The first session each subject participated in was an orientation session. Subjects were first presented with a PowerPoint presentation covering the instructions of the test procedure. Next, the subjects participated in a hearing screen and a practice arithmetic test session.

An audiometer was used to test hearing thresholds of both the left and right ear individually and was administered in the test room. Pure tones of each octave band between 125 and 8000 Hz were individually presented, first at 30 dB hearing level (HL). If the subject failed, the level was increased by 5 dB. If the subject correctly triggered that s/he heard the signal, the level decreased by 5 dB. This continued until a threshold was found where the subject can no longer hear the tone or 15 dB HL was reached. Subjects were required to have a threshold at or below 25 dB HL in each ear for all tested tones.

Next, the subject was introduced to the testing program. They were taught how to use the wireless keyboard for the test only using the number pad, arrow keys, enter, and backspace. They were reminded that they had to perform the task from memory and that they are only able to input answer in the single line text box.

Next, the subject took a five minute practice session with the test monitor present to answer any general questions about the test. The monitor also made sure that the subject was able to answer at least two questions correctly and that the subject felt comfortable with the task they were asked to perform. If necessary, the subjects were

allowed to take a second practice test to meet the two necessary conditions. If they were still unable to meet these conditions, they were asked not to participate in the rest of the testing sessions.

Five subjects were asked to not participate after not meeting these conditions during orientation. One subject was asked to not participate due to not meeting the hearing screen requirements. Five more subjects also dropped out further into testing due to other scheduling issues.

3.2.4. Statistical Analysis

Subjects' performance and perception results were statistically analyzed using Microsoft Excel and SPSS. The percent of questions answered correctly and the response times, or the time from presentation of the four-digit number to the subject submitting an answer, are the two types of performance data. Perception data are considered to be the subjective questionnaire responses, specifically those related to the acoustic conditions of the test environment.

Non-parametric tests were required for most cases; however, both parametric and non-parametric tests are used and presented. There are three requirements to be able to perform parametric tests: data must be measured at an interval (even point scale) or ratio level (like interval data, but with meaningful ratios between points on the scale), homogeneity of variance, and normal distribution of data. Homogeneity of variance means that the variance in all experimental conditions is roughly the same (Field and Hole, 2003).

3.2.4.1. Standard Error of the Mean

Standard error of the mean (SE) is a standard deviation of the sample means and is used to represent how accurate a sample can be. SE is reported in the form of error bars in results graphs in the following chapter. As the error bars grow wider (representing a larger SE), the variability of the sample means increases. SE is found by Equation 3.1:

$$SE = \frac{s}{\sqrt{N}} \tag{3.1}$$

where *s* is the sample standard deviation and *N* is the sample size (Field and Hole, 2003).

3.2.4.2. Parametric Tests

General relationships between a single dependent and independent variable were determined using Pearson Product Moment Correlations and linear mixed model analysis in SPSS. An example of a general relationship is the relationship between performance scores and noise condition presented in each session. All significant relationships were reported using these two statistical test methods.

The Pearson Product Moment Correlations reports the correlation, r, between the two variables, while the linear mixed model reports the F value with the degrees of freedom, df, of the numerator and denominator. These are reported, along with their respective significances, in the following format: $F_{dfn,dfd}$ =____, r = ____, where *dfn* is the numerator df and *dfd* is the denominator df as reported by SPSS.

A repeated measures ANOVA was used to compare a dependent variable to multiple independent variables. An example of this comparison is the relationship in loudness perception rating to the four noise conditions, gender, age, and/or noise sensitivity rating. Each repeated measures ANOVA test statistic, F, is reported with

significance in the following format: F(df,N) = ____, where df is degrees of freedom and N is sample size. The effect size, ω, was found by taking the square root of Equation 3.2:

$$\omega^2 = \frac{MS_M - MS_R}{MS_M + \left((n-1) * MS_R\right)} \qquad (3.2)$$

where MS_M is the mean sum of squares, MS_R is the mean squared error, and n is the sample size. When the F test statistic was found to be significant, Bonferroni post hoc tests were implemented to find significant differences between the test sessions (Field and Hole, 2003).

3.2.4.3. Non-Parametric Tests

In most cases, data were found to be not normally distributed, meaning that non-parametric tests were appropriate. The parametric tests described above may not be accurate when performed on non-normally distributed data because of a possible inaccurate P value. Therefore, some non-parametric tests were performed and compared to the parametric tests. A Spearman Correlation with significance, r, is reported in place of the Pearson Product Moment Correlation for general relationships between a single dependent and independent variable for non-parametric data.

A Friedman's ANOVA, which utilizes a Wilcoxon test, is used in place of the repeated measures ANOVA to compare a dependent variable to a single independent variable with multiple levels. For example, annoyance ratings are compared to the four test signals. Each Friedman test statistic is reported with df and significance in the following format: χ^2(df) = ____.

To find exactly where there are statistical differences between sessions, a Wilcoxon test is utilized with a Bonferroni correction. The Wilcoxon test statistic, T, is reported along with the effect size, r. Effect size is found using Equation 3.3:

$$r = \frac{Z}{\sqrt{N}} \qquad (3.3)$$

where Z is the z-score produced by SPSS and N is the total number of observations compared (Field and Hull, 2003).

3.2.4.4. Statistical Power Analysis

A power analysis is also utilized to determine the probability of each result presenting a genuine effect. This is reported as an observed power from 0 to 1, as reported by SPSS with $\alpha = 0.05$, and it is reported for each repeated measures ANOVA test.

Chapter 4: Fluctuations Results and Discussion

This chapter presents the analyzed results of the measured test signals, arithmetic task performance, and subjective perception. Test signals are analyzed and reported as discussed in Chapter 3. Task performance and subjective perception results are reported and analyzed using the statistical analysis methodology discussed in Chapter 3.

4.1. Demographic Results

27 subjects participated in this study consisting of 15 males and 12 females. The average subject age was 24 years old, with ages ranging from 19 to 38. Noise sensitivity questionnaires were also filled out during the subjects' final session. The responses were weighted and calculated in to sleep, work, residential, and total noise sensitivity percentages utilizing Schutte et al.'s NoiSeQ-R survey (2007). A histogram of subject responses is shown in Figure 4.1. and results to individual questions are shown in Figure 4.2. Total noise sensitivities ranged from 8% (not very sensitive) to 78% (very sensitive) with an average of 47.3% and a standard error of the mean of 3.5%. Noise sensitivity, gender, and age were considered as additional variables when analyzing complex relationships between noise conditions, task performance, and subjective perception and will be further discussed later in this chapter.

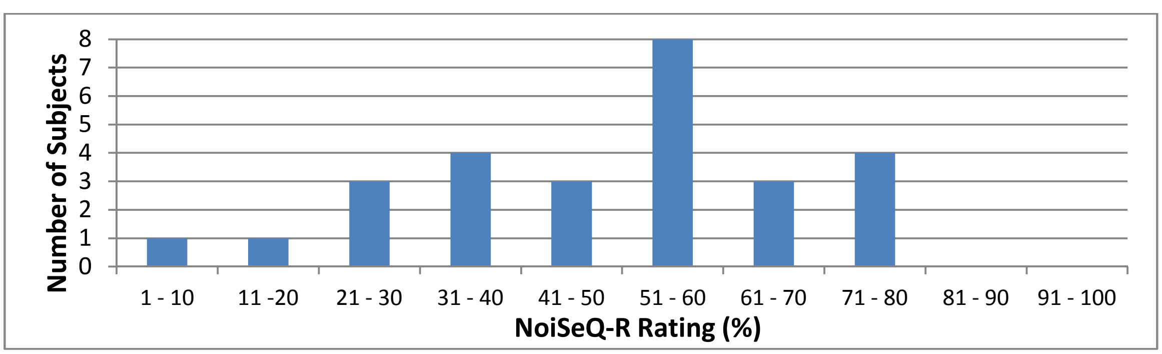


Fig. 4.1. Histogram of NoiSeQ-R responses.

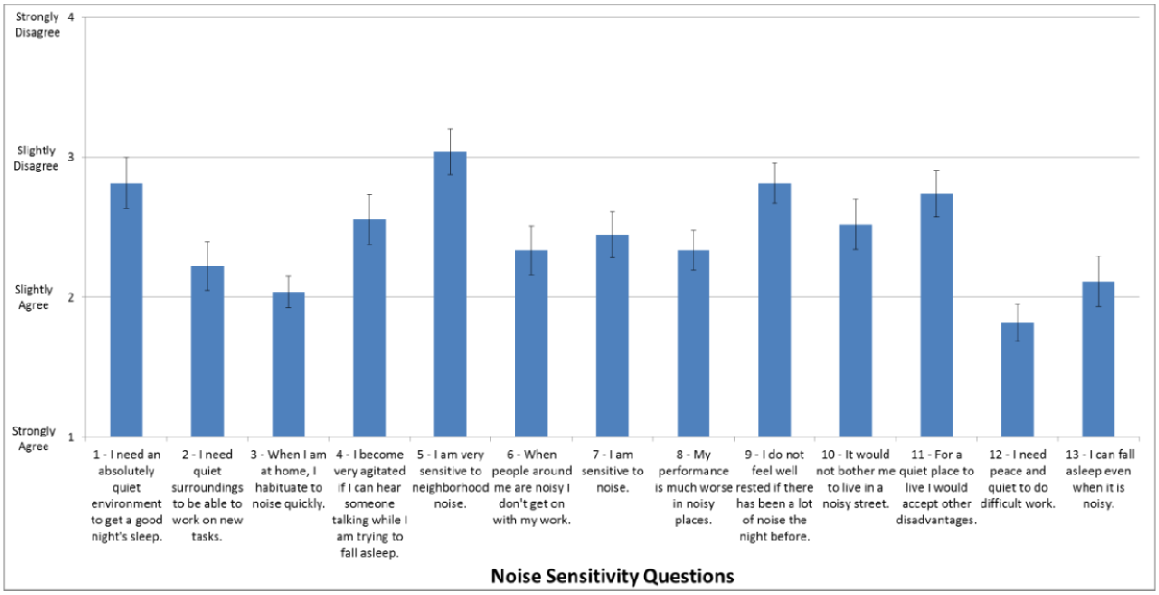


Fig. 4.2. Results of the NoiSeQ-R. Error bars represent the standard error of the mean.

4.2. Signal Results

The background noise level of the test room, measured as the equivalent continuous sound level (L_{eq}) over 10 seconds on a Larson-Davis 824 sound level meter, is shown in Figure 4.3. Although the noise level was too low to generate a room criteria (RC) reading on the sound level meter, it can be reported as an NCB-22 (H). This is too quiet for the purposes of this study, which is one reason why a generated background noise .wav file was implemented.

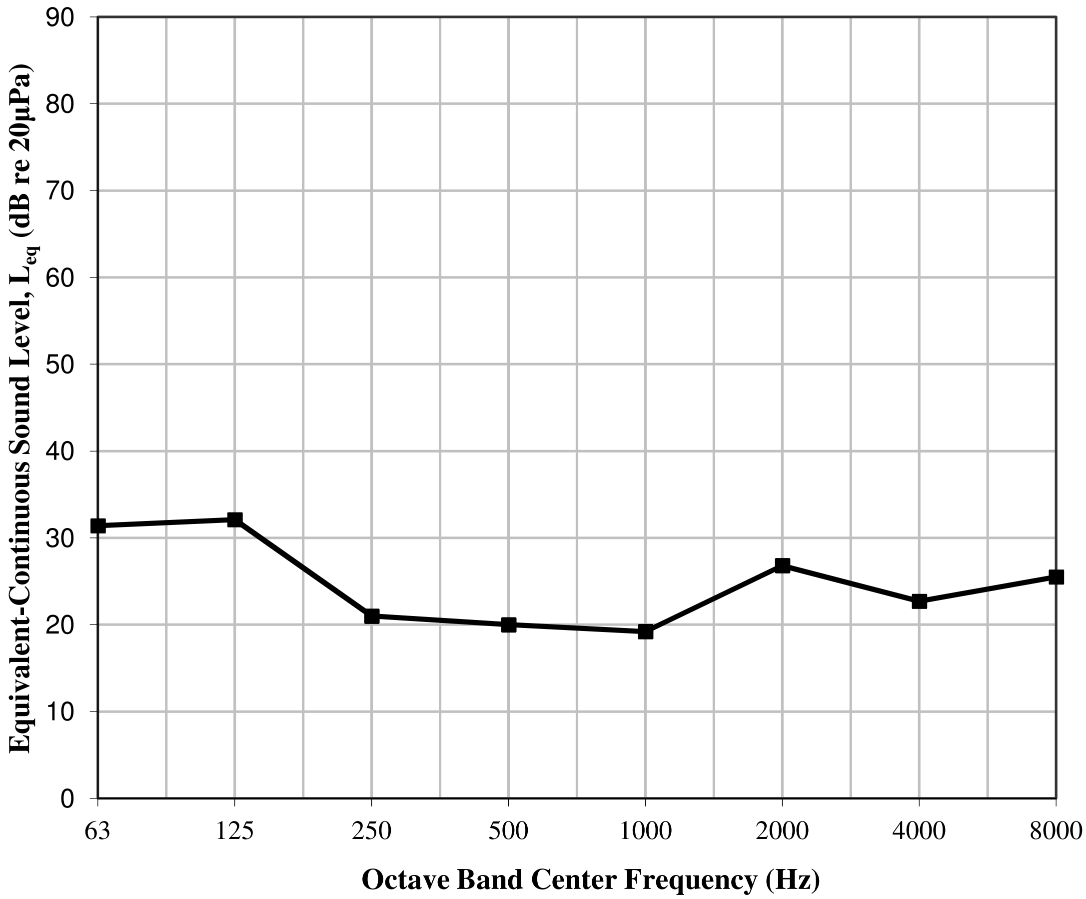


Fig. 4.3. Measurement of Leq across frequency in test room. Results yield an NCB-22 (H)

A higher ambient background noise level was generated as discussed in Section 3.2.1.2. Figure 4.4 reports the measured ambient level in the room with the generated noise as measured with a Larson-Davis 824 sound level meter over a 2 minute measurement period. The goal was RC-30, with a final result of RC-29 (H), 37 dBA.

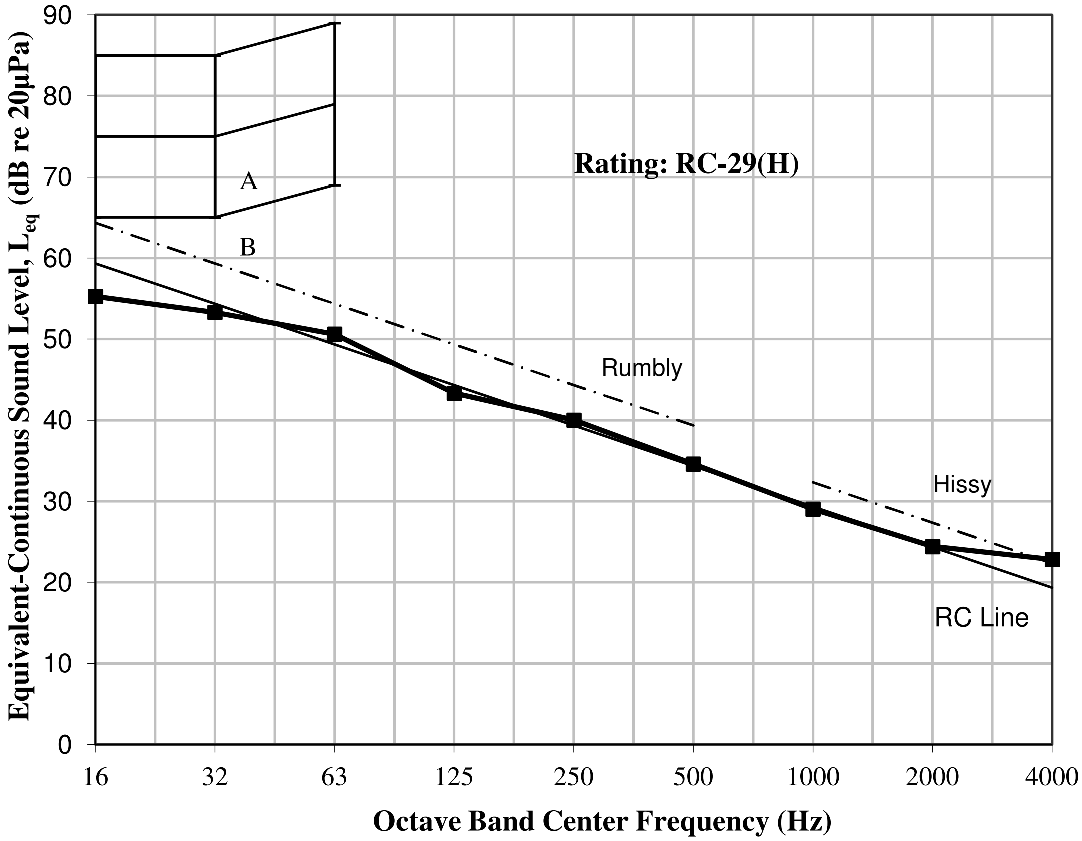


Fig. 4.4. Measurement of Leq across frequency of the ambient BNL .wav when played back in test room. Results yield an RC-29 (H).

An elevated background noise was also generated as discussed in Section 3.2.1.2. Figure 4.5 shows the measured ambient level in the room with this signal as measured with a Larson-Davis 824 sound level meter over a 2 minute measurement period. The final result was an RC-47 (RV), 56 dBA.

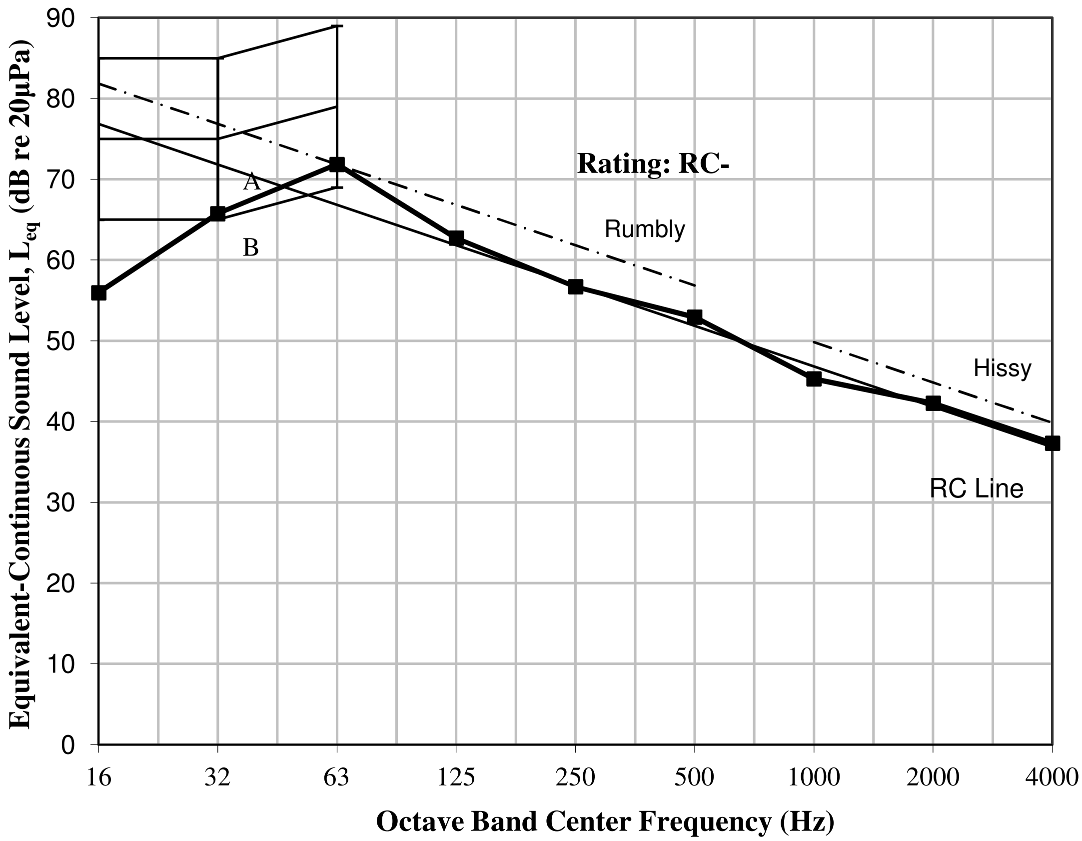


Fig. 4.5. Measurement of L_{eq} across frequency of the louder ambient BNL .wav when played back in test room. Results yield an RC-47 (RV).

4.3. Task Performance Results

Task performance was measured via the total percentage of correct answers and the average response time, in seconds, for each question. Statistical analysis was performed utilizing SPSS statistical software as discussed in Section 3.2.4. Results were tested for normal distribution via a Kolmogorov-Smirnov test and all were found to be non-normally distributed. Because of this, Friedman ANOVA and Spearman correlation coefficients are used in addition to the Pearson coefficient. Wilcoxon tests were used to analyze any relationships between noise conditions. A Bonferroni correction was applied, and all effects reported at a 0.05 level of significance (Field and Hole 2003).

Although it is a parametric test, repeated measures ANOVA results are also reported to support the non-parametric test results. Observed power, at $\alpha = 0.05$, is reported for each.

4.3.1. Task Performance Results across Noise Conditions

The overall task performance results across all analyzed test sessions are shown in Figures 4.6 and 4.7 as well as Tables 4.1 and 4.2. All results exhibited a non-normal distribution, as concluded by a Kolmogorov-Smirnov test. Therefore, the non-parametric Friedman ANOVA test was used to analyze these relationships, and Wilcoxon tests were used to further analyze the relationships between each noise condition.

The overall percentage of correct answers were significantly affected by the different noise conditions, $\chi 2(3) = 9.13$, $p<0.05$. A general trend can be seen in Figure 4.6 where subjects' percentage of correct responses decreases as the fluctuation times decrease. Also, the total percent correct on the ten minute interval sessions (87%) was found to be statistically significantly different than the total percent correct on the two minute interval sessions (80%).

The average response time was also significantly affected by the different noise conditions, $\chi 2(3) = 11.93$, $p<0.05$. However, no trend is apparent in the subjects' average response time compared to interval length. The average response times fell within 15 to 17 seconds across all four interval lengths. Also, the average response time on the five minute interval sessions (15.2 seconds) was found to be statistically significantly different than the average response time on the two minute interval sessions (16.8 seconds) and the ten minute interval sessions (17.0 seconds).

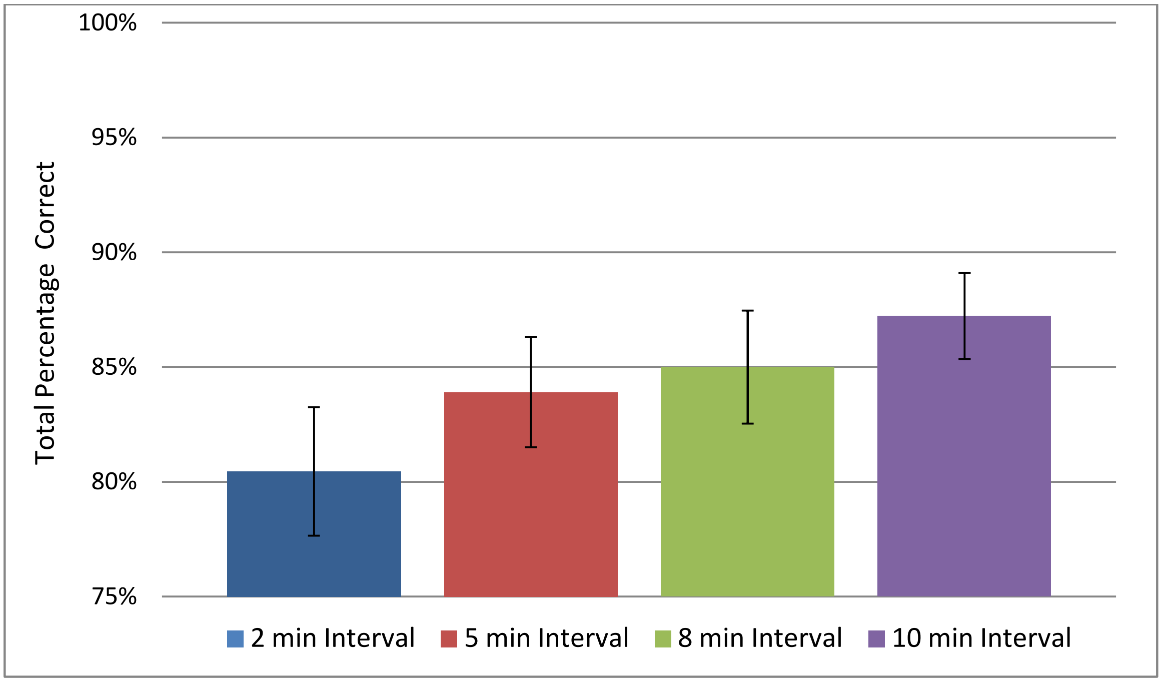


Fig. 4.6. Overall percentage of correct answers for each test session. Error bars represent the standard error of the mean.

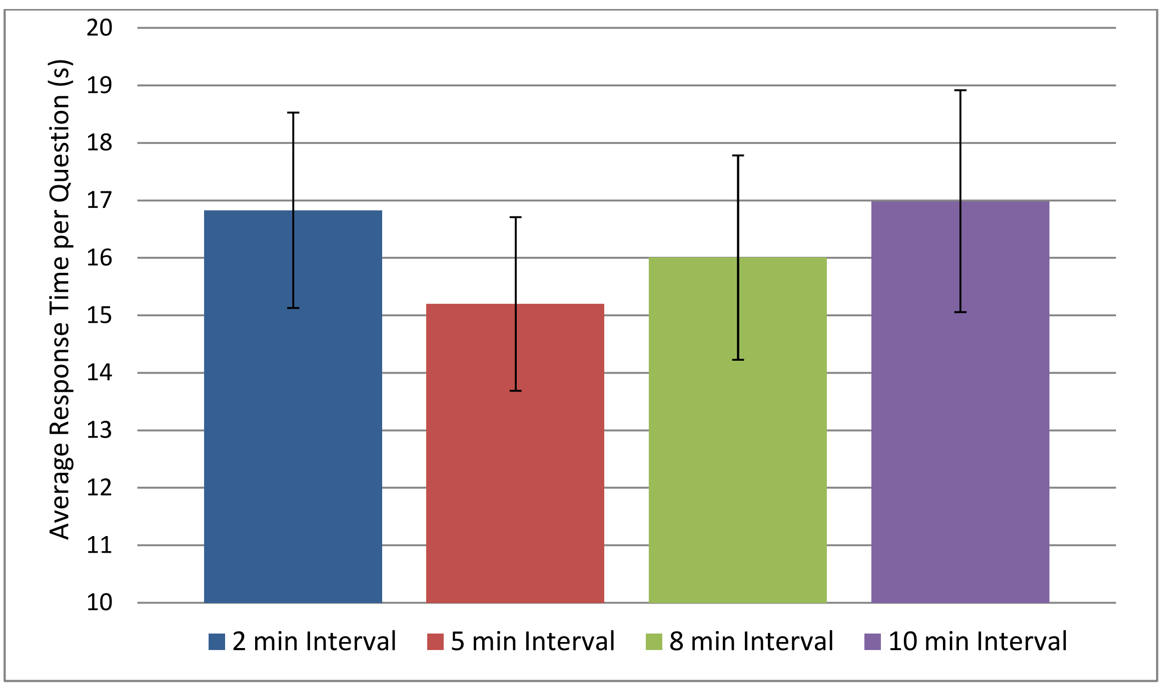


Fig. 4.7. Average response time, in seconds, for test questions in each test session averaged across all test sessions. Error bars represent the standard error of the mean.

Table 4.1. Wilcoxon Results between Noise Condition and Total Percentage Correct. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.005 level of significance.

Task Total % Correct				
		5 Minute interval	8 Minute Interval	10 Minute Interval
2 Minute Interval	T	*142.00*	*97.00*	38.00
	sig	*ns*	*ns*	**
	effect size	*-0.15*	*-0.27*	-0.41
5 Minute Interval	T		*128.50*	114.00
	sig		*ns*	ns
	effect size		*-0.16*	-0.25
8 Minute Interval	T			*114.50*
	sig			*ns*
	effect size			*-0.18*

** = Indicates the mean ranks between noise conditions is significant, p<0.005, ns = not significant.

Table 4.2. Wilcoxon Results between Noise Condition and Average Response Time. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.005 level of significance.

Task Total Average Time				
		5 Minute interval	8 Minute Interval	10 Minute Interval
2 Minute Interval	T	*87.00*	*107.00*	168.00
	sig	***	*ns*	ns
	effect size	*-0.33*	*-0.27*	-0.07
5 Minute Interval	T		*136.00*	56.50
	sig		*ns*	**
	effect size		*-0.17*	-0.43
8 Minute Interval	T			*110.50*
	sig			*ns*
	effect size			*-0.26*

** = Indicates the mean ranks between noise conditions is significant, p<0.005, ns = not significant.

4.3.2. Comparisons of Task Performance to Subjective Perception

Task performance and subjective perception were compared using Pearson and Spearman Correlation Coefficients and a linear mixed model analysis with results shown in Table 4.3. Significant relationships were found for all combinations of performance

and perception, p < 0.05, except between perceived distraction rating and total percentage correct.

Table 4.3. Correlations and Linear Mixed Model Analysis between Subjective Perception and Task Performance. The linear mixed model F values, Pearson correlation coefficients, and Spearman correlation coefficients between subjective perception of noise and performance of the task.

		Subject Questionnaire Results				
Task Performance Results	Statistical Measure	Loudness	Change in Noise	Rumble	Annoyance	Distraction
Average Time	$F_{1,134}$	ns	ns	4.35*	2.53*	ns
	Pearson (r)	.388**	.305**	.303**	ns	.204*
	Spearman (r)	.412**	.322**	.299**	.189*	.193*
% Correct	$F_{1,134}$	3.54**	ns	2.24*	3.63**	ns
	Pearson (r)	-.334**	-.235*	ns	-.203*	ns
	Spearman (r)	-.348**	-.287**	-.212*	ns	ns

*significant at p<0.05, ** significant at p<0.01, ns = not significant

The relationship between the loudness perception ratings and corresponding total percentage correct are shown in Figure 4.8. They show a small, negative correlation that means when ratings for loudness of noise increased, the total percentage of questions answered correctly generally decreased.

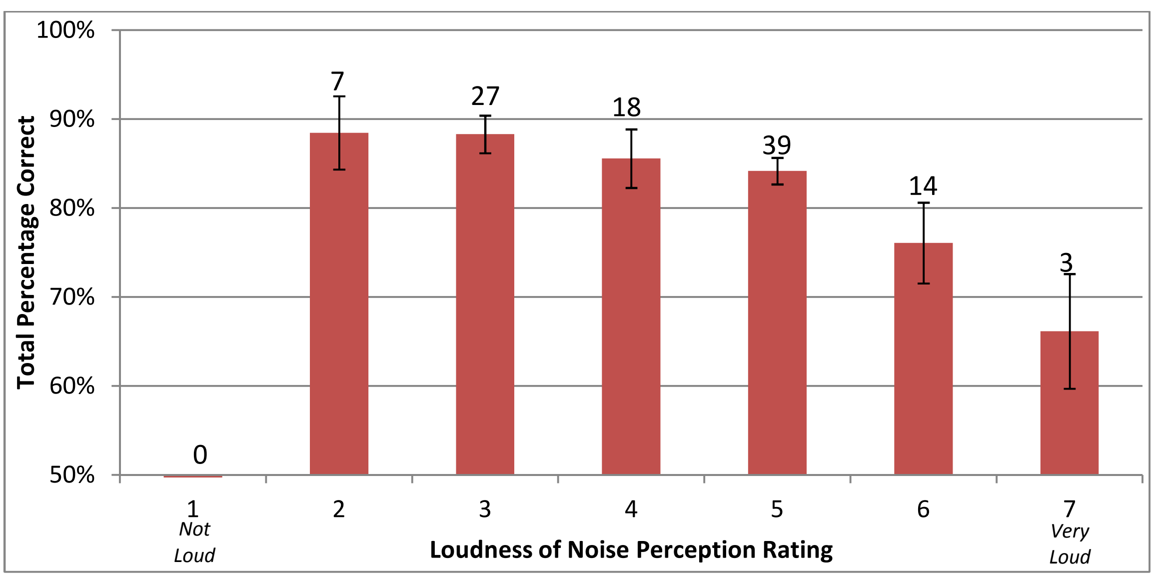


Fig. 4.8. Results of the loudness perception rating and the corresponding total percentage of correct responses given in a session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the loudness perception ratings and corresponding average response times are shown in Figure 4.9. They show a small, positive correlation that means when ratings for loudness of noise increased, the average response time generally increased.

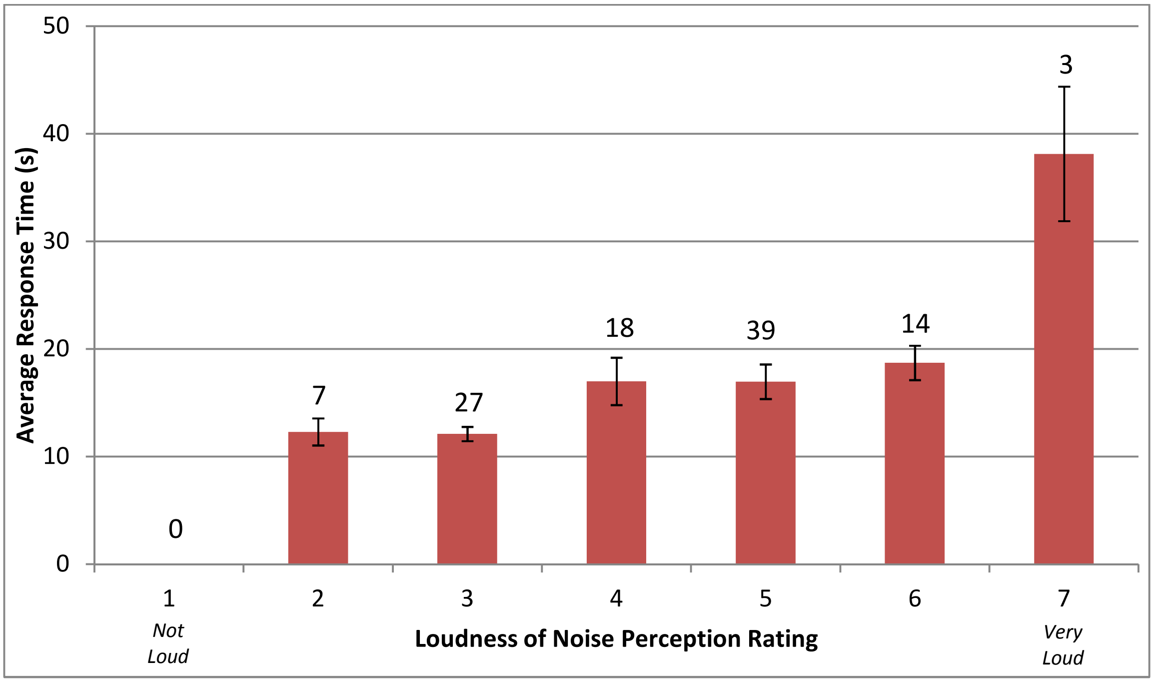


Fig. 4.9. Results of the perception ratings of changes in noise and the corresponding average time taken to solve task problems in a given session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the changes in noise perception ratings and corresponding total percentage correct are shown in Figure 4.10. They show a small, negative correlation that means when ratings for changes in the noise increased toward changing a lot, the total percentage of questions answered correctly generally decreased.

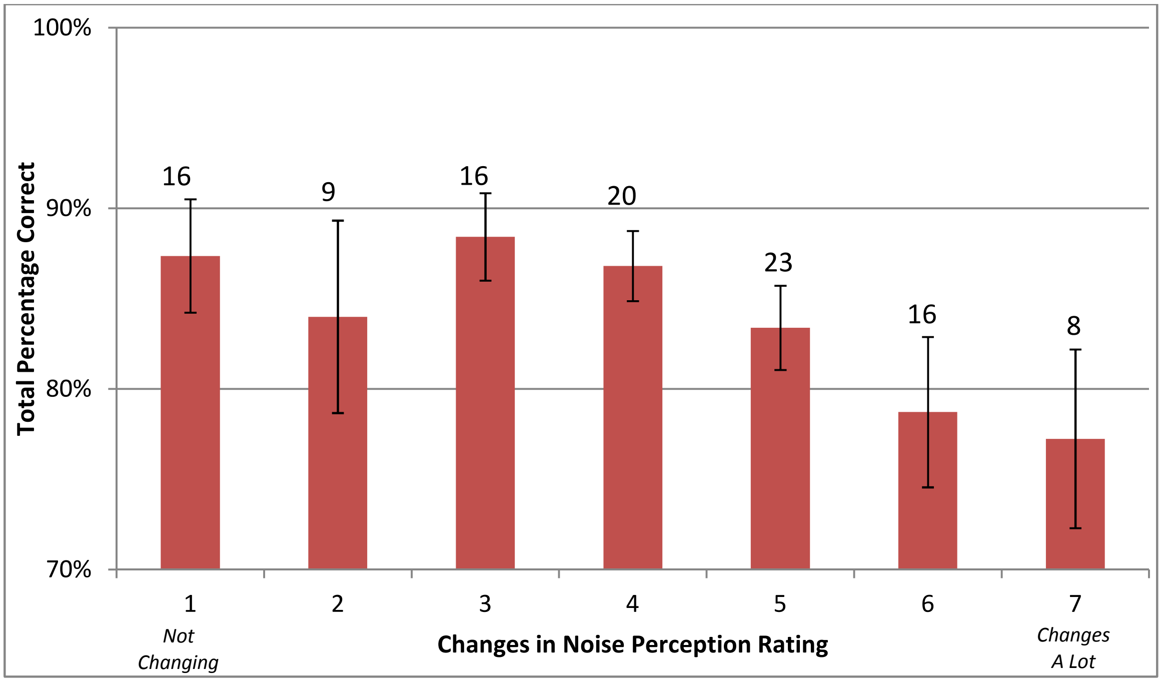


Fig. 4.10. Results of the perception ratings of changes in noise and the corresponding total percentage of correct responses given in a session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the changes in noise perception ratings and corresponding average response times are shown in Figure 4.11. They show a small, positive correlation that means when ratings for changes in the noise increased toward changing a lot, the average response time generally increased.

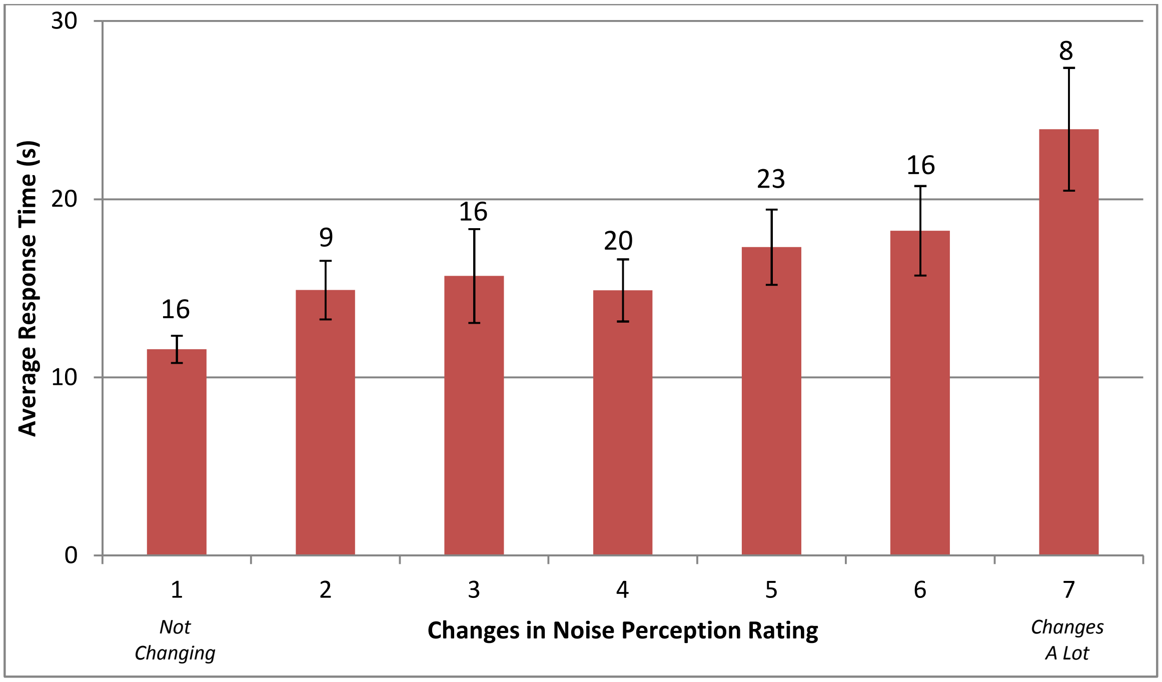
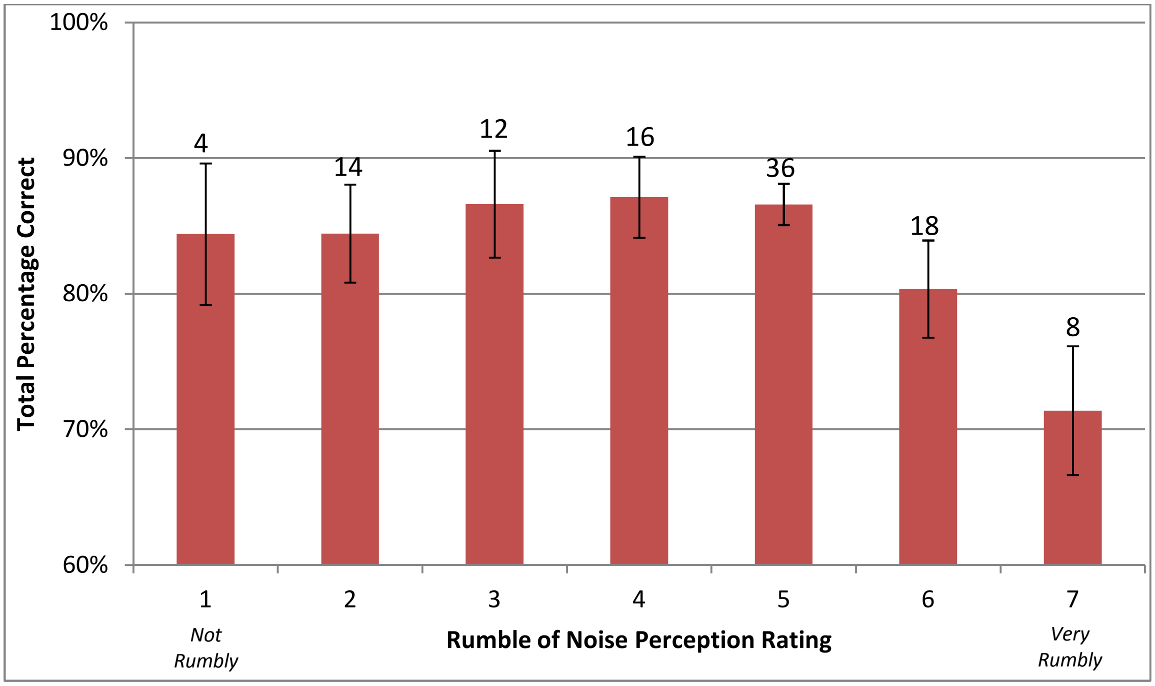


Fig. 4.11. Results of the perception ratings of changes in noise and the corresponding average time taken to solve task problems in a given session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the rumble of noise perception ratings and corresponding total percentage correct are shown in Figure 4.12. They show a small, negative correlation that means when ratings for rumble of noise increased toward very rumbly, the total percentage of questions answered correctly generally decreased.



<u>Fig. 4.12.</u> Results of the perception ratings of rumble of noise and the corresponding total percentage of correct responses given in a session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the rumble of noise perception ratings and corresponding average response times are shown in Figure 4.13. They show a small, positive correlation that means when ratings for rumble of noise increased toward very rumbly, the average response time generally increased.

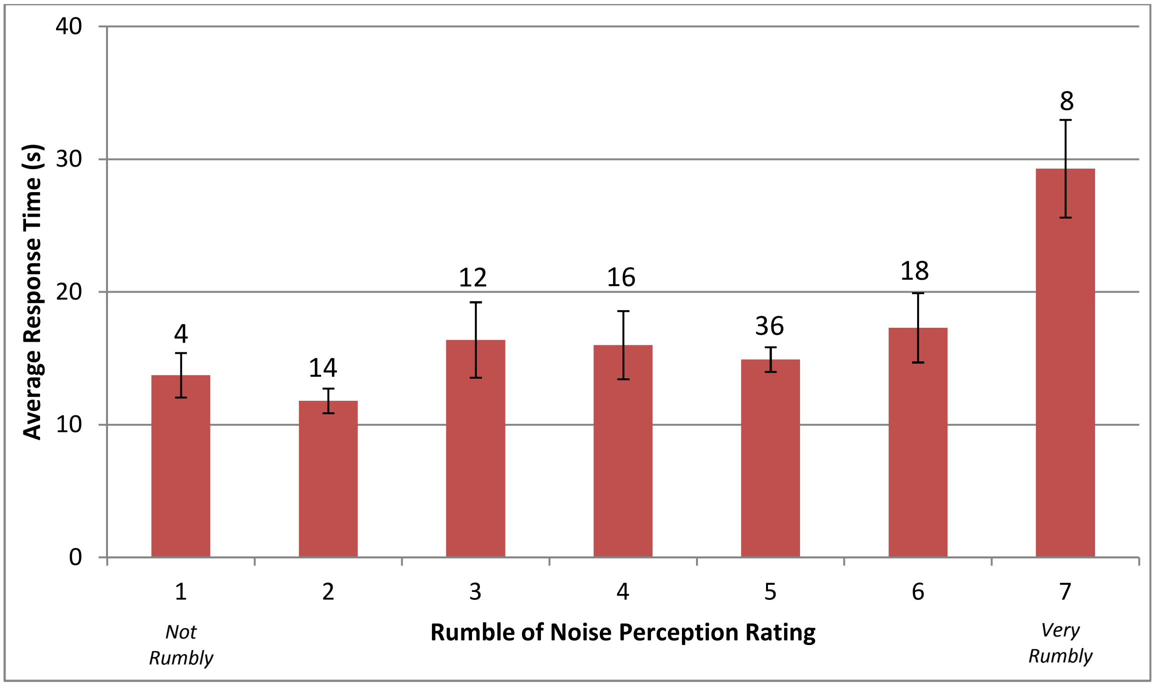


Fig. 4.13. Results of the perception ratings of rumble of noise and the corresponding average time taken to solve task problems in a given session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the annoyance to noise perception ratings and corresponding total percentage correct are shown in Figure 4.14. They show a small, negative correlation that means when ratings for annoyance to noise increased toward very annoying, the total percentage of questions answered correctly generally decreased.

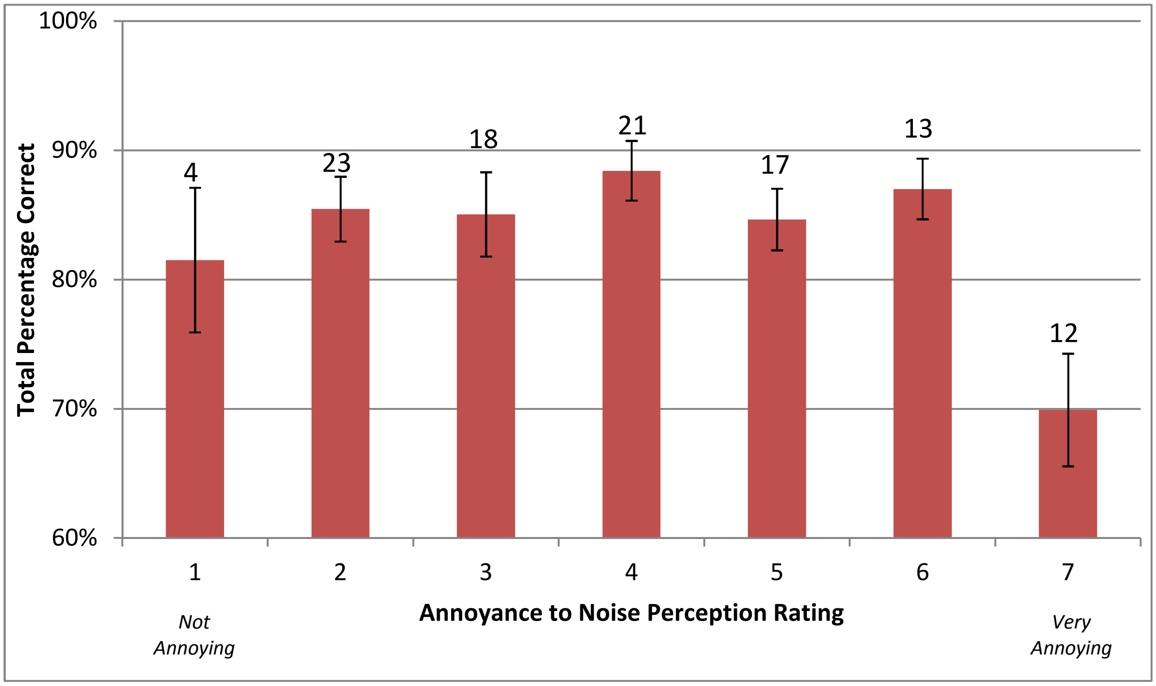


Fig. 4.14. Results of the perception ratings annoyance to noise and the corresponding total percentage of correct responses given in a session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the annoyance to noise perception ratings and corresponding average response times are shown in Figure 4.14. They show a small, positive correlation that means when ratings for annoyance to noise increased toward very annoying, the average response time generally increased.

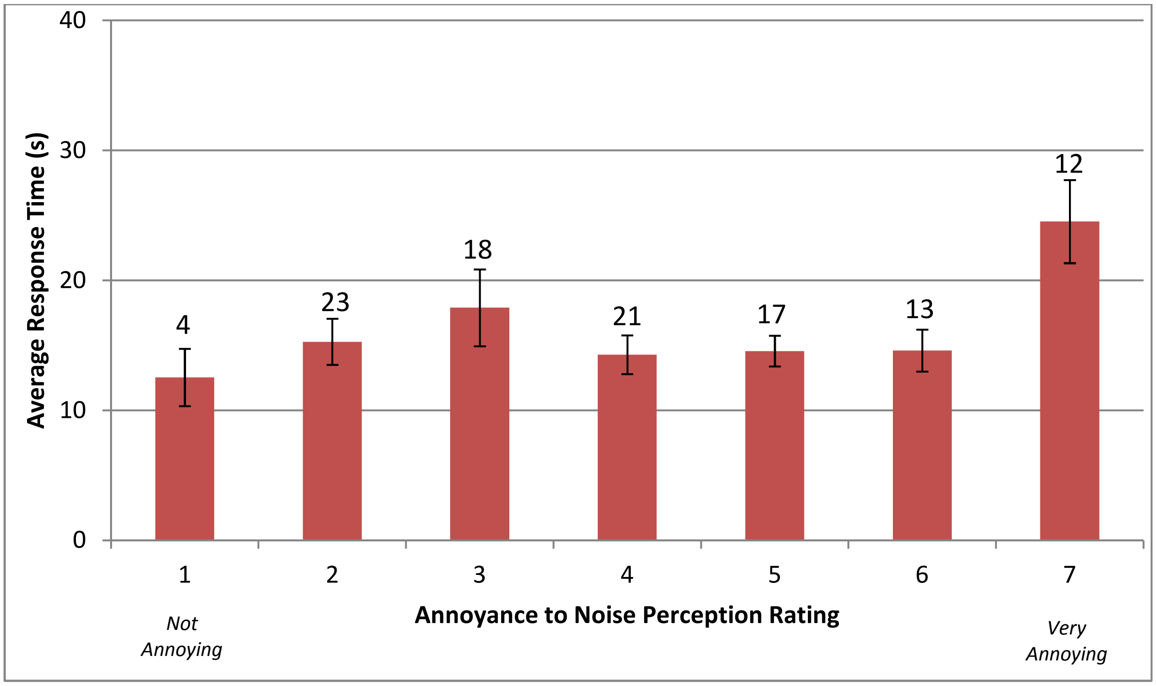


Fig. 4.15. Results of the perception ratings of annoyance to noise and the corresponding average time taken to solve task problems in a given session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the distraction of noise perception ratings and corresponding total percentage correct are shown in Figure 4.16. They show a small, negative correlation that means when ratings for distraction of noise increased toward very distracting, the total percentage of questions answered correctly generally decreased.

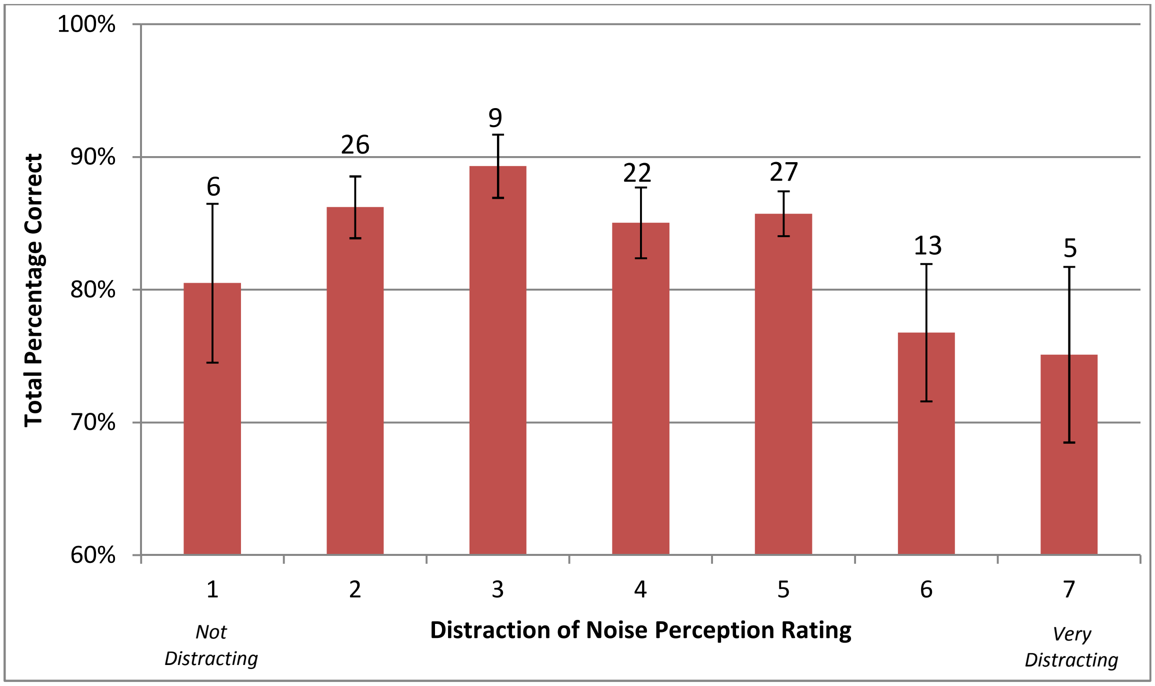


Fig. 4.16. Results of the perception ratings of distraction of noise and the corresponding total percentage of correct responses given in a session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the distraction of noise perception ratings and corresponding average response times are shown in Figure 4.17. They show a small, positive correlation that means when ratings for distraction of noise increased toward very distracting, the average response time generally increased.

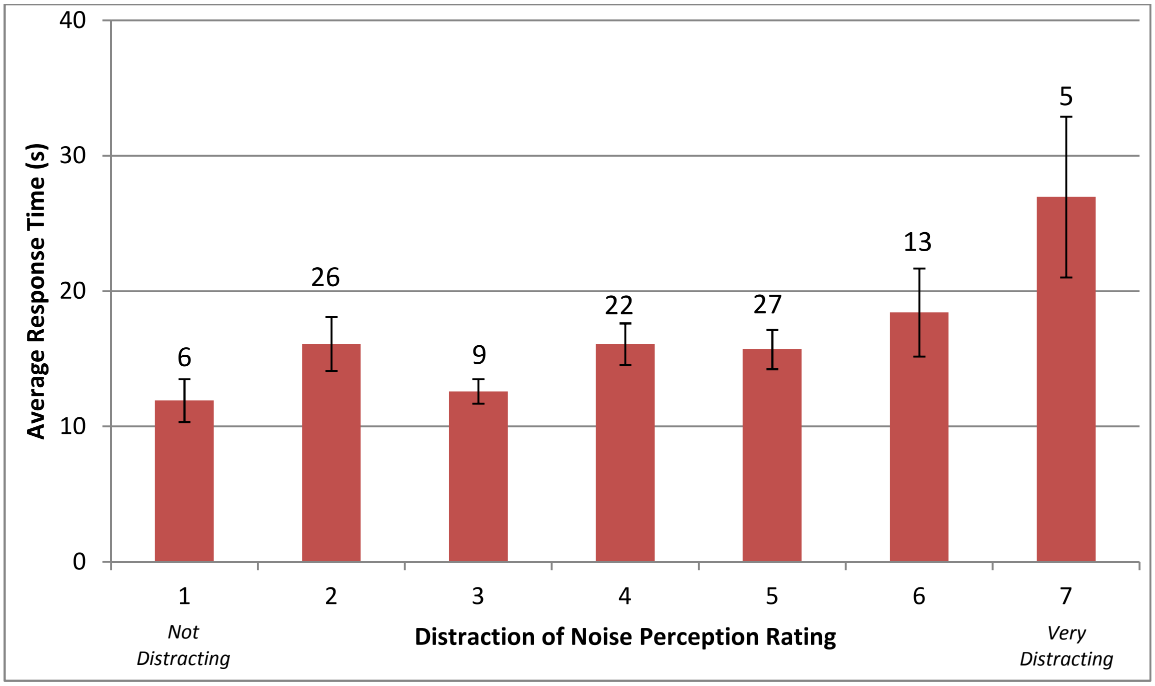


Fig. 4.17. Results of the perception ratings of distraction of noise and the corresponding average time taken to solve task problems in a given session. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

4.4. Subjective Perception Results

4.4.1. Subjective Perception Results across Noise Conditions

Subjects rated their perception of the loudness of noise, changes in time of noise, rumble of noise, annoyance to noise, and distraction of noise, as detailed in Section 3.2.3.2, for each session. These ratings were then compared to the four fluctuation intervals previously mentioned. A Kolmogorov-Smirnov test was run and showed that all results exhibited a non-normal distribution. Therefore, the Friedman ANOVA test was used to analyze these relationships, and Wilcoxon tests were used to further analyze the relationships between each noise condition.

Although it is a parametric test, repeated measures ANOVA results are also reported to support the non-parametric test results. Observed power, at $\alpha = 0.05$, is reported for each. Further results of the repeated measures ANOVA from SPSS,

including sum of squares, mean square, and degrees of freedom are reported in Section 4.4.3.

4.4.1.1. Loudness of Noise across Noise Conditions

The loudness of noise ratings were not affected by the different noise conditions, $\chi^2(3) = 0.69$, $p < 0.05$. As the fluctuation interval increased, the ratings of loudness of noise remained relatively constant. The average perception ratings for loudness of noise in each noise condition are shown in Figure 4.18. The results of the Wilcoxon test are shown in Table 4.4.

Table 4.4. Wilcoxon Results between Noise Condition and Loudness of Noise. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.005 level of significance.

Loudness ratings between sessions				
		5 Minute interval	8 Minute Interval	10 Minute Interval
2 Minute Interval	T	*54.50*	*45.00*	49.50
	sig	*ns*	*ns*	ns
	effect size	*-0.04*	*-0.07*	-0.03
5 Minute Interval	T		*77.50*	98.50
	sig		*ns*	ns
	effect size		*-0.10*	-0.03
8 Minute Interval	T			*63.00*
	sig			*ns*
	effect size			*-0.09*

** = Indicates the mean ranks between noise conditions is significant, p<0.005, ns = not significant.

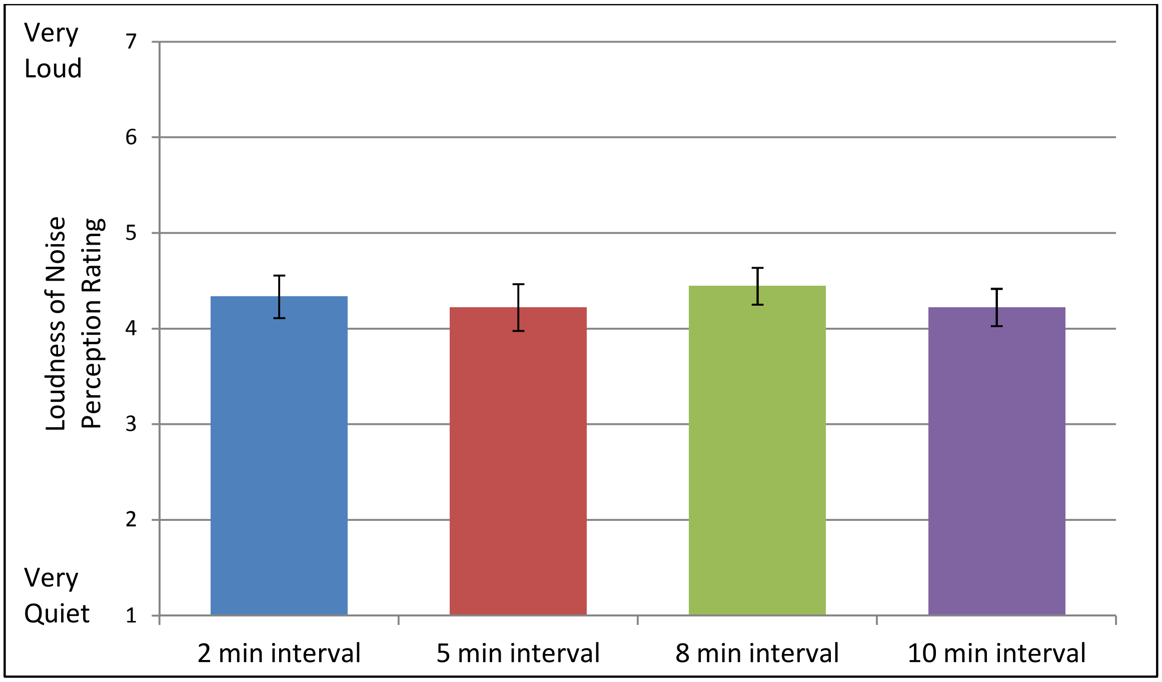


Fig. 4.18. Results of the average perception ratings of loudness of noise across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was not violated ($\chi^2(5) = 3.35$, $p<0.05$). Assuming sphericity, the results displayed a significant effect with a small to moderate effect size ($F(3,78) = 0.280$, $p<0.05$, $r = 0.10$). The observed power for this test was 0.101, or 10%, according to the results reported by SPSS with $\alpha = 0.05$. This suggests that the probability of the loudness rating vs. interval results presenting a genuine effect is poor.

4.4.1.2. Change in Noise over Time across Noise Conditions

The changes in noise over time ratings were significantly affected by the different noise conditions, $\chi^2(3) = 14.37$, $p < 0.05$. The perception ratings increase as the length of fluctuation interval increases. The average perception ratings for changes in noise over time in each noise condition are shown in Figure 4.19. The results of the Wilcoxon test are shown in Table 4.6.

Wilcoxon test results show that ratings of changes in noise over time were significantly lower at longest fluctuation interval compared to the two shortest intervals. The ratings of changes in noise over time were also significantly lower at the two minute interval compared to the eight minute interval.

Table 4.5. Wilcoxon Results between Noise Condition and Change of Noise over Time. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.005 level of significance.

Changes in noise ratings between sessions				
		5 Minute interval	8 Minute Interval	10 Minute Interval
2 Minute Interval	T	*40.00*	*47.50*	17.50
	sig	*ns*	**	**
	effect size	*-0.24*	*-0.30*	-0.41
5 Minute Interval	T		*16.50*	38.00
	sig		*ns*	**
	effect size		*-0.37*	-0.26
8 Minute Interval	T			*0.00*
	sig			*ns*
	effect size			*0.00*

** = Indicates the mean ranks between noise conditions is significant, p<0.005, ns = not significant.

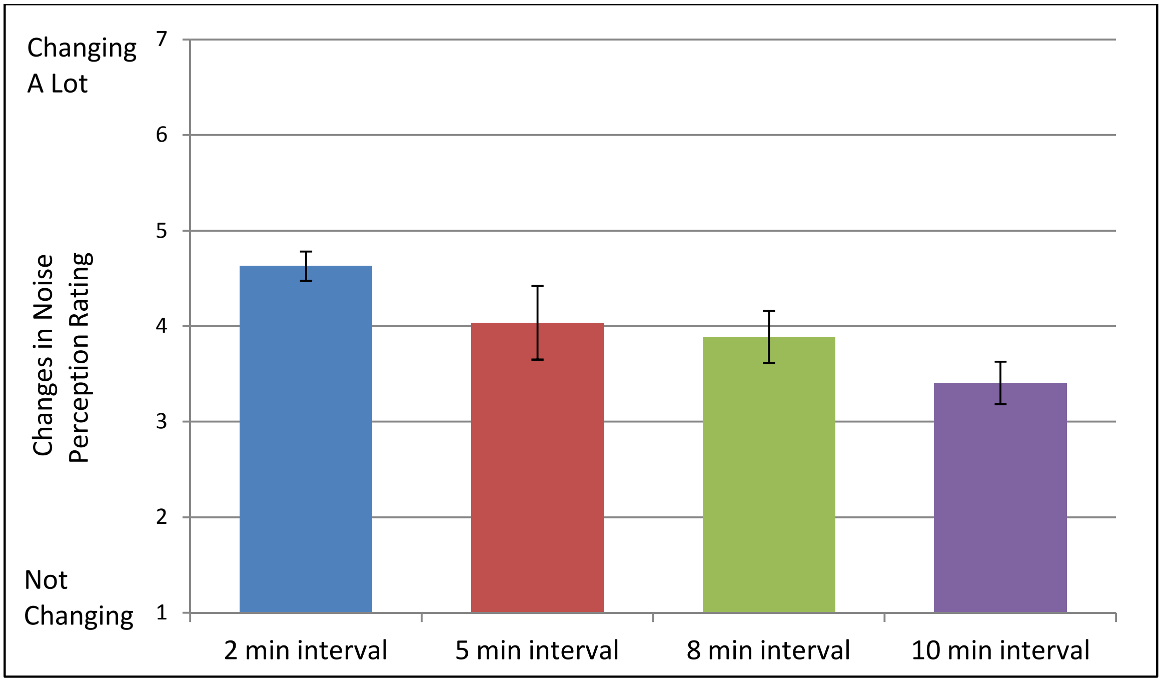


Fig. 4.19. Results of the average perception ratings of changes in noise over time across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was not violated ($\chi2(5) = 5.484$, $p<0.05$). Assuming sphericity, the results displayed a significant effect with a moderate to large effect size ($F(3,78) = 7.05$, $p<0.05$, $r = 0.45$). The observed power for this test was 0.976, or 97.6%, according to the results reported by SPSS with $\alpha = 0.05$. This means the probability of the changes rating vs. interval results presenting a genuine effect is very good.

Bonferroni post hoc tests found significant relationships as shown in Table 4.6. These significant relationships matched those found in the Wilcoxon test, except that no significant relationship between the two and eight minute fluctuation intervals was found in the repeated measures ANOVA results. Since the data has a non-normal distribution, the non-parametric test results are more likely to represent the accurate effects.

Therefore, the significant relationship difference the two and eight minute intervals probably still exists.

<u>Table 4.6.</u> Bonferroni Post Hoc Tests for Changes in Noise over Time Ratings across Noise Conditions.

Change in noise ratings between sessions			
	5 Minute interval	8 Minute Interval	10 Minute Interval
2 Minute Interval			**
5 Minute interval			**
8 Minute Interval			

**Indicates the mean difference between noise conditions is significant, $p<0.05$

4.4.1.3. Rumble of Noise across Noise Conditions

The rumble of noise ratings were not significantly affected by the different noise conditions, $\chi^2(3) = 2.88$, $p < 0.05$. As the noise level increases, the rumble perception ratings stayed relatively constant between four and five. The average perception ratings for rumble of noise over time in each noise condition are shown in Figure 4.20. The results of the Wilcoxon test are shown in Table 4.7.

<u>Table 4.7.</u> Wilcoxon Results between Noise Condition and Rumble of Noise. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.005 level of significance.

Rumble ratings between sessions				
		5 Minute interval	8 Minute Interval	10 Minute Interval
2 Minute Interval	T	*35.00*	*47.50*	67.00
	sig	*ns*	*ns*	ns
	effect size	*-0.15*	*-0.10*	-0.01
5 Minute Interval	T		*52.00*	56.50
	sig		*ns*	ns
	effect size		*-0.16*	-0.13
8 Minute Interval	T			*0.00*
	sig			*ns*
	effect size			*0.00*

** = Indicates the mean ranks between noise conditions is significant, $p<0.005$, ns = not significant.

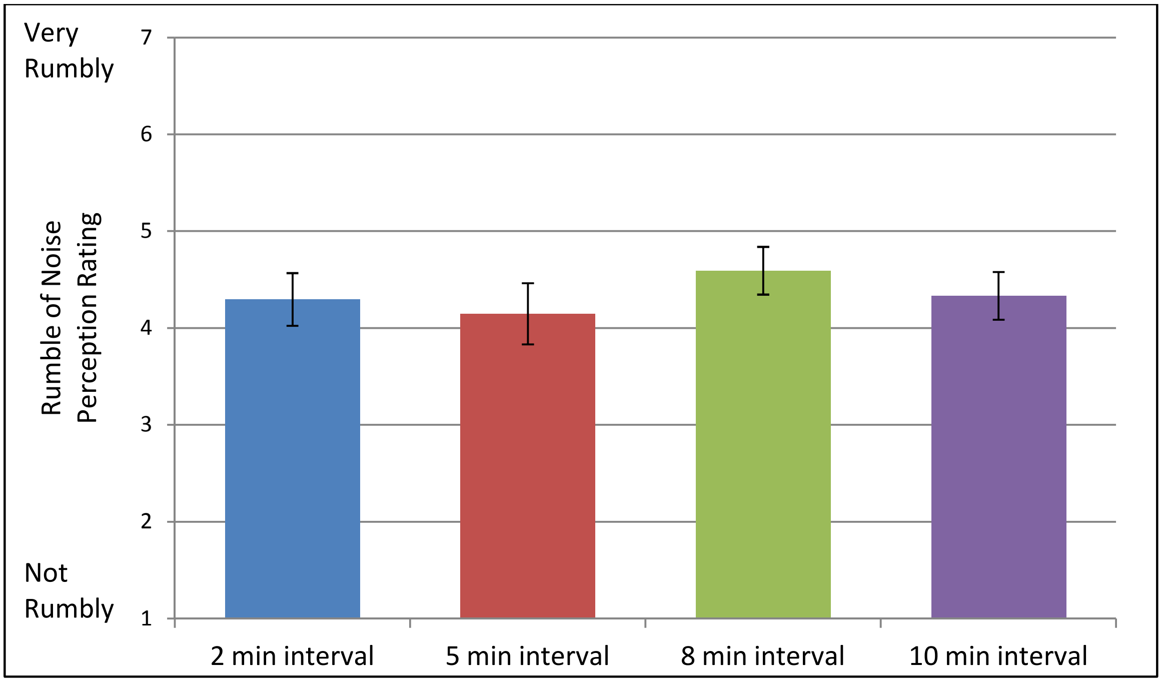


Fig. 4.20. Results of the average perception ratings of rumble of noise across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was not violated so sphericity was assumed ($F(3,78) = 0.95$, $p<0.05$, $r = 0.18$). The observed power for this test was 0.185, or 18.5%, according to the results reported by SPSS with $\alpha = 0.05$. This suggests the probability of the rumble perception vs. interval results presenting a genuine effect is poor. Bonferroni post hoc tests found no significant relationships.

4.4.1.4. Annoyance to Noise across Noise Conditions

The annoyance to noise ratings were not significantly affected by the different noise conditions, $\chi^2(3) = 0.28$, $p < 0.05$. As the fluctuation interval length increases, the perception ratings remain relatively constant. The average perception ratings for annoyance to noise over time in each noise condition are shown in Figure 4.21. The results of the Wilcoxon test are shown in Table 4.8.

Table 4.8. Wilcoxon Results between Noise Condition and Annoyance to Noise. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.005 level of significance.

Annoyance ratings between sessions				
		5 Minute interval	8 Minute Interval	10 Minute Interval
2 Minute Interval	T	*62.50*	*69.00*	102.50
	sig	*ns*	*ns*	ns
	effect size	*-0.04*	*-0.05*	-0.01
5 Minute Interval	T		*65.00*	88.00
	sig		*ns*	ns
	effect size		*-0.02*	-0.04
8 Minute Interval	T			*0.00*
	sig			*ns*
	effect size			*0.00*

** = Indicates the mean ranks between noise conditions is significant, p<0.005, ns = not significant.

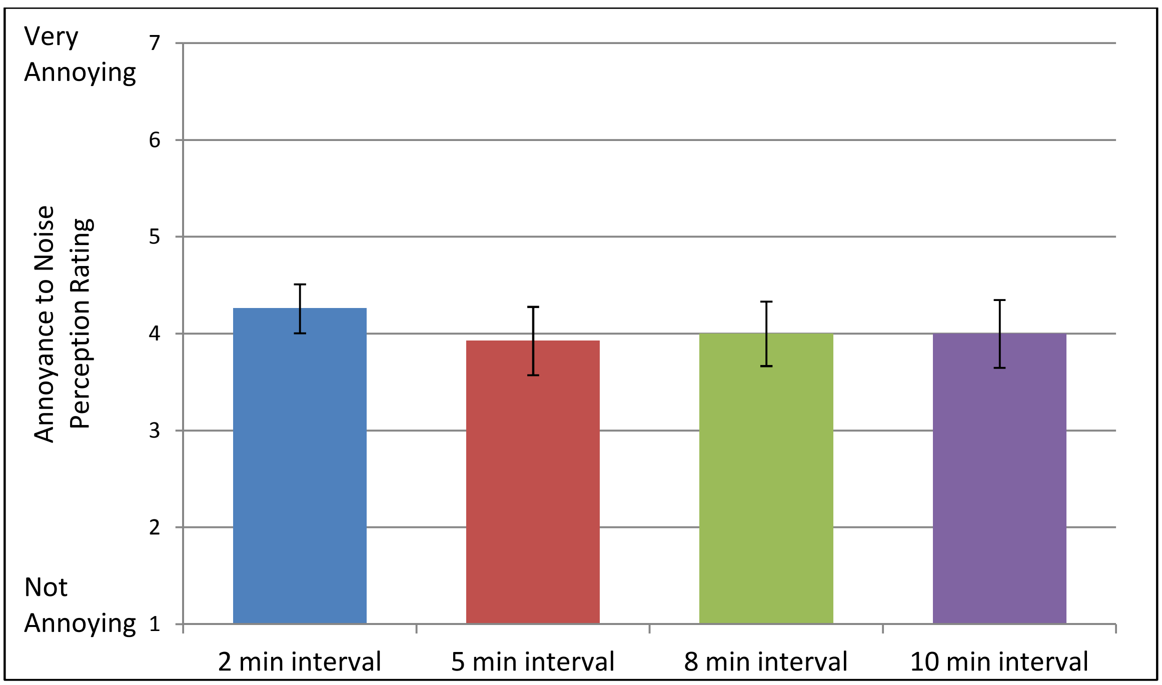


Fig. 4.21. Relationship between Noise Conditions and Annoyance to Noise Ratings. Results of the average perception ratings of annoyance to noise across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was not violated so sphericity was assumed

(F(3,78) = 0.070, p<0.05, r = 0.05). The effect size is considered small. The observed power for this test was 0.062, or 6.2%, according to the results reported by SPSS with $\alpha = 0.05$. This suggests the probability of the annoyance perception vs. interval results presenting a genuine effect is poor. Bonferroni post hoc tests found no significant relationships.

4.4.1.5. Distraction to Noise across Noise Conditions

The distraction to noise ratings were not significantly affected by the different noise conditions ($\chi^2(3) = 1.95$, $p < 0.05$). As the length of fluctuation interval increases, the perception rating stays relatively constant. The average perception ratings for distraction to noise over time in each noise condition are shown in Figure 4.22. The results of the Wilcoxon test are shown in Table 4.9.

Table 4.9. Wilcoxon Results between Noise Condition and Distraction to Noise. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.005 level of significance.

Distraction ratings between sessions				
		5 Minute interval	8 Minute Interval	10 Minute Interval
2 Minute Interval	T	*46.00*	*96.00*	108.50
	sig	*ns*	*ns*	ns
	effect size	*-0.16*	*-0.05*	-0.08
5 Minute Interval	T		*51.00*	107.00
	sig		*ns*	ns
	effect size		*-0.07*	-0.04
8 Minute Interval	T			*0.00*
	sig			*ns*
	effect size			*0.00*

** = Indicates the mean ranks between noise conditions is significant, p<0.005, ns = not significant.

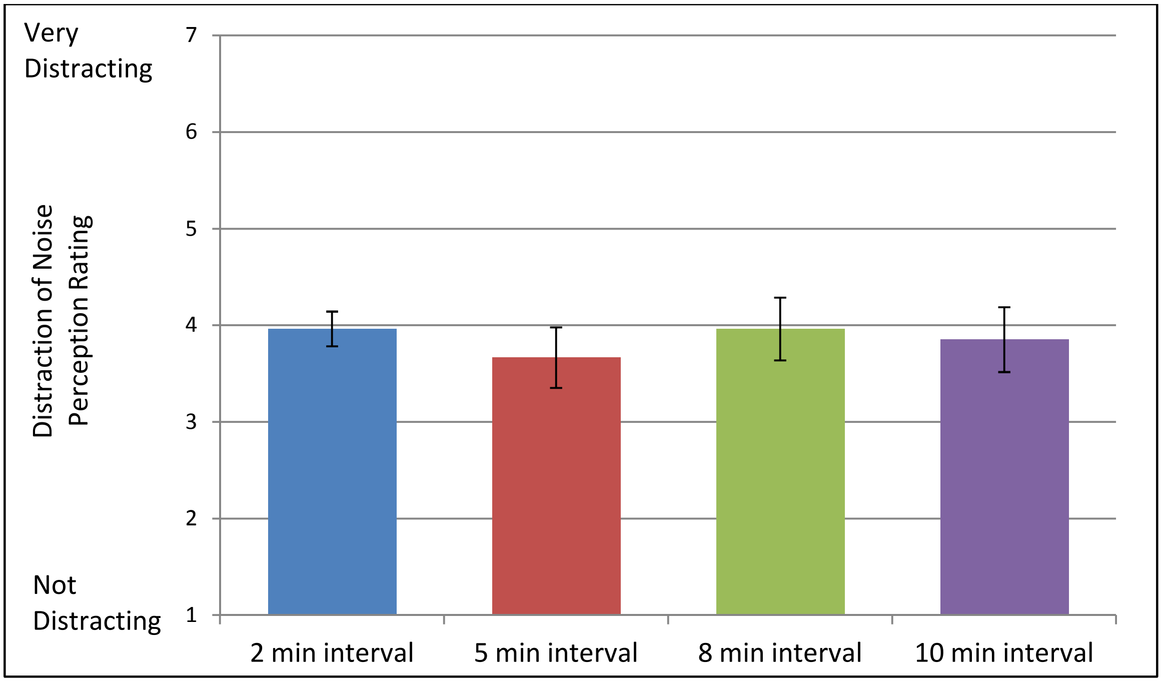


Fig. 4.22. Relationship between Noise Conditions and Distraction to Noise Ratings. Results of the average perception ratings of distraction to noise across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was not violated so sphericity was assumed ($F(3,78) = 0.558$, $p<0.001$, $r = 0.14$). The effect size is considered small. The observed power for this test was 0.134, or 13.4%, according to the results reported by SPSS with $\alpha = 0.05$. This suggests the probability of the distraction perception vs. interval results presenting a genuine effect is poor. Bonferroni post hoc tests found no significant relationships.

4.4.1.6. Discussion of Subjective Perception Results across Noise Conditions

No significant differences were found for all results of subjective perception ratings except changes in noise across noise conditions. This suggests subjects do not perceive these noise conditions as very different (except for the changes) by the standards

of the subjective questionnaire. However, as shown in Section 4.2.3, subjective perceptions of the noise seem tied to their performance. Subjects that rated the sounds higher (or harsher) tended to perform worse on the tests.

Subjects provided comments on the questionnaires at the end of each test session regarding their reactions to the bursts of noise. Since this study was done in conjunction with a study involving bursts of noise (Ainley, 2012), most subjects compared the bursts of noises from that test to the fluctuations of this test. Subjects that commented fell into two camps of about equal distribution: those that preferred bursts and those that preferred fluctuations. The subjects that preferred the fluctuations preferred them because it was more consistent and they did not get anxious anticipating bursts.

Multiple people commented that the periods of change between the two noise levels was the most distracting part of the sessions. Once in steady state, they adapted pretty quickly. A few noted that they still found it harder to concentrate in louder background noise. Although not evident in the subjective perception survey results, multiple subjects also singled out the two minute fluctuation interval as more distracting than the rest of the sessions.

4.4.2. Relationships of Subjective Perception Results across Noise Conditions with Gender, Age, and Noise Sensitivity as Covariates

There were very few significant relationships found between subjective perception results and noise conditions as shown in Section 4.4.1. However, additional variables were not factored into these results. Gender, age, and noise sensitivity were other independent variables collected during this study. It is necessary to look at the

effects of these variables on the relationship between subjective perception ratings to the four noise conditions to get a better understanding of everything affecting the results.

These additional independent variables are difficult to include in a non-parametric test like the ones used in this study. However, repeated measures ANOVA with covariates can still be analyzed to study these relationships with multiple independent variables. Because of the non-normal distributions of the subjective perception ratings, these results are presented with caution. Additionally, the observed power, as reported by SPSS with $\alpha = 0.05$, is reported for each repeated measures ANOVA test.

The SPSS outputs for each subjective perception rating with each covariate combination are shown in Tables 4.10 to 4.14. Almost all relationships remain insignificant, $p < 0.05$, with the addition of covariates. The relationship between noise condition and perceived changes in noise, which is significant without covariates, becomes insignificant with the addition of any or all analyzed covariates. However, there is one interesting outlier to note: the addition of age and noise sensitivity makes the relationship between perceived loudness of the noise and the noise condition significant (Table 4.10).

Observed power depends on the number of independent variables and the sample size. In all cases, except for the significant relationship between perceived changes and the noise condition, the observed power is below the recommended 0.8. A larger sample size is desired to increase the observed power – increasing the probability that these tests show a genuine effect (Field and Hole 2003). Therefore, these results are presented with caution.

Table 4.10. Analysis of variance for loudness to noise ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Loudness to Noise Ratings across Noise Conditions	0.694	3.000	0.231	0.280	0.840	0.011	0.839	0.101
Gender		1.252	3.000	0.417	0.494	0.688	0.019	1.481	0.146
Age		4.489	3.000	1.496	1.860	0.144	0.069	5.580	0.464
Noise Sensitivity		3.381	3.000	1.127	1.376	0.257	0.052	4.127	0.352
Gender and Noise Sensitivity		1.301	3.000	0.434	0.520	0.670	0.021	1.560	0.151
Gender and Age		2.758	3.000	0.919	1.111	0.350	0.044	3.334	0.288
Age and Noise Sensitivity		7.346	3.000	2.449	3.119	**0.031**	0.115	9.356	0.704
Gender, Age, and Noise Sensitivity		4.598	3.000	1.533	1.897	0.138	0.076	5.691	0.471

[a]Computed using alpha = .05

Table 4.11. Analysis of variance for change in noise over time ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Change in Noise over Time Ratings across Noise Conditions	20.843	3.000	6.948	7.046	0.000	0.213	21.139	0.976
Gender		0.663	3.000	0.221	0.223	0.880	0.009	0.668	0.090
Age		1.951	3.000	0.650	0.637	0.593	0.025	1.911	0.178
Noise Sensitivity		3.297	3.000	1.099	1.092	0.358	0.042	3.277	0.284
Gender and Noise Sensitivity		1.437	2.596	0.553	0.471	0.676	0.019	1.223	0.134
Gender and Age		0.711	3.000	0.237	0.230	0.875	0.010	0.691	0.091
Age and Noise Sensitivity		1.778	3.000	0.593	0.568	0.638	0.023	1.705	0.162
Gender, Age, and Noise Sensitivity		1.318	3.000	0.439	0.416	0.742	0.018	1.249	0.129

[a]Computed using alpha = .05

Table 4.12. Analysis of variance for rumble of noise ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Rumble of Noise Ratings across Noise Conditions	3.704	3.000	1.235	0.951	0.420	0.035	2.852	0.251
Gender		0.199	3.000	0.066	0.049	0.985	0.002	0.148	0.058
Age		3.542	3.000	1.181	0.899	0.446	0.035	2.697	0.238
Noise Sensitivity		8.832	3.000	2.944	2.331	0.081	0.085	6.994	0.565
Gender and Noise Sensitivity		1.813	3.000	0.604	0.464	0.708	0.019	1.391	0.139
Gender and Age		2.146	3.000	0.715	0.525	0.667	0.021	1.574	0.152
Age and Noise Sensitivity		7.129	3.000	2.376	1.873	0.142	0.072	5.618	0.466
Gender, Age, and Noise Sensitivity		3.838	3.000	1.279	0.972	0.411	0.041	2.915	0.254

[a]Computed using alpha = .05

Table 4.13. Analysis of variance for annoyance to noise ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Annoyance to Noise Ratings across Noise Conditions	0.324	3.000	0.108	0.070	0.976	0.003	0.209	0.062
Gender		1.404	3.000	0.468	0.294	0.830	0.012	0.881	0.104
Age		2.159	3.000	0.720	0.454	0.715	0.018	1.363	0.137
Noise Sensitivity		4.194	3.000	1.398	0.911	0.440	0.035	2.732	0.241
Gender and Noise Sensitivity		2.266	3.000	0.755	0.478	0.698	0.020	1.435	0.142
Gender and Age		2.848	3.000	0.949	0.584	0.627	0.024	1.753	0.165
Age and Noise Sensitivity		3.853	3.000	1.284	0.799	0.499	0.034	2.397	0.214
Gender, Age, and Noise Sensitivity		3.853	3.000	1.284	0.799	0.499	0.034	2.397	0.214

[a]Computed using alpha = .05

Table 4.14. Analysis of variance for distraction to noise ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Distraction to Noise Ratings across Noise Conditions	2.843	3.000	0.948	0.558	0.644	0.021	1.675	0.160
Gender		0.883	3.000	0.294	0.168	0.918	0.007	0.505	0.080
Age		0.481	3.000	0.160	0.091	0.965	0.004	0.274	0.066
Noise Sensitivity		2.025	3.000	0.675	0.385	0.764	0.015	1.156	0.123
Gender and Noise Sensitivity		0.255	3.000	0.085	0.047	0.986	0.002	0.141	0.058
Gender and Age		1.201	3.000	0.400	0.221	0.881	0.009	0.663	0.090
Age and Noise Sensitivity		0.515	3.000	0.172	0.095	0.963	0.004	0.284	0.066
Gender, Age, and Noise Sensitivity		0.582	3.000	0.194	0.104	0.958	0.004	0.311	0.068

[a]Computed using alpha = .05

Chapter 5: Rattle Methodology

The purpose of this study was to evaluate the performance and perception of humans under noise bursts of varying amplitude with and without a rattle element. Subjects completed an arithmetic test under different acoustic conditions for eight different test sessions and filled out subjective questionnaires over the test environment at the end of each session. Each session lasted a total of thirty minutes and was comprised of three parts: (1) a five minute practice period, (2) a twenty minute test period, and (3) five minutes for a subjective questionnaire.

Test subjects experienced eight different noise conditions separated into two groups: (1) four sessions involving noise bursts and (2) four sessions involving the same noise bursts accompanied by an additional rattle noise. For all tests, a consistent background noise with a room criteria rating of RC-29(H) was introduced to the room. The bursts of noise were synthesized broadband noise signals presented at four peak A-weighted sound pressure levels (L_{Apk}) ranging from 55 dBA to 70 dBA. The rattle noise was a pre-recorded rattle of a wall hung mirror provided by NASA Langley Research Center and was presented four dBA higher than the accompanying burst noise, with L_{Apk} levels ranging from 59 dBA to 74 dBA. The level of the noise burst, or noise burst and rattle, remained constant within a single session but varied across all of the sessions.

5.1. Facilities

5.1.1. New Nebraska Test Chamber

All testing was carried out in the new Nebraska Test Chambers in room 131 of the Peter Kiewit Institute (PKI) on the University of Nebraska campus – a different space than the one where the previously discussed fluctuation testing was conducted. The test

chambers are acoustically isolated from the nearby spaces with a field sound transmission class (FSTC) rating of 30 between the test room and monitor room. The test room measures approximately 9' x 13' with the long slanted wall approximately six degrees off angle and the short slanted wall approximately 8 degrees off angle. The test room has gypsum board walls, carpet, and acoustical ceiling tiles (ACT). The room is further acoustically treated with one-inch thick 4' x 8' Tectum panels hung on the left and rear walls and four bass traps – one in the front right corner, one in the rear right corner, and two in the rear left corner. The average mid-frequency reverberation time was measured as 0.22 seconds. The layout of the New Nebraska Test Chambers is shown in Figure 5.1.

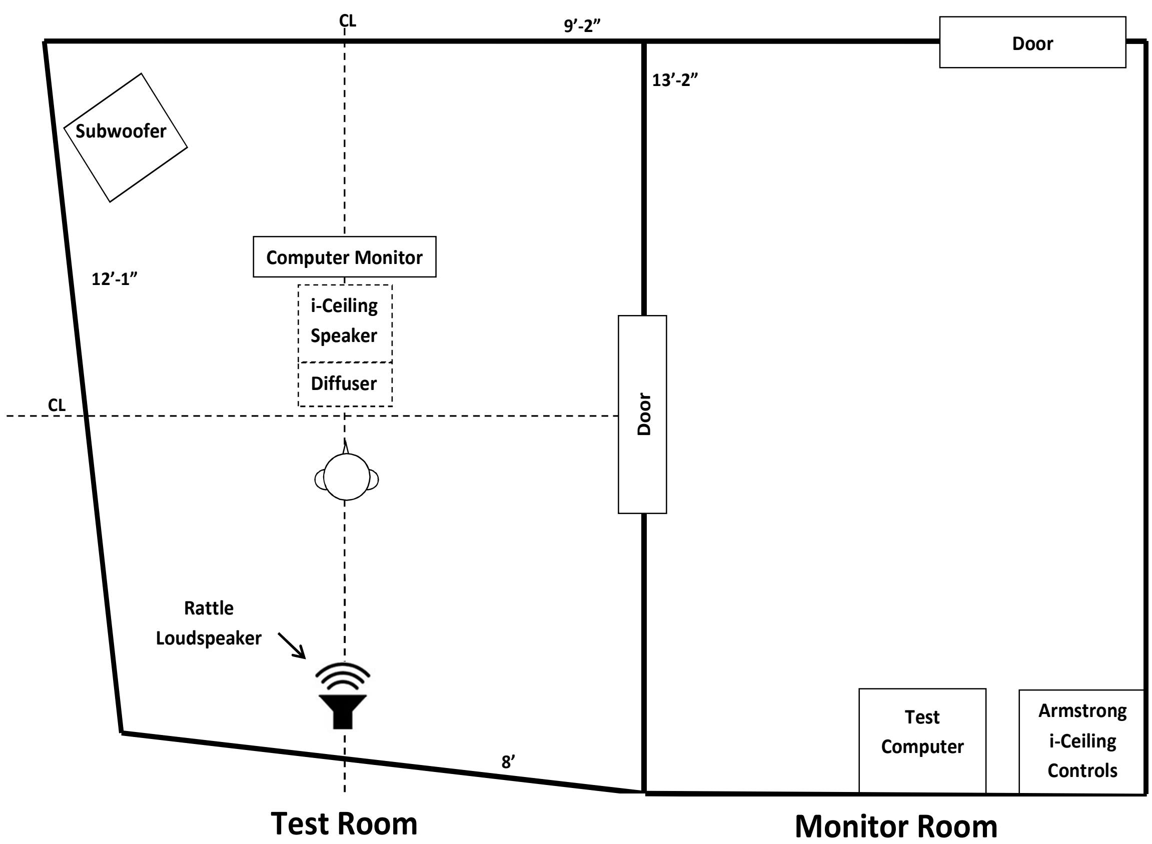


Fig. 5.1. The layout of the new Nebraska Test Chambers showing locations of the subject, test equipment, and loudspeakers used in this study (not to scale). Room height is 8'-5".

The test room contains a chair with a built-in desk, a wireless keyboard to input answers, and a computer monitor to display test questions. The chair was oriented in the room so that the subject’s head was 4’ 3” from the wall shared with the monitor room, 5’ 2” from the back wall and 3’ 6” off the ground. It should be noted that this is the location of the sound level meter used for the measurements mentioned in Section 5.2.2. This is considered the approximate subject head location due to variations in subject height and head movements during testing. Subject’s head position was not monitored during the testing. The subject sat approximately 4’ away from the 23.5” computer monitor. For legibility, all fonts displayed on the monitor were sized to be at least 36 point.

The loudspeakers used to implement the noise conditions are an Armstrong i-ceiling loudspeaker, a JBL Northridge ESeries subwoofer, and a JBL LSR6300 Series studio monitor. The i-ceiling loudspeaker resembles an ordinary ACT and is situated next to a diffuser in the ceiling. The subwoofer, covered in fabric and situated in the corner of the room, is necessary to provide the low-frequency content of the impulse signal. The JBL studio monitor, covered in fabric and located 3’ 7” directly behind and 9” above the subjects head, is used to introduce the rattle noise. Two loudspeakers, utilized in an unrelated test, are also in the room. They are covered in fabric, and the subjects were told to ignore them for the test. Two photographs of the interior of the test room are shown in Figure 5.2.

<u>Fig. 5.2.</u> Pictures of the interior of the test room.

The monitor room is located adjacent to the test room. It contains the test computer and the power amplifier for the loudspeakers. It is also the room that the test monitor works from and that the subjective questionnaires are filled out in.

The test room was controlled for temperature as best as possible. An average of 74.7 ^{o}F was measured across all test sessions.

5.1.2. Sound and Computer Systems

A diagram of the Nebraska Test Chamber testing and sound system is shown in Figure 5.3. The test computer runs the arithmetic test and generates the sound signals. With the computer and loudspeaker controls in the monitor room, the loudspeakers are the only sources of noise in the test room. The JBL LSR6300 Series studio monitor is a powered loudspeaker; the rattle signal does not need to go through the Armstrong i-Ceiling DSP and power amplifier controls.

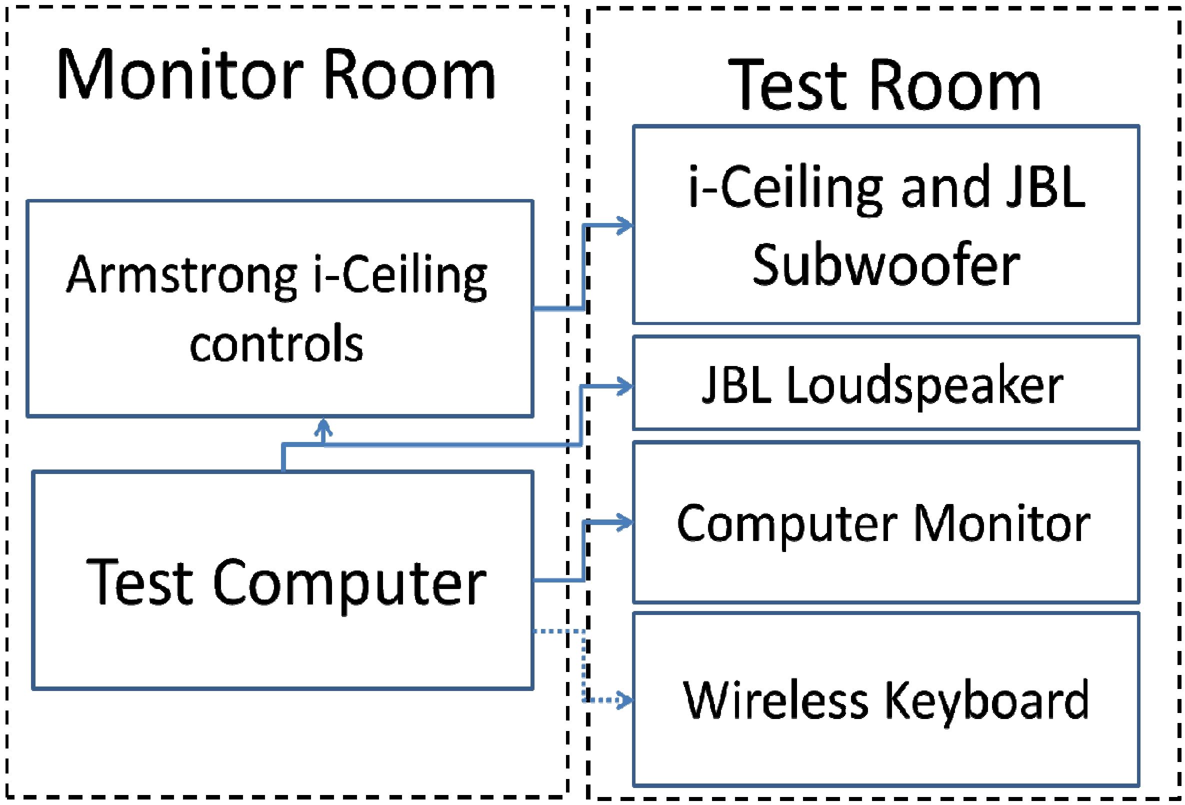


Fig. 5.3. A diagram of the new Nebraska Test Chamber system showing both the testing system and the sound system.

5.2. Experimental Methods

This section reviews the experimental methodology and is separated into four subsections: (1) creation of the signals used, (2) recording and measurement procedures used to analyze each signal, (3) procedure involved with the creation and running of the test sessions, and (4) statistical analysis.

5.2.1. Signal Creation

Nine sound files of three types were utilized in this study. One type, the broadband background noise, was synthesized to an approximate room criteria rating of RC-30(N). This level was selected to be representative of a quiet workspace. A ten second file was looped and calibrated using the equalizer in CoolEdit until a RC-29(H),

37 dBA, was measured in the test chamber while being played back over the JBL subwoofer and the Armstrong i-ceiling loudspeaker.

Four levels of a broadband noise burst and four levels of a broadband noise burst including a rattle element were also required for the study. These two types of signals will be discussed in the following subsections.

5.2.1.1. Impulse Sound Signals

Four broadband impulse sound signals of varying intensity were created for this test of 250 ms in length, a typical length of a sonic boom (Shepherd and Sullivan 1991). White noise bursts were created in CoolEdit at four different intensities with equal dBA across all octave bands.

Each signal was looped and calibrated using the equalizer function in CoolEdit until two criteria were met: (1) the A-weighted sound pressure level at each octave band was about the same (within 2 dB of each other) and (2) the total sound pressure levels equaled 55, 60, 65, or 70 dBA. These levels were selected to test a narrower and finer range than in a previous study that used impulses in 10 dBA increments from 50 to 70 dBA and found an annoyance cutoff around 70 dBA (Ainley 2012). The synthesized background noise of RC-29(H) was also played during calibration, since it would be present during testing.

The impulse levels of 55, 60, 65, and 70 dBA were initially established while the signals were played back as a continuous loop and not as a single 250 ms impulse, as presented to the subjects. Final frequency analysis results will be presented in the next chapter.

5.2.1.2. Impulse Plus Rattle Sound Signals

Four more signals were created utilizing the impulse signals, but a rattle element was added to these. The rattle signal utilized was a pre-recorded rattle noise obtained from NASA Langley Research Center titled "mirror_hung_on_metal_wire.wav" (Loubeau et al. 2013, Rathsam et al. 2013). This file was played back over the JBL studio monitor and adjusted using the CoolEdit amplify function until the A-weighted peak sound pressure level (L_{Apk}) consistently returned the desired level. It was desired to have the rattle be 4 dB louder than the impulse, so 59, 64, 69, and 74 dBA rattle files were created.

The individual impulse and rattle mono sound files were then combined in Audacity. For playback purposes, it was desirable for the impulse to come from the JBL subwoofer and Armstrong i-ceiling loudspeaker while the rattle comes from the JBL studio monitor, both being triggered at the same time by the testing program. A stereo cable connected the test computer (a mono cable was used for impulse alone tests) with one channel running to the Armstrong controls and the other to the powered JBL studio monitor.

The impulse signal was hard-panned to the left channel of a stereo file. The onset of the rattle waveform was offset by 10 ms and the entire rattle signal was hard-panned to the right channel of the same stereo file. This yielded four impulse plus rattle sound files titled "impulse 55(L) + rattle 59(R)", "impulse 60(L) + rattle 64(R)", "impulse 65(L) + rattle 69(R)", and "impulse 70(L) + rattle 74(R)". An Audacity screenshot of the "impulse 70(L) + rattle 74(R)" stereo waveform is shown in Figure 5.6.

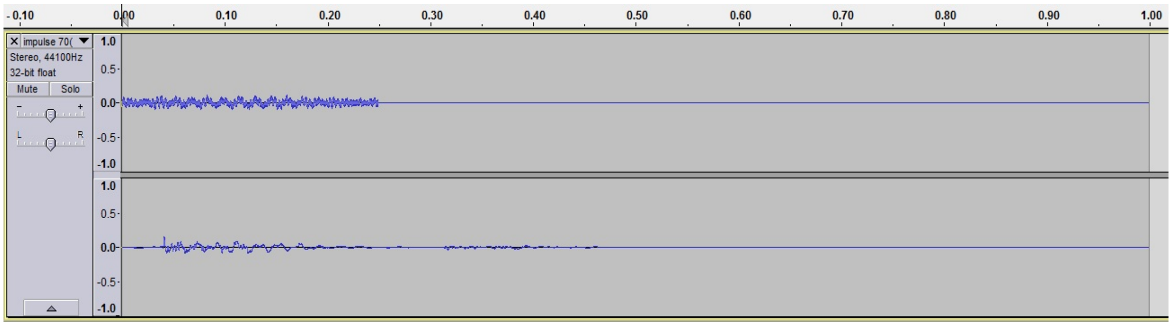


Fig. 5.4. An Audacity screenshot of the "impulse 70(L) + rattle 74(R)" stereo waveform with the impulse (left) channel on top and the rattle (right) channel on the bottom.

5.2.2. Signal Recordings and Measurements

All signals were recorded and measured in the test room at the head position of the subject (mentioned in Section 5.1.1) using a Larson-Davis 824 sound level meter (SLM). The recordings and measurement procedures for the signals are reported in the following subsections.

5.2.2.1. Signal Recordings

All signals, as played back in the room over the Armstrong i-ceiling loudspeaker, the JBL Northridge subwoofer, and the JBL LSR6300 Series studio monitor, were recorded to .wav files using Presonus Studio One recording software for archival purposes. The above mentioned SLM was used as a microphone and connected to a Presonus AudioBox44VSL external sound card. The recording computer was kept in the monitor room as to not add extra noise to the recording. A diagram of the audio playback system and recording system is shown in Figure 5.7.

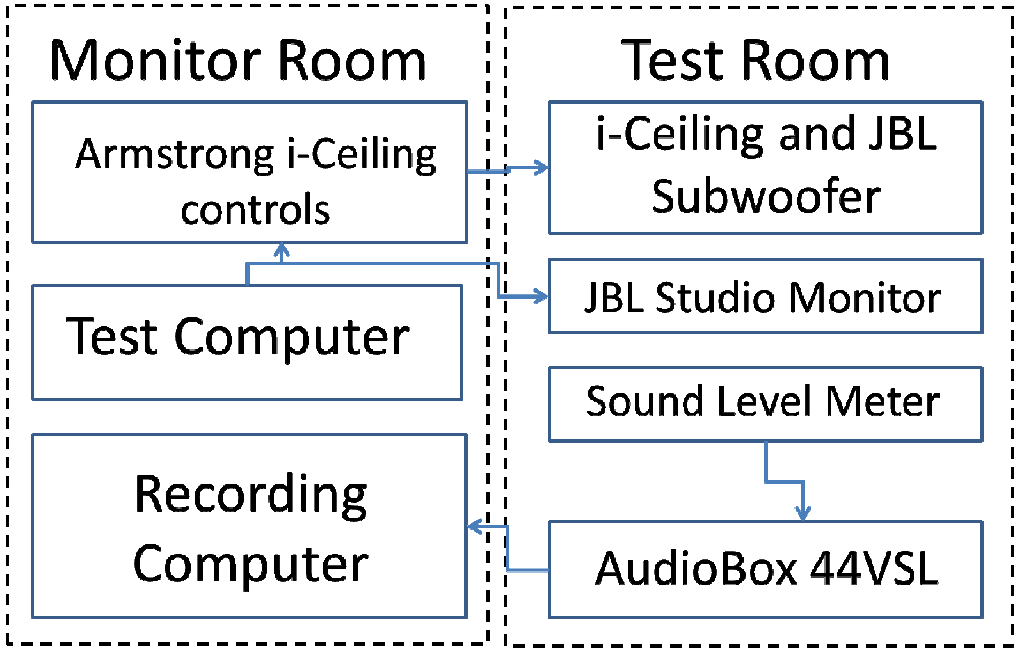


Fig. 5.5. A diagram of the new Nebraska Test Chamber system showing the equipment used for recording test signals in the test room.

5.2.2.2. Signal Measurements

Signals were measured using the Larson-Davis 824 SLM using the settings shown in Figure 5.8. The impulses, rattle, and impulses plus rattle were measured at the one minute mark of a two minute measurement period with the background noise level of RC-29(H) playing constantly throughout. A two minute time period of the RC-29(H) alone was also analyzed.

Sound Level Meter / RTA Settings	
Bandwidth:	1/3
Detector:	Fast
Weighting:	Flat
Peak-1 Weighting:	Flat
Second Display:	TWA
Gain:	0
RTA Detector:	Fast
RTA Weighting:	Flat
Filter Range	12.5-20k

Intervals	
Intervals:	Enabled
Interval Time Sync:	No
Interval Save Ln:	Yes
Interval Save Ln Table:	No
Interval Auto Stop:	Yes
Interval Period:	0:10:20
Interval Threshold:	0
Interval Exchange Rate:	3 dB
Interval Spectra Option:	At Max

Ln	
Ln:	Enabled
Ln Start Level:	15 dB
Spectral Ln Option:	Interval
Ln Percentiles	
Ln Percentiles	
L 1.0	
L 10.0	
L 50.0	
L 90.0	
L 95.0	
L 99.0	

Time History	
Time History:	Enabled
Time History Period:	4
Time History Units:	1/32 seconds
Resolution:	0.1 dB

Fig. 5.6. A list of settings used for measurement of signals with a Larson-Davis 824 SLM.

Measurements were made every 125 ms – the shortest measurement interval available on this SLM. The SLM was set to "fast" mode, and 1/3 octave band data were recorded. The data were then exported to Excel for calculations.

5.2.3. Test Session Procedure

This section details the preparation and implementation of testing procedures. It contains three subsections: test session scheduling, test session design and procedure, and recruitment and orientation procedure.

5.2.3.1. Test Session Scheduling

The overall test consists of an orientation session and eleven regular test sessions, each of which is 30 minutes long. The eleven sessions are broken up into three groups: three sessions with just the RC-29(H) background noise, four sessions with impulse noises, and four sessions with impulse plus rattle noises. Subjects first experience all three sessions with just the background noise, in an attempt to counteract a learning curve that was observed in a previous study (Ainley 2012), before moving on to the sessions with noises. Subjects are only allowed to participate in one session per day. However, a few exceptions were made due to scheduling issues.

The test presentation order was determined with a Latin square design to avoid a test order bias. For the Latin square design, there were eight test sessions and expected to be up to 30 subjects. Three 8x8 squares were used for the first 24 subjects. The order for the last six test subjects was determined with a random order function in Microsoft Excel.

5.2.3.2. Test Session Design and Procedure

This subsection details the design and procedure used for the arithmetic test. Every test session consists of a five minute practice period, a twenty minute test period, and five minutes allotted to fill out a subjective questionnaire. Each five minute practice uses its own unique set of questions. Scores are not recorded during this time, as it is just for the subject to reacquaint themselves with the arithmetic task. After five minutes, the subjects are notified that the practice session is completed and prompted to begin the main test. The arithmetic task utilized is the same as detailed in Section 3.2.3.2.

Due to the specific requirements of controlling borrows, test questions were recycled from the previous study (Ainley 2012), while the remaining questions were

written from scratch. An Excel spreadsheet that automatically calculated the number of times borrowing is required was developed to aid in the writing and repurposing of these questions.

Unique test question sets were created for each practice period and each test period. Unlike the fluctuation tests (Chapter 3), test question sets were not presented with the same noise condition to each subject. This was done to control for bias due to possible difficulty of individual test questions or the whole test question set not accounted for by controlling the number of borrows. Test question sets were named “Test 5”, “Test 6”, “Test 7”, “Test 8”, “Test 5’”, “Test 6’”, “Test 7’”, and “Test 8’” with the prime (‘) tests associated with impulse plus rattle signals and the non-prime tests associated with impulse alone signals. As designed for 30 subjects, subjects one through seven and 29-30 were scheduled to see test order one, subjects eight through 14 were scheduled to see order number two, subjects 15 through 21 were scheduled to see order number three, and subjects 22-28 were scheduled to see order number four. The test number and associated signal presented for all four orders is shown in Figure 5.9.

Order # 1				
w/o rattle:	Test 5	Test 6	Test 7	Test 8
w/ rattle:	Test 5'	Test 6'	Test 7'	Test 8'
Impulse Level:	55 dBA	60 dBA	65 dBA	70 dBA

Order # 3				
w/o rattle:	Test 7	Test 8	Test 5	Test 6
w/ rattle:	Test 7'	Test 8'	Test 5'	Test 6'
Impulse Level:	55 dBA	60 dBA	65 dBA	70 dBA

Order # 2				
w/o rattle:	Test 6	Test 7	Test 8	Test 5
w/ rattle:	Test 6'	Test 7'	Test 8'	Test 5'
Impulse Level:	55 dBA	60 dBA	65 dBA	70 dBA

Order # 4				
w/o rattle:	Test 8	Test 5	Test 6	Test 7
w/ rattle:	Test 8'	Test 5'	Test 6'	Test 7'
Impulse Level:	55 dBA	60 dBA	65 dBA	70 dBA

Fig. 5.7. The four test orders for which test number is associated with which noise condition. Prime (‘) tests are associated with impulse plus rattle signals and the non-prime tests are associated with impulse alone signals. Impulse level refers to either the impulse alone signal or the associated impulse plus rattle signal. For example, 55 dBA refers to both “Impulse 55” and the associated “impulse 55(L) + rattle 59(R)” signal.

As before, based on Woodhead's findings with regards to how quickly subjects can complete this test, the five minute practice session sets contain 20 questions, while the 20 minute session sets contain 45 questions (1964). No subject was able to complete all questions in a practice or regular session within the allotted time.

Specific questions were selected in advance to be linked to impulses and impulses plus rattle (noise conditions). These sound files were played four seconds after the six-digit number appears on the screen, as done in previous tests (Woodhead 1964, Ainley 2012). The noise conditions were designed to randomly occur four times within the first 20 questions of each session so that an average subject would experience the maximum number of noise conditions. Figure 5.10 shows the order of linked questions. Although subjects worked at different paces, almost every subject experienced all four noise conditions presented in each session. There are three instances of a subject only experiencing three noise conditions in a session. The final numbers for each subject's exposure to noise conditions is presented in the next chapter.

Question #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	…	45
Test 5	x							x				x								x		
Test 6				x		x			x						x							
Test 7					x		x						x			x						
Test 8		x					x			x						x						
Test 5'		x						x					x		x							
Test 6'		x							x			x						x				
Test 7'	x					x			x							x						
Test 8'					x			x				x		x								

Fig. 5.8. The order of test questions linked to bursts where X's represent the test questions linked with an impulse for each respective session, and blank cells represent randomized non-impulse-presented questions.

Every noise condition linked question is presented once with an impulse, once with the associated level of impulse plus rattle, and once not linked to any noise condition

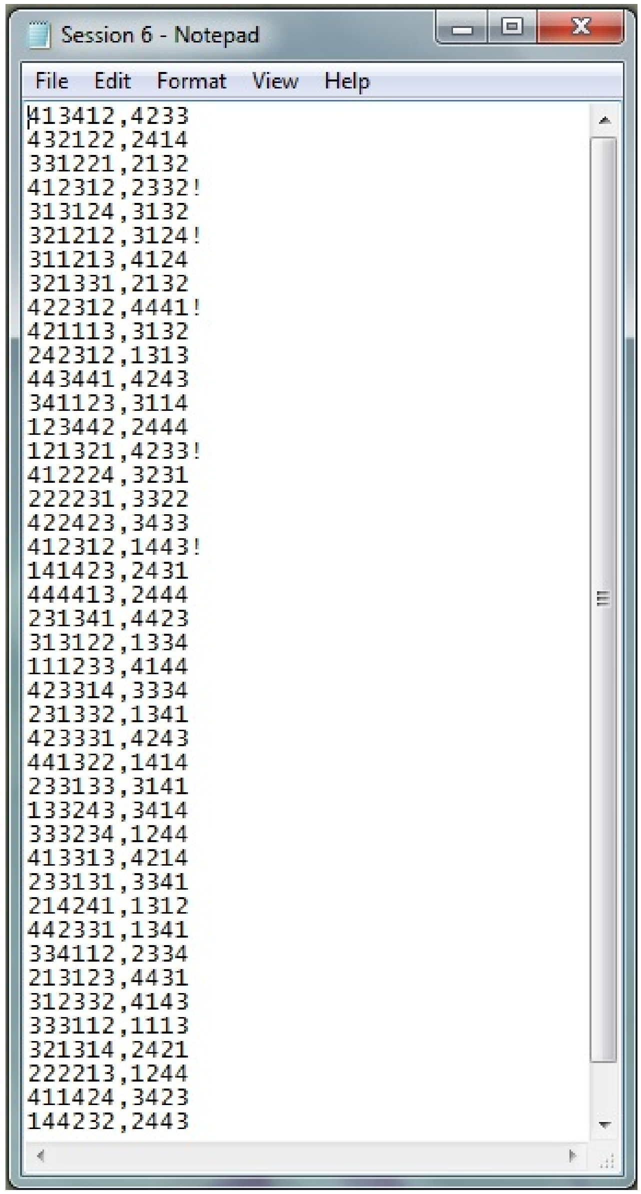
(control question). The control questions are randomly located throughout the first 20 questions of the impulse sessions, not including the session that the impulse linked question is in. Figure 5.11 shows the location of all three iterations of the same questions. Capital letters (e.g. A) refer to impulse presented questions. Capital prime letters (e.g. A') refer to impulse plus rattle presented questions. Lowercase letters (e.g. a) refer to control questions.

Question #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	...	45
Test 5	A	f			j			B	l			C	o						m	D		
Test 6			a	E		F	i		G	c				b	H							
Test 7					I		J	N'	h			g	K			L				p		
Test 8		M			d		N		e	O					k	P						
Test 5'		D'						A'					B'		C'							
Test 6'		E'							G'			H'						F'				
Test 7'	I'					K'			L'							J'						
Test 8'					M'			N'				P'		O'								

Fig. 5.9. The order of impulse-presented questions (capital letter), corresponding impulse + rattle (capital prime letter) and corresponding control questions (lower-case letters). Blank cells represent randomized non-impulse-presented questions.

The same Java program as before was used to automatically conduct the arithmetic test. For this test, the program also triggers the noise condition.

Test questions were written in the form of a .txt file that could be imported into the test program. Each test problem was written on a single row with a comma separating the six-digit and four-digit numbers. An exclamation point at the end of a row indicates a sound file should be triggered during that question. An example text file can be seen in Figure 5.12. A unique text file was required for each five minute practice, each twenty minute main test portion of each session, and two for the orientation session.



Session 6 - Notepad

File Edit Format View Help

```
413412,4233
432122,2414
331221,2132
412312,2332!
313124,3132
321212,3124!
311213,4124
321331,2132
422312,4441!
421113,3132
242312,1313
443441,4243
341123,3114
123442,2444
121321,4233!
412224,3231
222231,3322
422423,3433
412312,1443!
141423,2431
444413,2444
231341,4423
313122,1334
111233,4144
423314,3334
231332,1341
423331,4243
441322,1414
233133,3141
133243,3414
333234,1244
413313,4214
233131,3341
214241,1312
442331,1341
334112,2334
213123,4431
312332,4143
333112,1113
321314,2421
222213,1244
411424,3423
144232,2443
```

Fig. 5.10. Screenshot of the Test 6 questions text file.

The final five minutes of each test session are allotted for the subjects to fill out a subjective questionnaire about their experience in the room on that day. Space is also provided for the subjects to add any additional comments about that specific session.

During their final test session, subjects are additionally asked to complete a noise sensitivity questionnaire – the same as discussed in Section 3.2.3.2.

5.2.3.3. Recruitment and Orientation Procedure

Subjects were recruited by fliers posted in on the University of Nebraska – Omaha campus. The first session each subject participates in is the same orientation session as detailed in Section 3.2.3.3.

No subjects failed to meet the conditions discussed in Section 3.2.3.3 during orientation or failing the hearing screen. However, three subjects dropped out further into testing due to other scheduling issues.

5.2.4. Statistical Analysis

Subjects' performance and perception results were statistically analyzed using Microsoft Excel and SPSS utilizing the same methodology discussed in Section 3.2.4.

Chapter 6: Impulse Plus Rattle Results and Discussion

This chapter presents the analyzed results of the measured test signals, arithmetic task performance, and subjective perception. Test signals are analyzed and reported as discussed in Chapter 5. Task performance and subjective perception results from the tests consisting of short noise bursts with rattle are reported and analyzed using the statistical analysis methodology discussed in Chapter 5.

6.1. Demographic Results

17 subjects participated in this study consisting of eight males and nine females. Two of these subjects also participated in the fluctuations portion of this study. The average subject age was 26 years old, with ages ranging from 19 to 40. Noise sensitivity questionnaires were also filled out during the subjects' final session. The responses were weighted and calculated in to sleep, work, residential, and total noise sensitivity percentages utilizing Schutte et al.'s NoiSeQ-R survey (2007). A histogram of subject responses is shown in Figure 6.1 and results to individual questions are shown in Figure 6.2. Total noise sensitivities ranged from 17% (not very sensitive) to 89% (very sensitive) with an average of 63.1% and a standard error of the mean of 5.7%. Noise sensitivity, gender, and age were considered as additional variables when analyzing complex relationships between noise conditions, task performance, and subjective perception and will be further discussed later in this chapter.

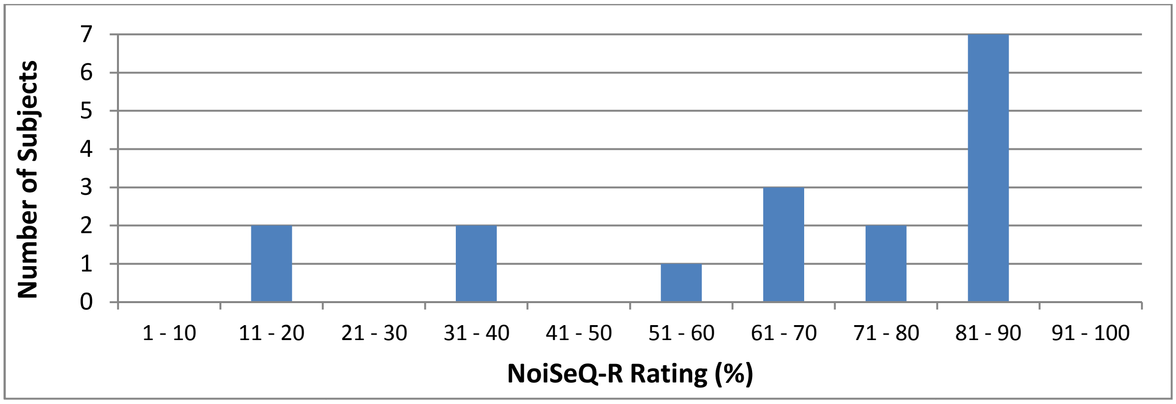


Fig. 6.1. Histogram of subjects total NoiSeQ-R ratings.

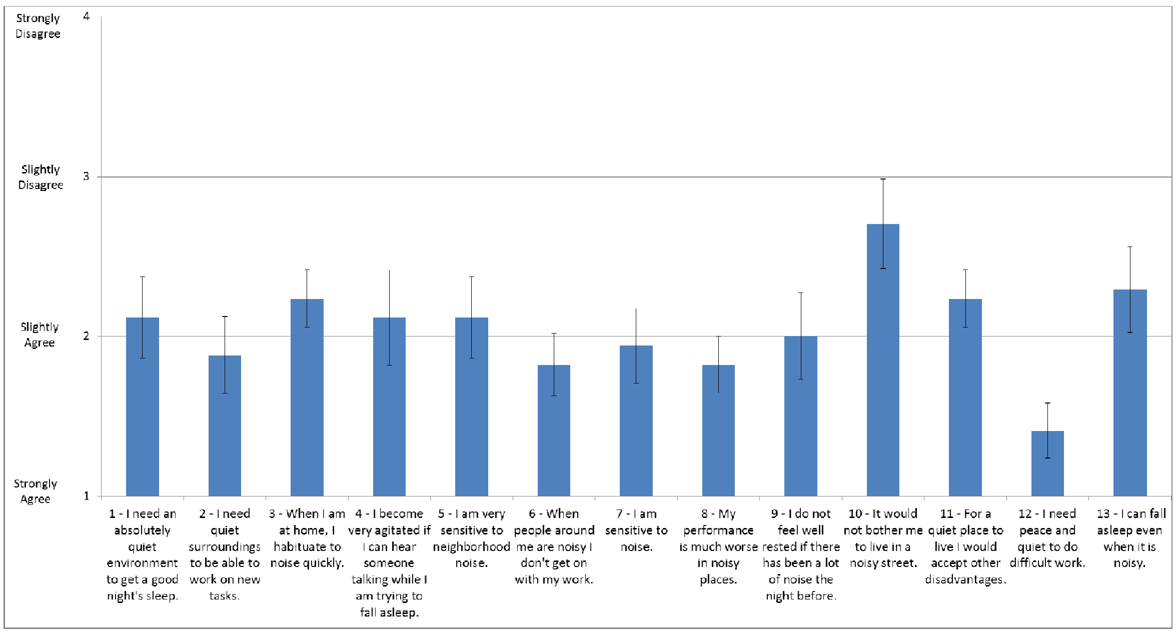


Fig. 6.2. Results of noise sensitivity questionnaires averaged across all subjects. Error bars represent the standard error of the mean.

6.2. Signal Results

6.2.1. Background Noise Results

The background noise level of the test room, as L_{eq} measured over 10 seconds on a Larson-Davis 824 sound level meter, is shown in Figure 6.2. Although the noise level was too low to generate a room criteria (RC) reading on the sound level meter, it can be

reported as an NCB-26. This is too quiet for the purposes of this study, which is one reason why a generated background noise .wav file was implemented.

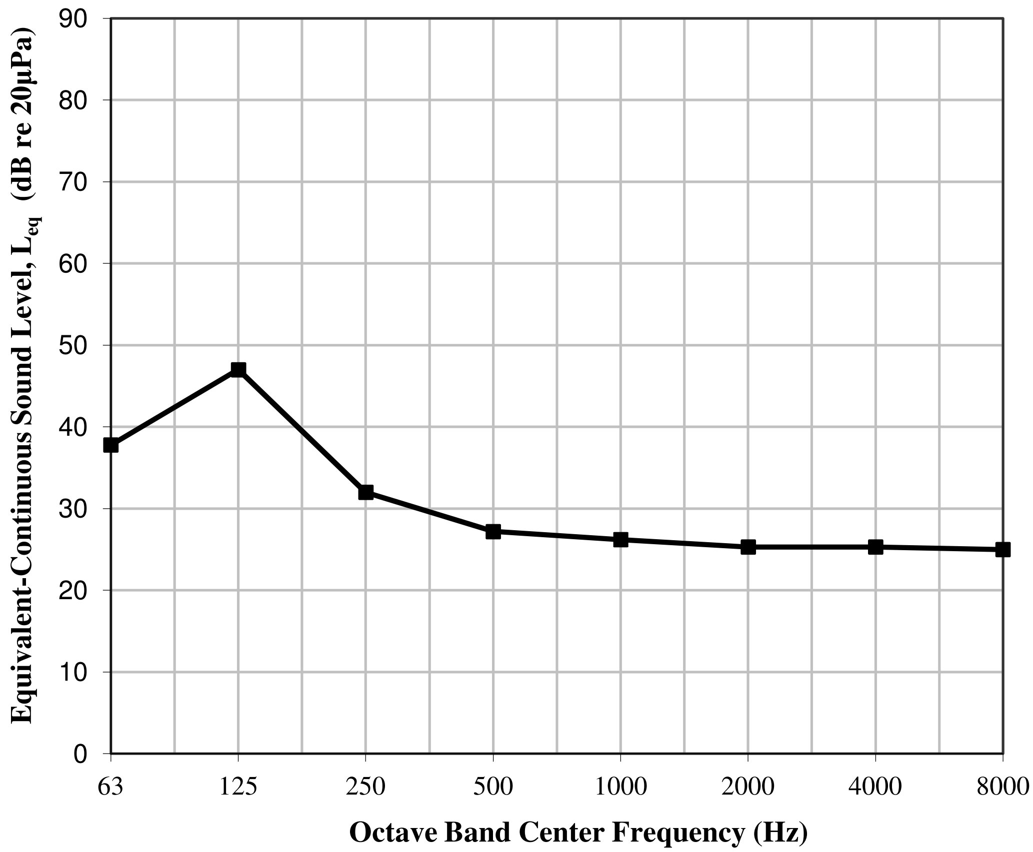


Fig. 6.3. Measurement of Leq across frequency in test room. Results yield an NC-29.

A higher ambient background noise level was generated as discussed in Section 5.2.1.2. Figure 6.4 reports the measured ambient level in the room with the generated noise as measured with a Larson-Davis 824 sound level meter over a 2 minute measurement period. The goal was RC-30, with a final result of RC-30 (H), or 38 dBA.

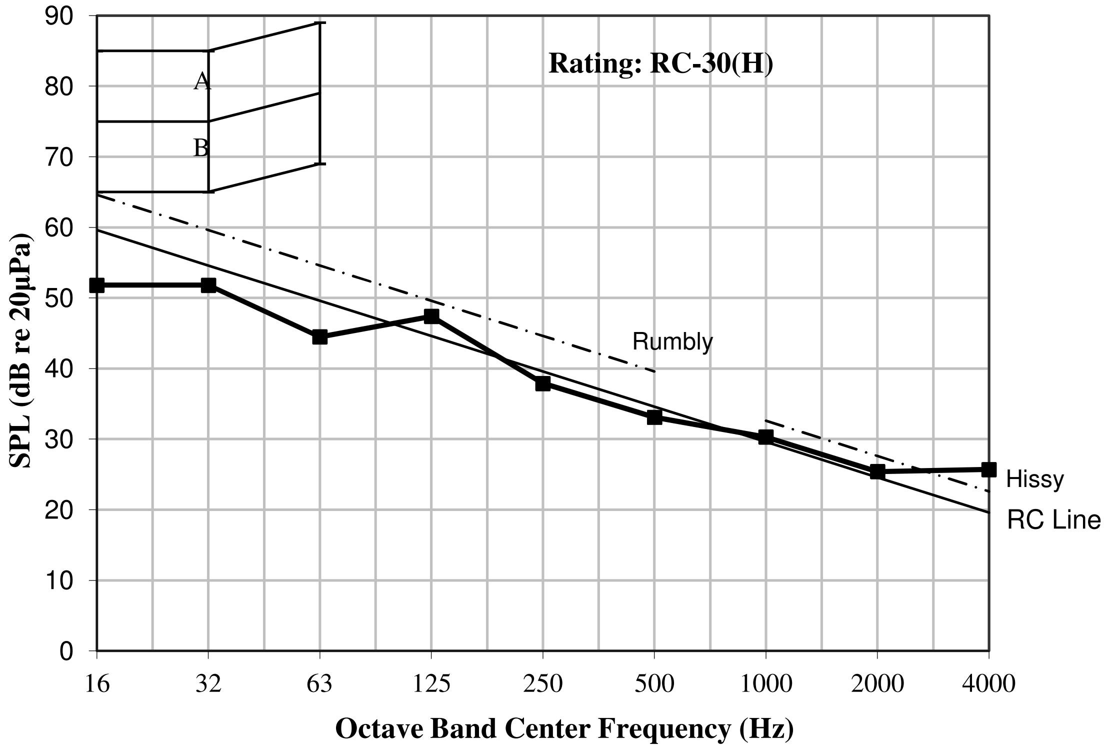


Fig. 6.4. Measurement of L_{eq} across frequency of the ambient BNL .wav when played back in test room. Results yield an RC-30 (H).

6.2.2. Impulse Results

Impulses were measured as discussed in Section 5.2.2.2. Original impulse calibrations, made while the signal was played continuously, had overall sound pressure levels with A-weightings to be approximately 55, 60, 65 and 70 dBA. Therefore, the impulses were titled, “Impulse 55”, “Impulse 60”, “Impulse 65”, and “Impulse 70”, respectively.

Peak sound pressure levels (L_{pk}) were analyzed for each impulse. Total peak values correspond to the maximum instantaneous sound pressure level during the measurement. The spectral L_{pk} results of each impulse are shown in Figure 6.5.

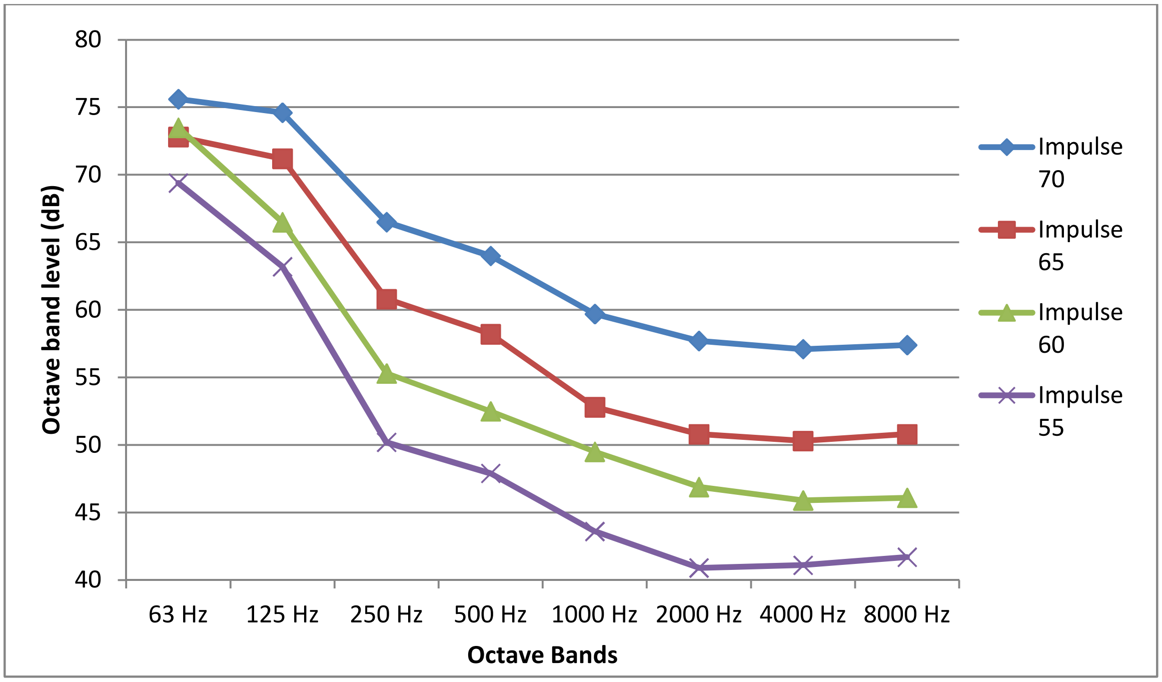


Fig. 6.5. Spectra of peak readings for the four impulse signals. Results yield overall peak SPL of 70, 74, 75, and 79 dB respectively.

The peak A-weighted sound pressure levels (L_{Apk}) were additionally analyzed for each impulse. These values correspond to the maximum instantaneous A-weighted sound pressure level during the measurement. Note that the L_{pk} and the L_{Apk} do not necessarily occur at the same point in time during the measurement. The overall L_{Apk} for the four levels of the impulse were 53, 57, 62, and 67 dBA, and the spectral L_{Apk} results of each signal are shown in Figure 6.6.

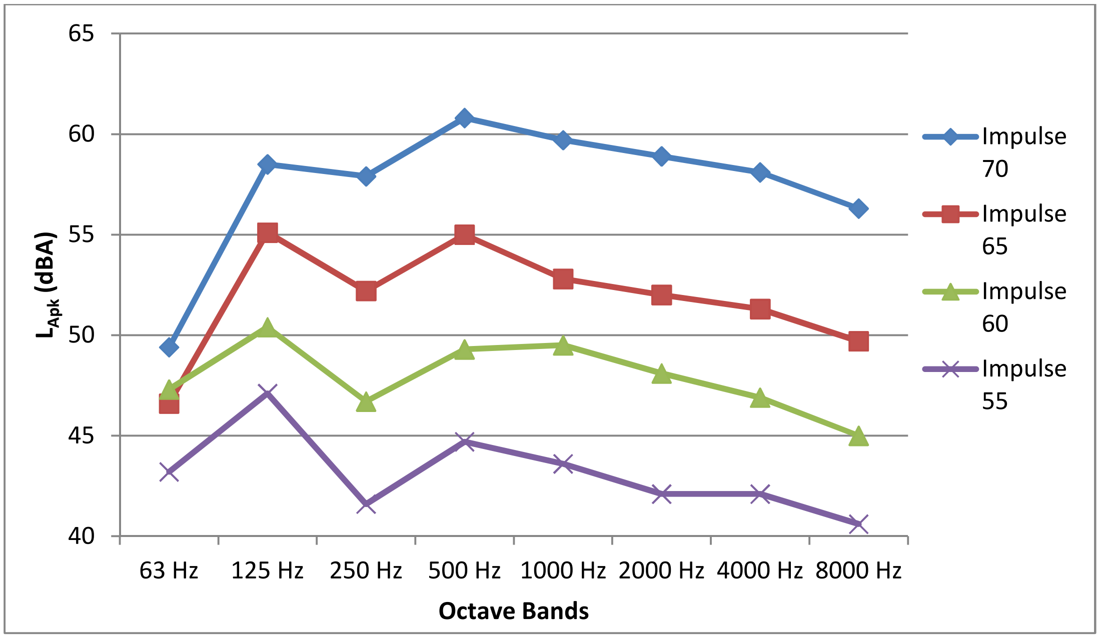


Fig. 6.6. Spectra of peak A-weighted readings for the four impulse signals. Results yield overall A-weighted peak SPL of 53, 57, 62, and 67 dBA respectively.

The ambient background noise, RC-30 (H), was also analyzed for its L_{Apk} values which was 40 dBA.

6.2.3. Rattle Results

Rattle signals were measured as discussed in Section 5.2.2.2. Original rattle calibrations were made by continuously playing the individual rattle signal until the L_{Apk} of the overall signal met the desired level. The spectral content of the rattle signals as created and as measured can be seen in Figure 6.7.

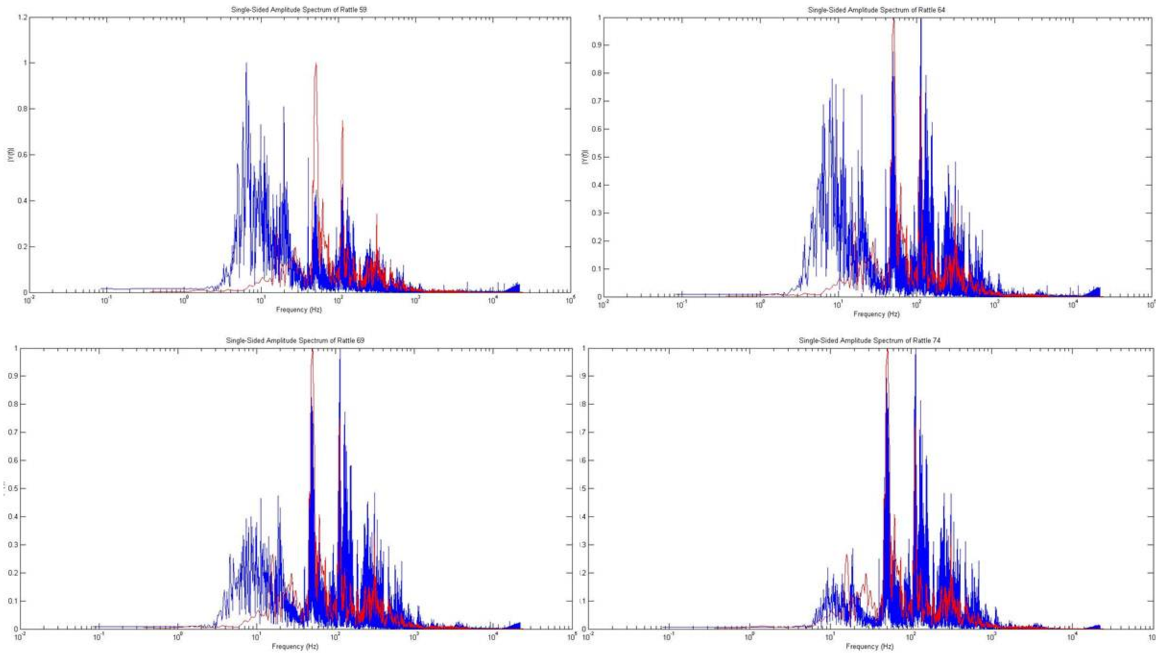
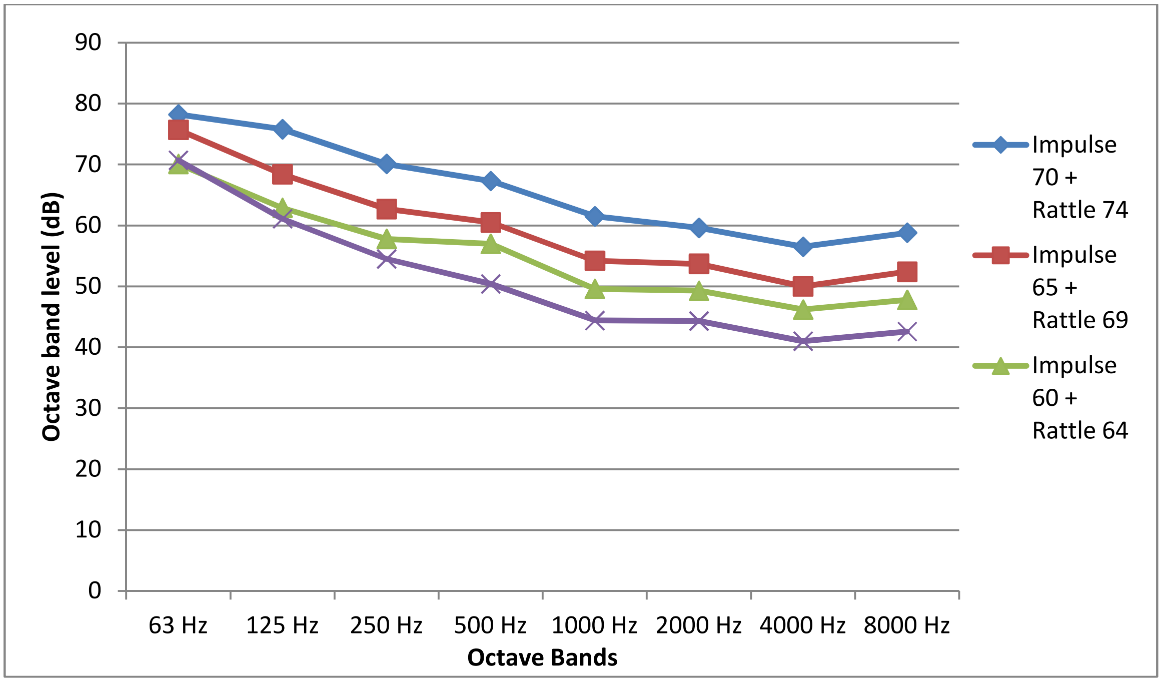


Fig. 6.7. Spectra rattle signals as generated (red) and as recorded in the test facility (blue).

6.2.4. Impulse plus Rattle Results

Impulse plus rattle signals were measured as discussed in Section 5.2.2.2. Original impulse plus rattle calibrations were made by continuously playing each element (impulse and rattle) separately and then combining the signals after calibration. The overall sound pressure levels of the impulse with A-weightings were measured to be approximately 55, 60, 65 and 70 dBA. The overall peak sound pressure levels of the rattles with A-weightings were measured to be approximately 59, 64, 69, and 74 dBA. Therefore, the impulses plus rattle signals were titled, “Impulse 55 + Rattle 59”, “Impulse 60 + Rattle 64”, “Impulse 65 + Rattle 69”, and “Impulse 70 + Rattle 74”, respectively.

Peak sound pressure levels (L_{pk}) were analyzed for each impulse plus rattle signal. Total peak values correspond to the maximum instantaneous sound pressure level during the measurement. The overall L_{pk} for the four levels of the impulse plus rattle were 71, 72, 77, and 81, and the spectral L_{pk} results of each signal are shown in Figure 6.8.



<u>Fig. 6.8.</u> Spectra of peak readings for the four impulse signals. Results yield overall peak SPL of 71, 72, 77, and 81 dB respectively.

The peak A-weighted sound pressure levels (L_{Apk}) were additionally analyzed for each impulse plus rattle signal. These values correspond to the maximum instantaneous A-weighted sound pressure level during the measurement. Note that the L_{pk} and the L_{Apk} do not necessarily occur at the same point in time during the measurement. The overall L_{Apk} for the four signals were 54, 58, 63, and 70 dBA, and the spectral L_{Apk} results of each signal are shown in Figure 6.9.

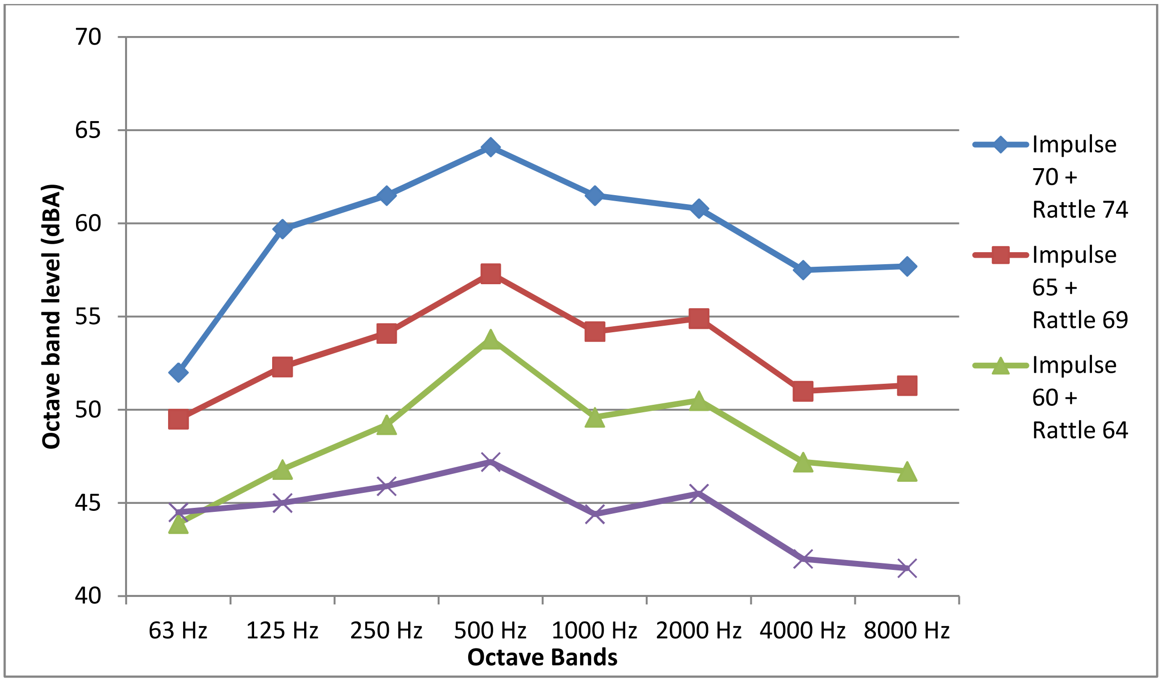


Fig. 6.9. Spectra of peak A-weighted readings for the four impulse plus rattle signals. Results yield overall peak SPL of 54, 58, 63, and 70 dBA respectively.

6.3. Task Performance Results

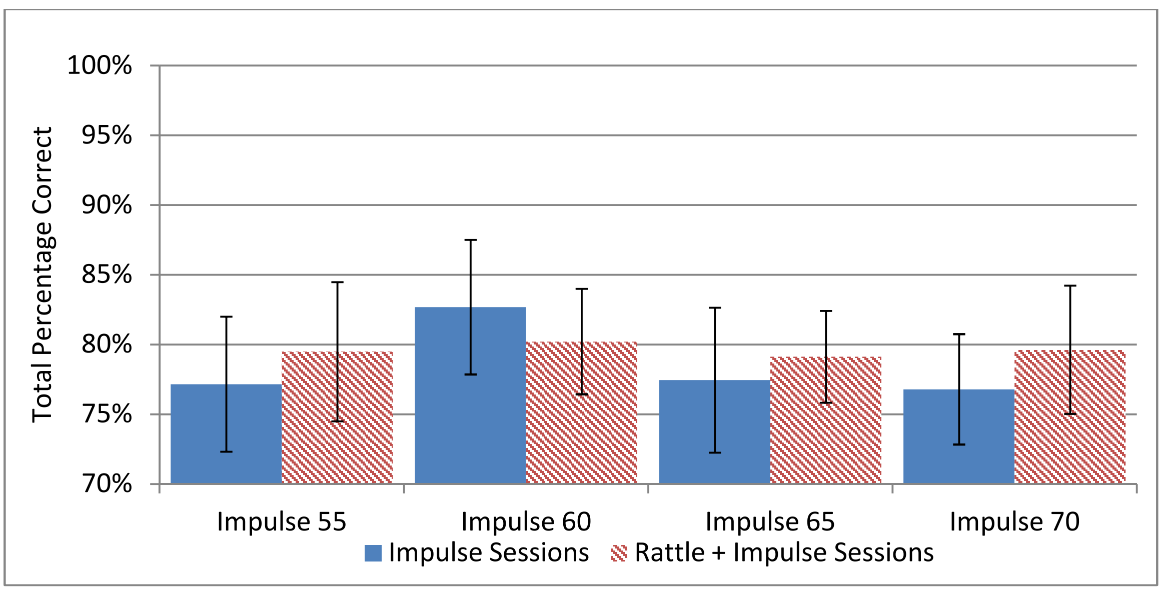
Task performance was measured via the total percentage of correct answers and the average response time, in seconds, for each question. Statistical analysis was performed utilizing SPSS statistical software as discussed in Section 5.2.4. Results were tested for normal distribution via a Kolmogorov-Smirnov test and all were found to be non-normally distributed. Because of this, Friedman ANOVA and Spearman correlation coefficient are used in addition to the Pearson coefficient. Wilcoxon tests were used to analyze any relationships between noise conditions. A Bonferroni correction was applied, and all effects reported at a 0.05 level of significance (Field and Hole 2003).

Although it is a parametric test, repeated measures ANOVA results are also reported to back up the non-parametric test results. Observed power, at $\alpha = 0.05$, is reported for each.

6.3.1. Task Performance Results across Noise Conditions

The overall task performance results across all analyzed test sessions are shown in Figures 6.10 and 6.11 as well as Tables 6.1 and 6.2. All results exhibited a non-normal distribution, as concluded by a Kolmogorov-Smirnov test. Therefore, the non-parametric Friedman ANOVA test was used to analyze these relationships, and Wilcoxon tests were used to further analyze the relationships between each noise condition.

No apparent general trends can be seen. The total percentage of correct answers all fall between 75% and 85%. All average response times fall between 20 and 25 seconds. The Wilcoxon test results show no significant relationships between the total percentage of correct answers or the average time taken for each question across different impulse sessions and different impulse plus rattle sessions.



<u>Fig. 6.10.</u> Overall percentage of correct answers for each test session. Solid represents sessions where impulses were presented alone. Hashed represents sessions where impulses were presented with the rattle noise. Error bars represent the standard error of the mean.

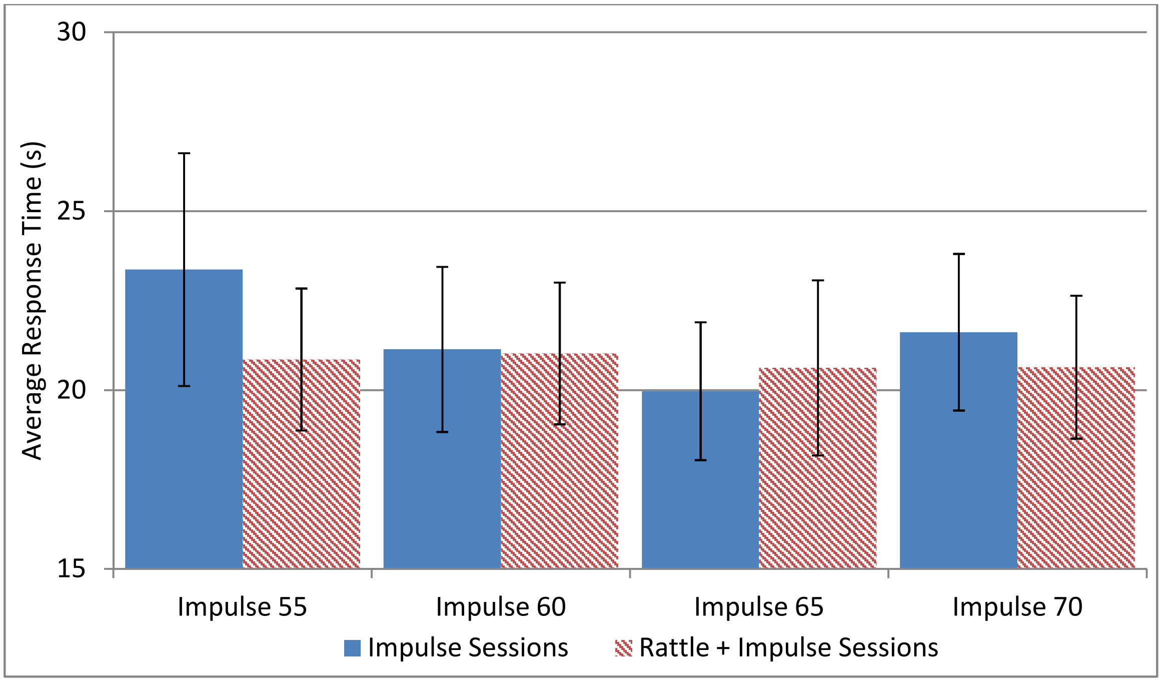


Fig. 6.11. Average response time, in seconds, for test questions in each test session averaged across all test sessions. Blue represents sessions where impulses were presented alone. Red represents sessions where impulses were presented with the rattle noise. Error bars represent the standard error of the mean.

Table 6.1. Wilcoxon Results between Noise Condition and Total Percentage Correct. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.05 level of significance.

Total percentage correct between sessions								
		Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55	T	*41.00*	*73.00*	*56.00*	*72.00*	*47.00*	*73.00*	*62.00*
	sig	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
	effect size	*-0.29*	*-0.03*	*-0.11*	*-0.04*	*-0.24*	*-0.03*	*-0.12*
Impulse 60	T		*47.00*	*31.00*	*58.00*	*46.00*	*40.00*	*44.00*
	sig		*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
	effect size		*-0.24*	*-0.33*	*-0.15*	*-0.14*	*-0.25*	*-0.21*
Impulse 65	T			76.00	57.00	64.00	72.00	56.00
	sig			ns	ns	ns	ns	ns
	effect size			0.00	-0.10	-0.10	-0.04	-0.11
Impulse 70	T				56.00	42.00	70.00	50.00
	sig				ns	ns	ns	ns
	effect size				-0.17	-0.23	-0.05	-0.16
Impulse 55 + Rattle 59	T					66.50	59.00	68.00
	sig					ns	ns	ns
	effect size					-0.08	-0.14	-0.07
Impulse 60 + Rattle 64	T						55.00	52.00
	sig						ns	ns
	effect size						-0.05	-0.08
Impulse 65 + Rattle 69	T							71.00
	sig							ns
	effect size							-0.04

** = Indicates the mean ranks between noise conditions is significant, $p<0.005$, ns = not significant.

Table 6.2. Wilcoxon Results between Noise Condition and Average Response Time. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.05 level of significance.

Average response time between sessions								
		Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55	T	*55.00*	*53.00*	*66.00*	*56.00*	*52.00*	*54.00*	*56.00*
	sig	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
	effect size	*-0.17*	*-0.19*	*-0.09*	*-0.17*	*-0.20*	*-0.18*	*-0.17*
Impulse 60	T		*55.00*	*66.00*	*66.00*	*76.00*	*74.00*	*65.00*
	sig		*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
	effect size		*-0.17*	*-0.09*	*-0.09*	*0.00*	*-0.02*	*-0.09*
Impulse 65	T			49.00	57.00	63.00	64.00	64.00
	sig			ns	ns	ns	ns	ns
	effect size			-0.17	-0.16	-0.11	-0.10	-0.10
Impulse 70	T				64.00	60.00	63.00	66.00
	sig				ns	ns	ns	ns
	effect size				-0.10	-0.13	-0.11	-0.09
Impulse 55 + Rattle 59	T					74.00	76.00	75.00
	sig					ns	ns	ns
	effect size					-0.02	0.00	-0.01
Impulse 60 + Rattle 64	T						65.00	73.00
	sig						ns	ns
	effect size						-0.03	-0.03
Impulse 65 + Rattle 69	T							75.00
	sig							ns
	effect size							-0.01

** = Indicates the mean ranks between noise conditions is significant, p<0.005, ns = not significant.

6.3.2. Comparison between Test Questions Linked to Impulses, the Same Test Questions Linked to Impulses plus Rattle, and the Same Test Questions Presented Without Impulses or Rattle

Specific test questions were also linked to impulse-presented questions and impulse-plus-rattle-presented questions. These linked questions came in sets of three with the same question linked to an impulse-presented question, an impulse-plus-rattle-presented question, and a question presented without any impulse or rattle. The impulse and impulse plus rattle questions were presented between three and four times in a given session depending on the subject's pace while answering questions. There were only three out of 136 total sessions where a subject experienced three impulses. If a subject did not complete a certain impulse or impulse-plus-rattle-presented question, the results of the corresponding questions were removed from the final analysis. The total

percentage correct and average response time, in seconds, for questions linked to an impulse, the same question linked to an impulse plus rattle, and the same questions presented without an impulse are shown in Figures 6.12 and 6.13, respectively.

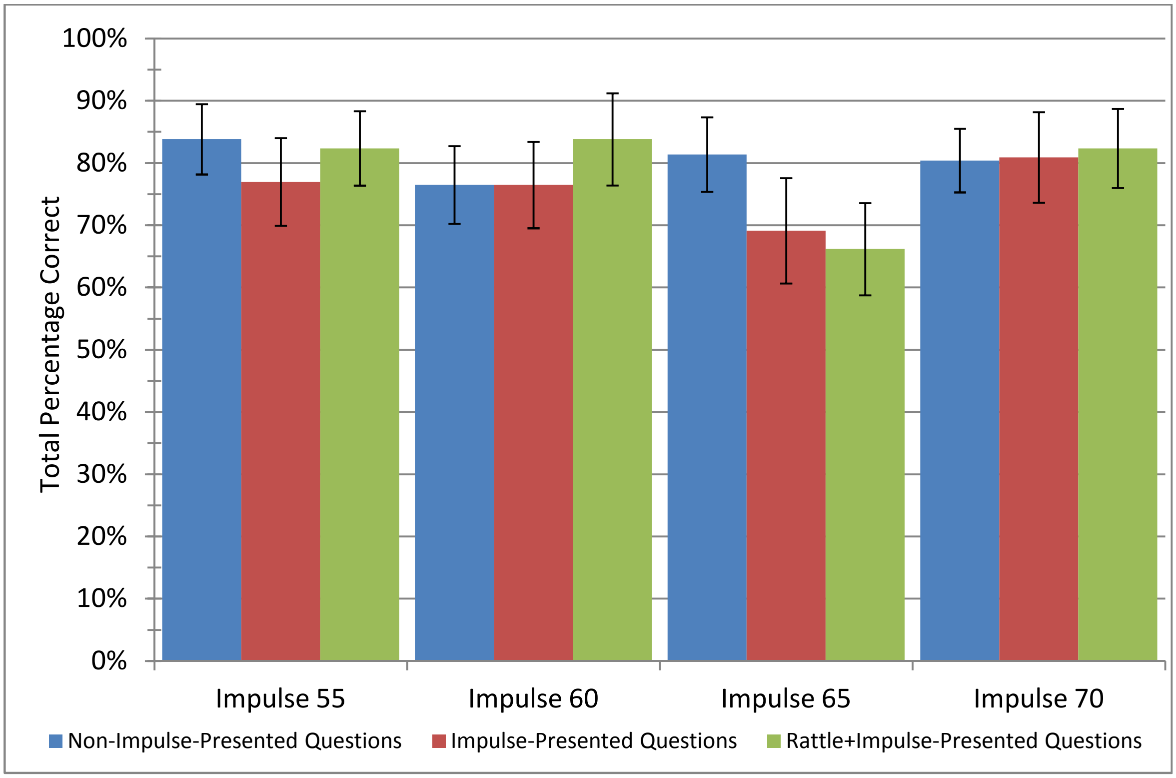


Fig. 6.12. Results of the overall percentage of correct answers for questions linked to respective impulses, impulses plus rattle, and again for those same questions when presented without an impulse. Error bars represent the standard error of the mean.

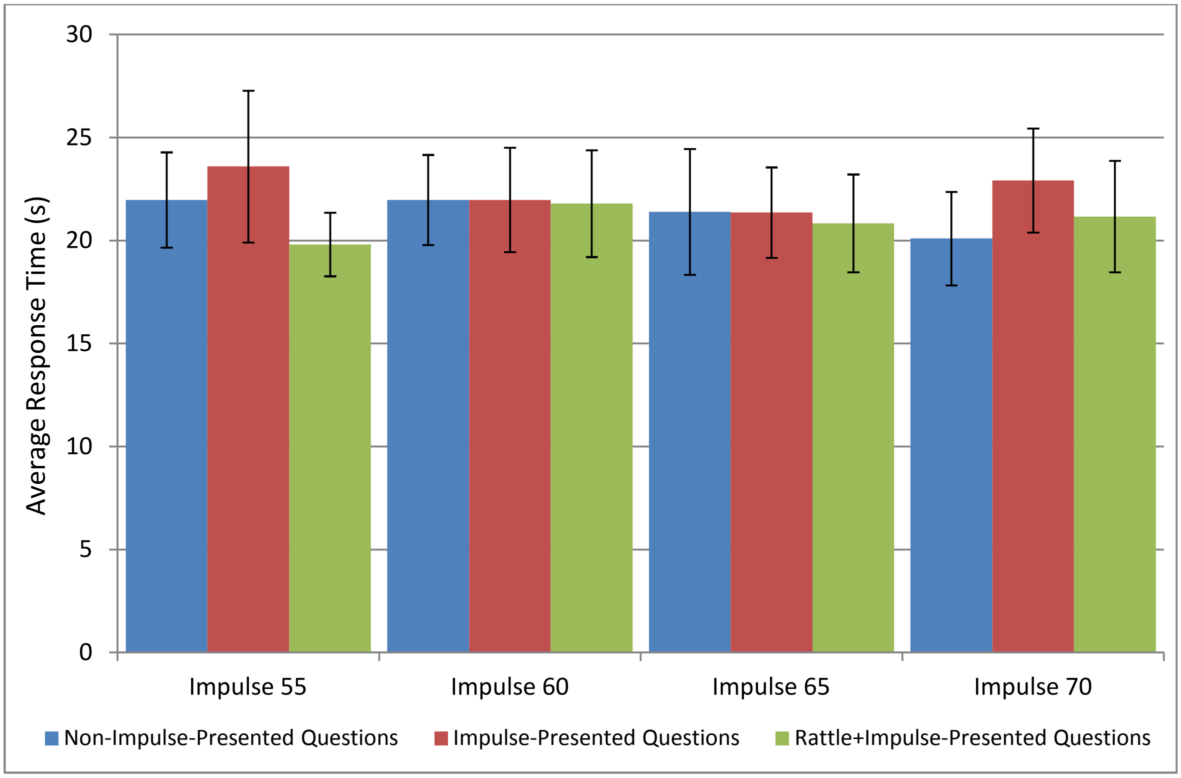


Fig. 6.13. Results of the average time taken in seconds for questions linked to respective impulses, impulse plus rattle, and again for those same questions when presented without an impulse. Error bars represent the standard error of the mean.

Again, there seem to be no apparent trends in this data. The relationships across test sessions were further analyzed using repeated measures ANOVA and Friedman ANOVA. All statistical tests found no significant relationships across test sessions. Previous research found no statistically significant relationships between the performance between impulse-presented questions and non-impulse-presented questions (Ainley 2012). However, Ainley found trends showing a decrease in total percentage of correct answers of impulse-presented questions from non-impulse-presented questions.

6.3.3. Comparisons of Task Performance to Subjective Perception

Task performance and subjective perception were compared using Pearson and Spearman Correlation Coefficients and a linear mixed model analysis with results shown

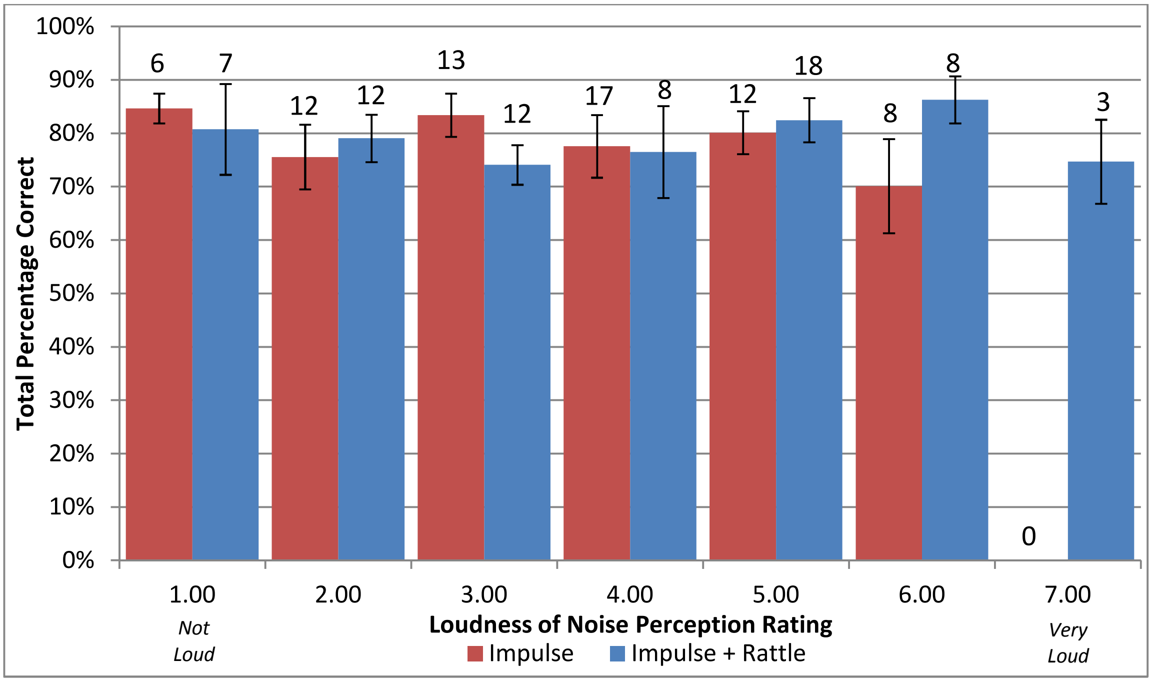
in Table 6.3. No significant relationships were found with the Pearson and Spearman Correlation Coefficients. However, a few were found with the linear mixed model analysis results, at $p < 0.05$.

Table 6.3. Correlations and Linear Mixed Model Analysis between Subjective Perception and Task Performance. The linear mixed model F values, Pearson correlation coefficients, and Spearman correlation coefficients between subjective perception of noise and performance of the task. Results are split between impulse sessions and impulse plus rattle sessions.

		Subject Questionnaire Results				
Task Performance Results	Statistical Measure	Loudness	Change in Noise Over Time	Rumble	Annoyance	Distraction
Impulse Average Time	$F_{1,61}$	ns	2.64*	ns	ns	ns
	Pearson (r)	ns	ns	ns	ns	ns
	Spearman (r)	ns	ns	ns	ns	ns
Impulse Total % Correct	$F_{1,61}$	ns	ns	ns	3.13*	2.57*
	Pearson (r)	ns	ns	ns	ns	ns
	Spearman (r)	ns	ns	ns	ns	ns
Impluse + Rattle Average Time	$F_{1,61}$	2.30*	ns	ns	ns	ns
	Pearson (r)	ns	ns	ns	ns	ns
	Spearman (r)	ns	ns	ns	ns	ns
Impluse + Rattle Total % Correct	$F_{1,61}$	ns	ns	ns	ns	ns
	Pearson (r)	ns	ns	ns	ns	ns
	Spearman (r)	ns	ns	ns	ns	ns

**significant at p<0.01, *significant at p<0.05, ns = not significant

The relationship between the loudness perception ratings and corresponding total percentage correct are shown in Figure 6.14. They show no correlation.



<u>Fig. 4.14.</u> Results of the loudness perception rating and the corresponding total percentage of correct responses given in a session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the loudness perception ratings and corresponding average response times are shown in Figure 6.15. The linear mixed model analysis results show a small, positive correlation in the impulse plus rattle ratings, meaning when ratings for loudness of noise increased, the average response time generally increased.

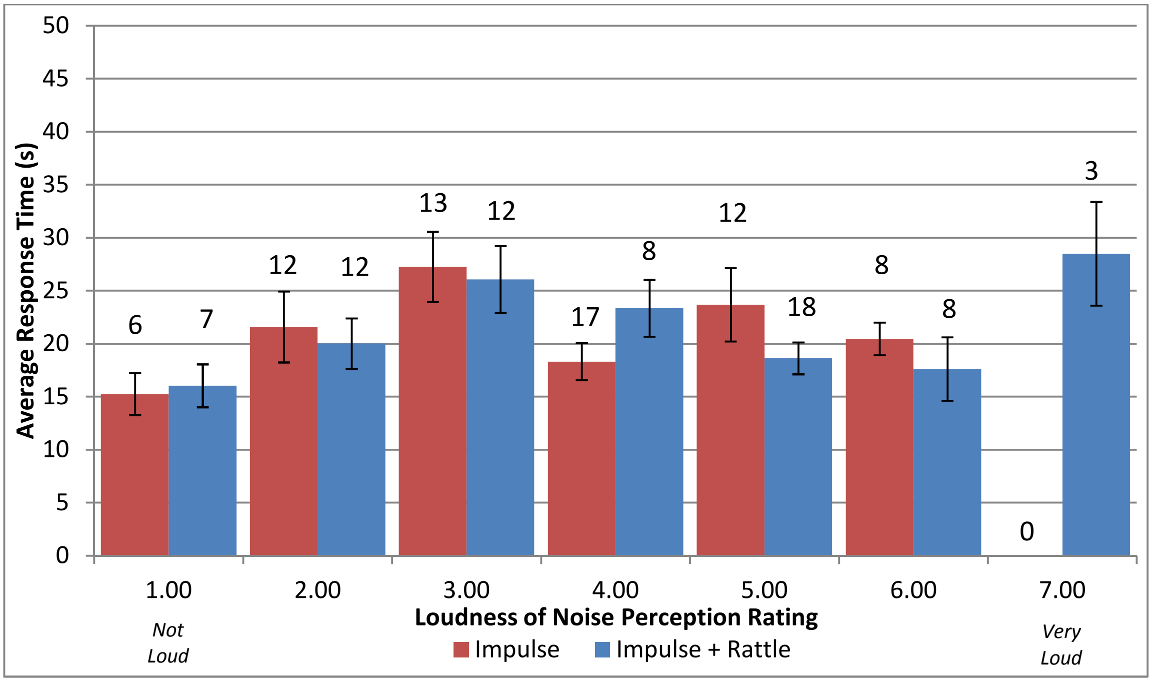
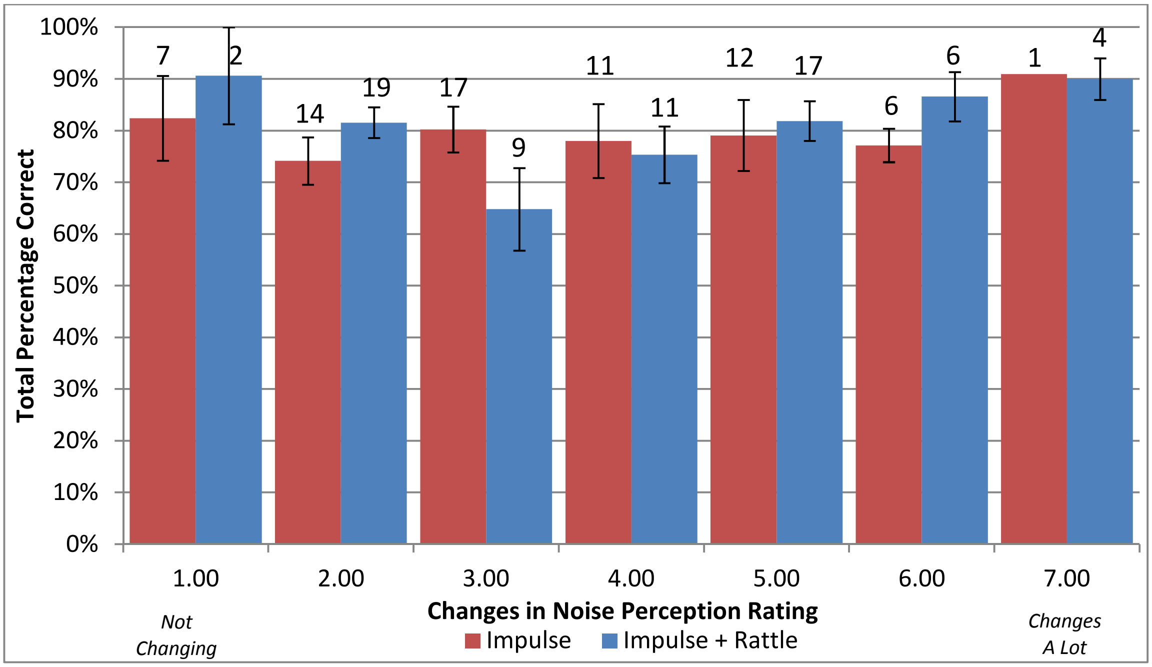


Fig. 4.15. Results of the perception ratings of changes in noise and the corresponding average time taken to solve task problems in a given session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the changes in noise perception ratings and corresponding total percentage correct are shown in Figure 6.16. They show no correlation.



<u>Fig. 6.16.</u> Results of the perception ratings of changes in noise and the corresponding total percentage of correct responses given in a session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the changes in noise perception ratings and corresponding average response times are shown in Figure 6.17. The linear mixed model analysis results show a small, positive correlation in the impulse ratings, meaning when ratings for changes in noise increased, the average response time generally increased.

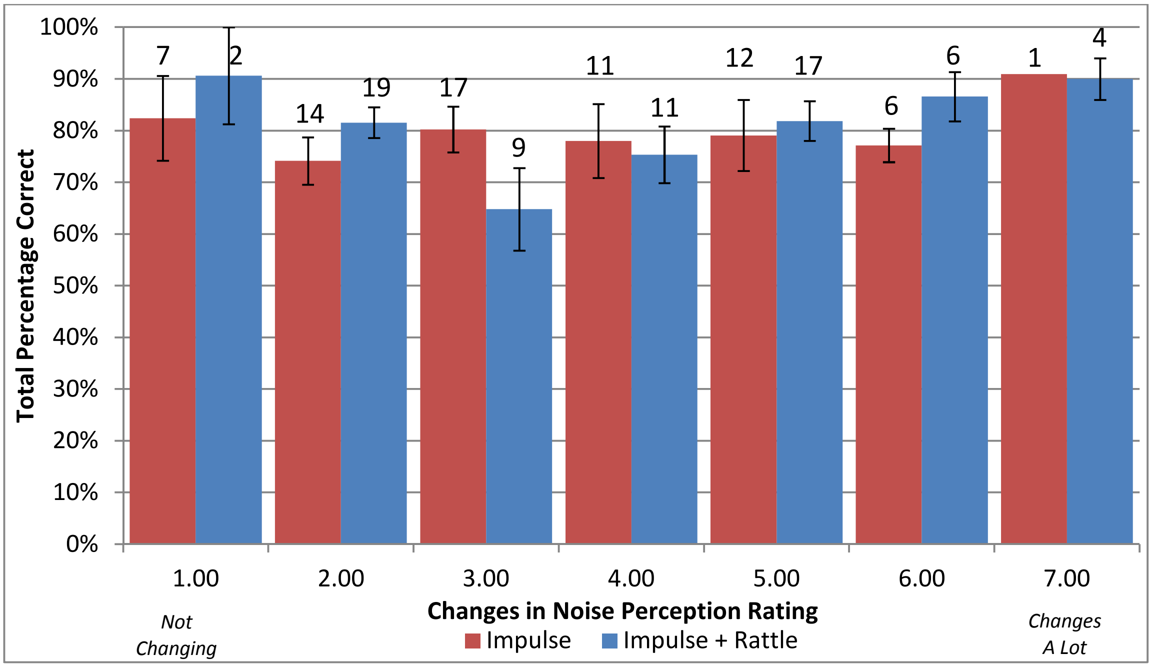
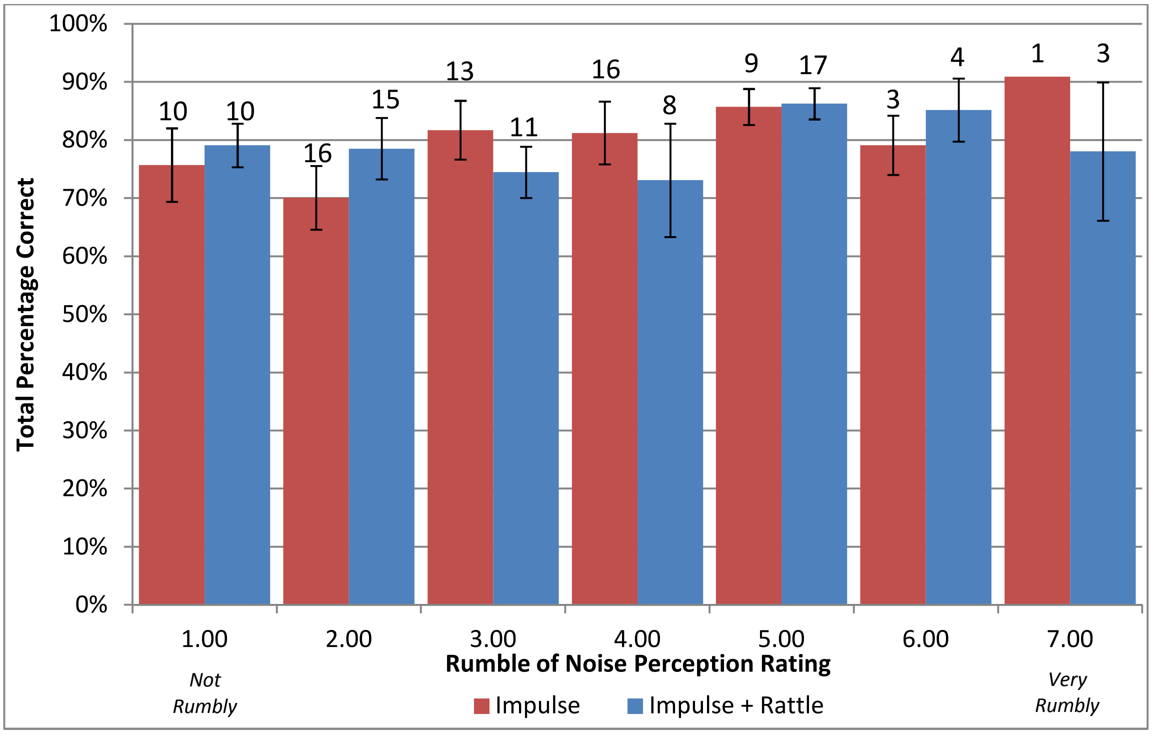


Fig. 6.17. Results of the perception ratings of changes in noise and the corresponding average time taken to solve task problems in a given session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the rumble of noise perception ratings and corresponding total percentage correct are shown in Figure 6.18, and the relationship between the rumble of noise perception ratings and corresponding average response times are shown in Figure 6.19. No correlation is found in either relationship.



<u>Fig. 6.18.</u> Results of the perception ratings of rumble of noise and the corresponding total percentage of correct responses given in a session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

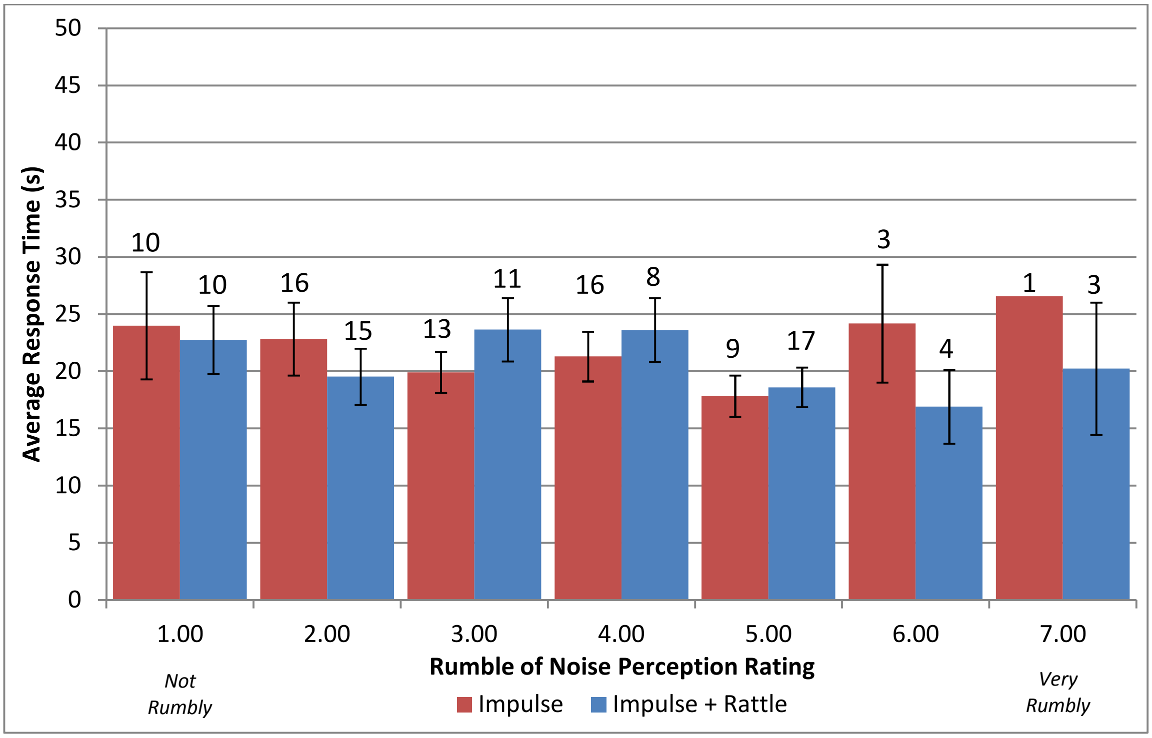


Fig. 6.19. Results of the perception ratings of rumble of noise and the corresponding average time taken to solve task problems in a given session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the annoyance to noise perception ratings and corresponding total percentage correct are shown in Figure 6.20. They show a small, negative correlation that means when ratings for annoyance to noise increased toward very annoying, the total percentage of questions answered correctly generally decreased.

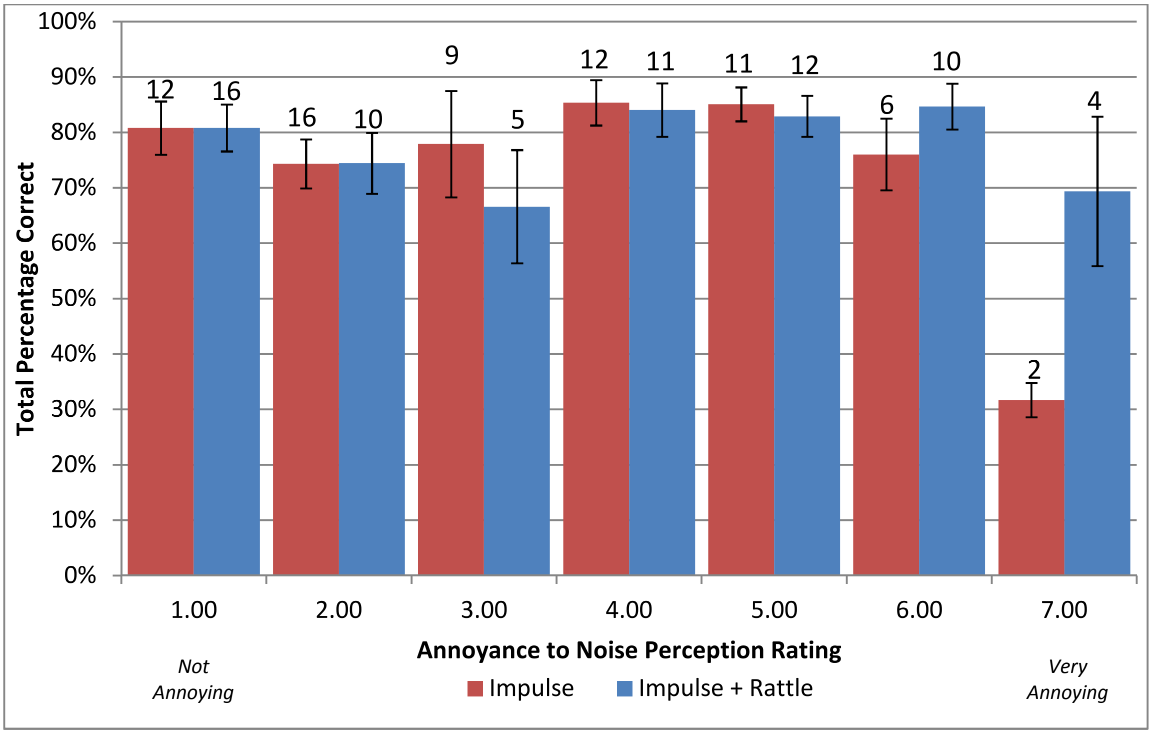


Fig. 6.20. Results of the perception ratings annoyance to noise and the corresponding total percentage of correct responses given in a session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the annoyance to noise perception ratings and corresponding average response times are shown in Figure 6.21. The linier mixed model analysis results show a small, negative correlation in the impulse ratings, meaning when ratings for annoyance to noise increased, the total percent correct generally decreased.

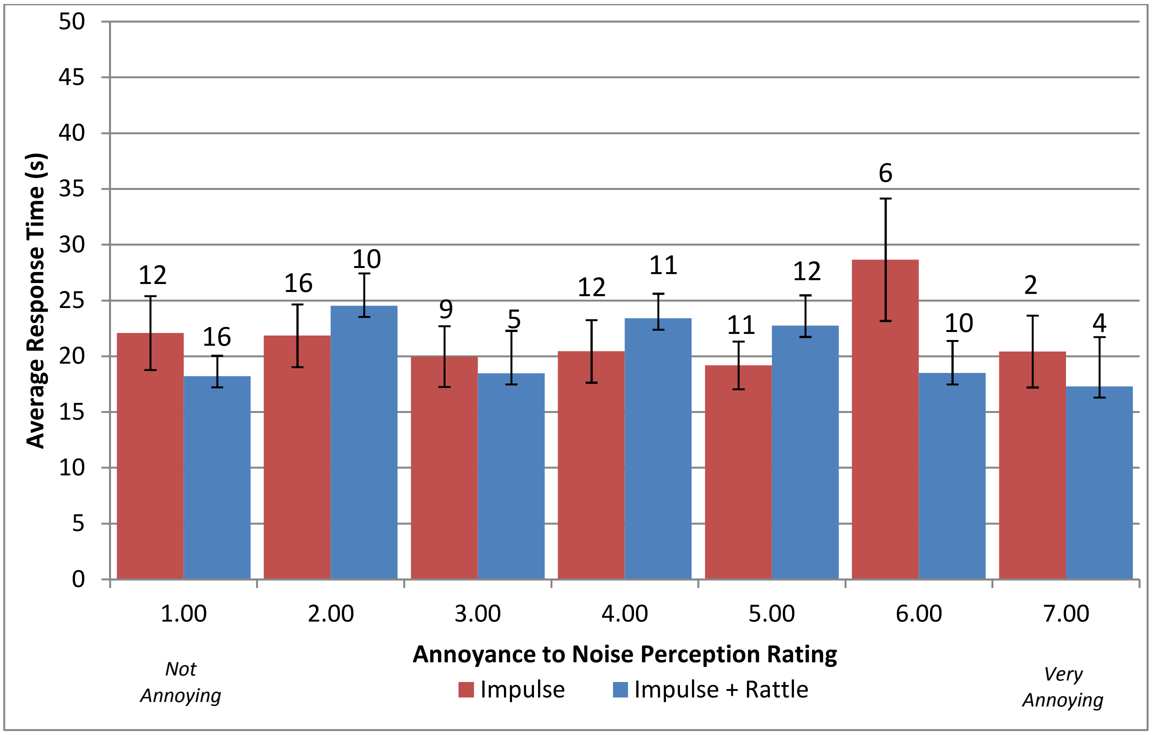


Fig. 6.21. Results of the perception ratings of annoyance to noise and the corresponding average time taken to solve task problems in a given session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the distraction of noise perception ratings and corresponding total percentage correct are shown in Figure 4.22. They show a small, not statistically significant, negative trend for both impulse and impulse-plus-rattle-presented questions. When ratings for distraction of noise increased, the total percentage of correct generally decreased.

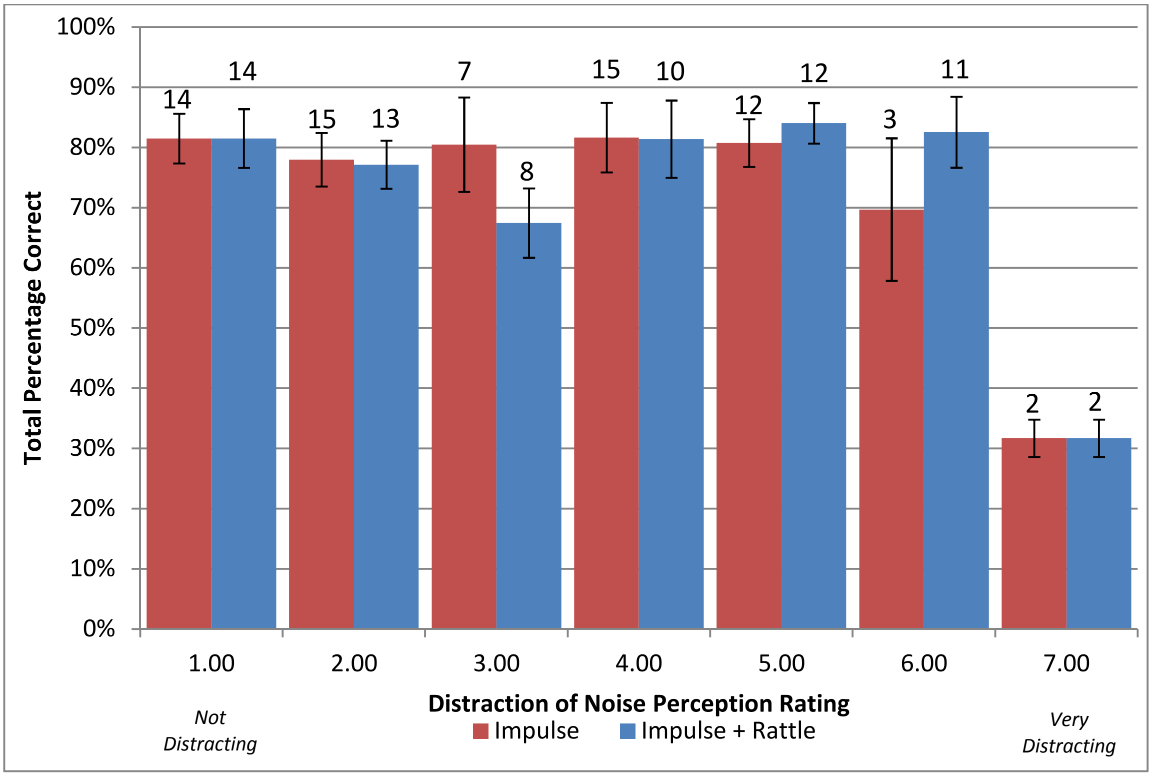


Fig. 6.22. Results of the perception ratings of distraction of noise and the corresponding total percentage of correct responses given in a session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

The relationship between the distraction of noise perception ratings and corresponding average response times are shown in Figure 6.23. They show no correlation.

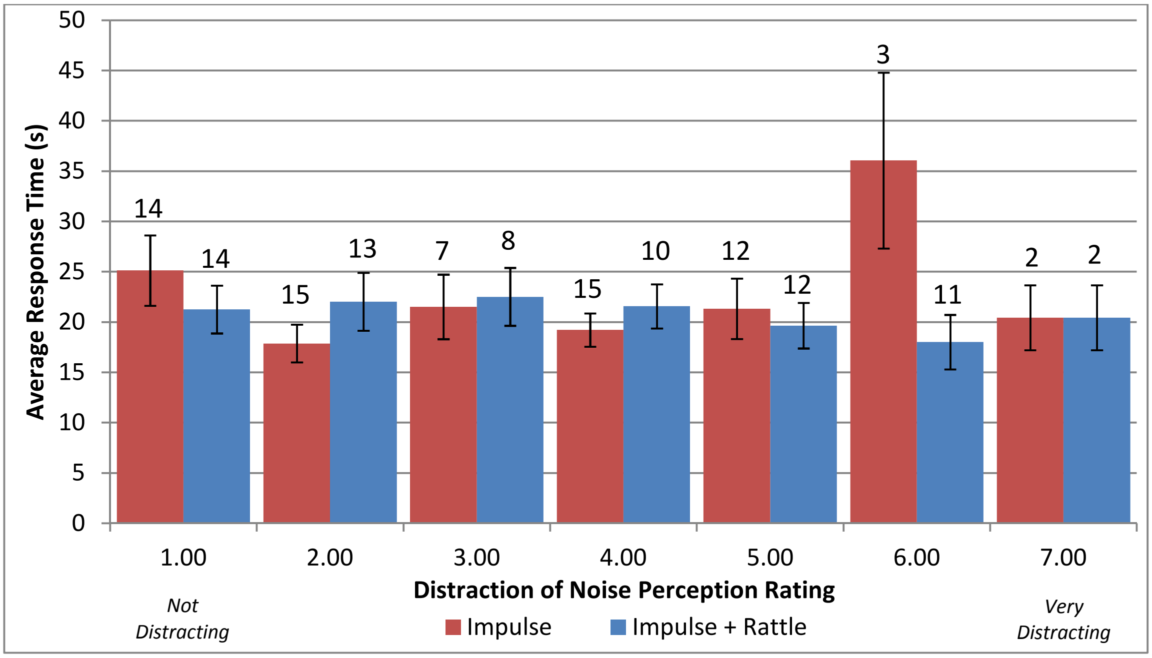


Fig. 6.23. Results of the perception ratings of distraction of noise and the corresponding average time taken to solve task problems in a given session. Results are split between impulse sessions and impulse plus rattle sessions. Error bars represent the standard error of the mean. Numbers represent the sample size for each response.

6.4. Subjective Perception Results

6.4.1. Subjective Perception Results across Noise Conditions

Subjects rated their perception of the loudness of noise, changes in time of noise, rumble of noise, annoyance to noise, and distraction of noise, as detailed in Section 5.2.3.2, for each session. These ratings were then compared to the eight noise conditions previously mentioned. A Kolmogorov-Smirnov test was run and showed that all results exhibited a non-normal distribution. Therefore, the Friedman ANOVA test was used to analyze these relationships, and Wilcoxon tests were used to further analyze the relationships between each noise condition.

Although it is a parametric test, repeated measures ANOVA results are also reported to back up the non-parametric test results. Observed power, at $\alpha = 0.05$, is reported for each. Further results of the repeated measures ANOVA from SPSS,

including sum of squares, mean square, and degrees of freedom are reported in Section 5.4.3.

6.4.1.1. Loudness of Noise across Noise Conditions

The loudness of noise ratings were significantly affected by the different noise conditions ($\chi^2(7) = 15.77$, $p < 0.05$). As the level of impulse plus rattle increased, the ratings of loudness of increased. A change in impulse level did not affect the ratings of loudness of noise for the impulse alone sessions, but a positively correlated trend can be seen in the impulse-plus-rattle-presented conditions. The average perception ratings for loudness of noise in each noise condition are shown in Figure 6.24. The results of the Wilcoxon test are shown in Table 6.4.

Table 6.4. Wilcoxon Results between Noise Condition and Loudness of Noise. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.05 level of significance.

Loudness ratings between sessions								
		Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55	T	14.50	30.00	27.50	27.50	27.00	11.50	16.00
	sig	ns	ns	ns	ns	ns	ns	**
	effect size	-0.09	-0.05	-0.09	-0.22	-0.01	-0.23	-0.40
Impulse 60	T		22.00	29.50	22.50	22.00	14.50	16.50
	sig		ns	ns	ns	ns	ns	**
	effect size		-0.01	-0.13	-0.17	-0.10	-0.30	-0.39
Impulse 65	T			43.00	19.00	37.50	10.00	16.50
	sig			ns	ns	ns	ns	**
	effect size			-0.10	-0.15	-0.02	-0.33	-0.44
Impulse 70	T				28.00	36.00	24.50	23.00
	sig				ns	ns	ns	ns
	effect size				-0.22	-0.04	-0.05	-0.27
Impulse 55 + Rattle 59	T					20.50	11.00	12.00
	sig					ns	**	**
	effect size					-0.20	-0.44	-0.44
Impulse 60 + Rattle 64	T						11.50	18.00
	sig						ns	**
	effect size						-0.23	-0.34
Impulse 65 + Rattle 69	T							22.00
	sig							ns
	effect size							-0.29

** = Indicates the mean ranks between noise conditions is significant, p<0.05, ns = not significant.

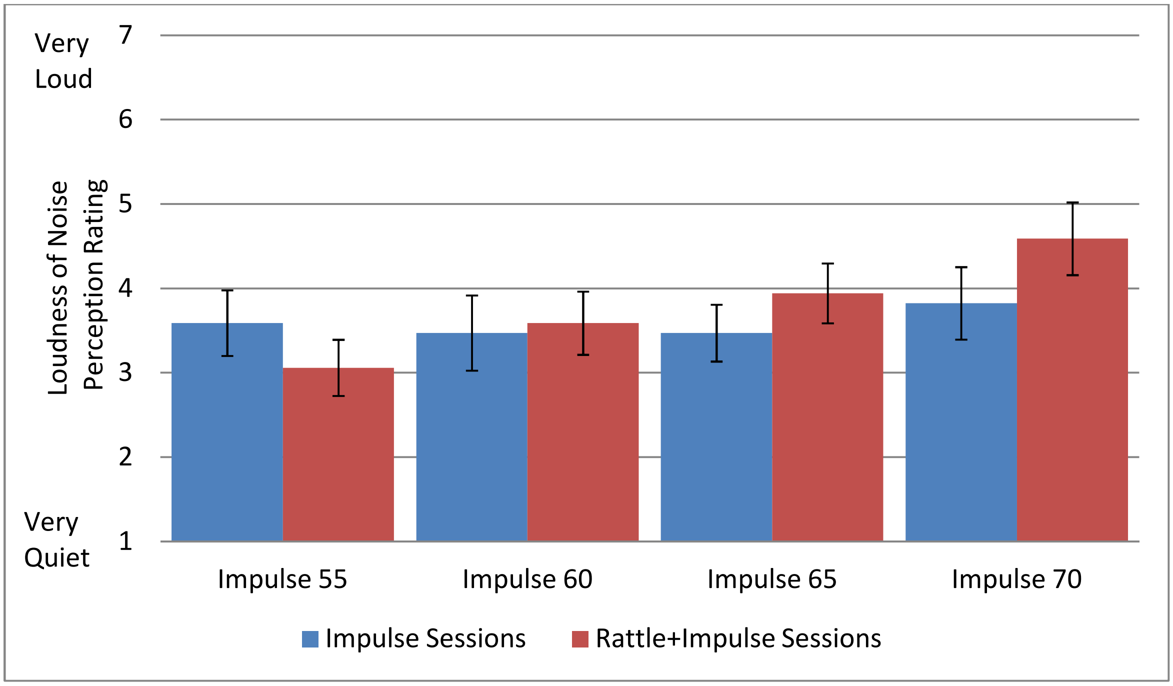


Fig. 6.24. Results of the average perception ratings of loudness of noise across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was not violated ($\chi^2(27) = 39.97$, $p<0.05$). Assuming sphericity, the results displayed a significant effect with a small effect size ($F(7,112) = 2.67$, $p<0.05$, $r = 0.20$). The observed power for this test was 0.886, or 88.6%, according to the results reported by SPSS with $\alpha = 0.05$. This suggests that the probability of these results presenting a genuine effect is great.

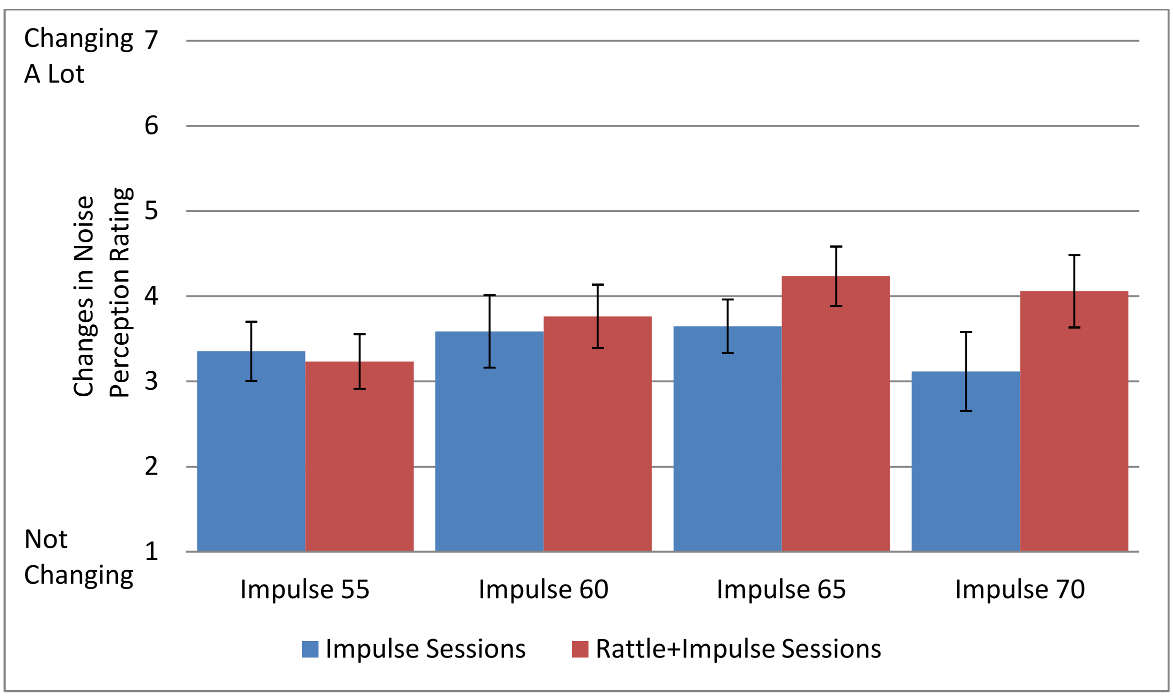
6.4.1.2. Change in Noise over Time across Noise Conditions

The changes in noise over time ratings were significantly affected by the different noise conditions ($\chi^2(7) = 17.31$, $p < 0.05$). In general, the perceived changes in noise ratings increase as the level of impulse or impulse plus rattle increased. The average perception ratings for changes in noise over time in each noise condition are shown in Figure 6.25. The results of the Wilcoxon test are shown in Table 6.5.

<u>Table 6.5.</u> Wilcoxon Results between Noise Condition and Change of Noise over Time. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.05 level of significance.

		Changes in noise ratings between sessions						
		Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55	T	*26.50*	*24.00*	*59.00*	*45.50*	*36.50*	*11.00*	*38.00*
	sig	*ns*	*ns*	*ns*	*ns*	*ns*	**	*ns*
	effect size	*-0.10*	*-0.14*	*-0.08*	*-0.08*	*-0.17*	*-0.34*	*-0.27*
Impulse 60	T		*28.00*	*20.50*	*20.00*	*24.00*	*30.00*	*7.00*
	sig		*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
	effect size		*-0.08*	*-0.20*	*-0.14*	*-0.14*	*-0.25*	*-0.33*
Impulse 65	T			16.00	12.00	28.00	4.00	34.50
	sig			ns	ns	ns	**	ns
	effect size			-0.32	-0.28	-0.08	-0.40	-0.13
Impulse 70	T				28.00	9.00	3.00	9.00
	sig				ns	**	**	**
	effect size				-0.08	-0.38	-0.49	-0.37
Impulse 55 + Rattle 59	T					11.50	8.00	10.50
	sig					ns	**	**
	effect size					-0.29	-0.43	-0.35
Impulse 60 + Rattle 64	T						19.50	33.00
	sig						ns	ns
	effect size						-0.27	-0.08
Impulse 65 + Rattle 69	T							34.00
	sig							ns
	effect size							-0.07

** = Indicates the mean ranks between noise conditions is significant, p<0.05, ns = not significant.



<u>Fig. 6.25.</u> Results of the average perception ratings of changes in noise over time across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was violated ($\chi^2(27) = 42.676$, $p<0.05$). Greenhouse-Geisser estimates of sphericity were used to correct the degrees of freedom and still displayed a significant effect with a small effect size ($F(3.62, 57.95) = 2.380$, $p<0.001$, $r = 0.19$). The observed power for this test was 0.621, or 62.1%, according to the results reported by SPSS with $\alpha = 0.05$.

Bonferroni post hoc tests found one significant relationship as shown in Table 6.6. This significant relationship matches one of those found in the Wilcoxon test. This significant difference is between the impulse-70-presented condition and the impulse-65-plus-rattle-69-presented condition, and does not really make sense. Since the data had a non-normal distribution, the non-parametric test results are more likely to represent the accurate effects. Therefore the other significant relationships in the Wilcoxon test results are presented with caution.

Table 6.6. Bonferroni Post Hoc Tests for Changes in Noise over Time Ratings across Noise Conditions.

Changes in noise ratings between sessions							
	Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55							
Impulse 60							
Impulse 65							
Impulse 70						**	
Impulse 55 + Rattle 59							
Impulse 60 + Rattle 64							
Impulse 65 + Rattle 69							

**Indicates the mean difference between noise conditions is significant, p<0.05

6.4.1.3. Rumble of Noise across Noise Conditions

The rumble of noise ratings were significantly affected by the different noise conditions ($\chi^2(7) = 28.95$, $p < 0.05$). In general, the perceived rumble of noise ratings increase as the level of impulse or impulse-plus-rattle increased, and the impulse-plus-rattle conditions were rated higher than the impulse conditions. The average perception ratings for rumble of noise in each noise condition are shown in Figure 6.26. The results of the Wilcoxon test are shown in Table 6.7.

Table 6.7. Wilcoxon Results between Noise Condition and Rumble of Noise. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.05 level of significance.

Rumble noise ratings between sessions								
		Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55	T	9.00	6.50	25.00	18.00	17.50	17.00	3.00
	sig	ns	ns	ns	ns	ns	**	**
	effect size	-0.15	-0.28	-0.31	0.00	-0.24	-0.39	-0.44
Impulse 60	T		18.50	10.00	14.00	12.00	13.00	7.00
	sig		ns	ns	ns	ns	**	**
	effect size		-0.16	-0.32	-0.10	-0.22	-0.40	-0.41
Impulse 65	T			33.50	0.00	22.50	12.00	12.00
	sig			ns	**	ns	ns	ns
	effect size			-0.08	-0.39	0.00	-0.33	-0.33
Impulse 70	T				11.00	46.00	7.00	28.00
	sig				ns	ns	ns	ns
	effect size				-0.40	-0.08	-0.28	-0.22
Impulse 55 + Rattle 59	T					7.00	0.00	4.50
	sig					ns	**	**
	effect size					-0.33	-0.52	-0.51
Impulse 60 + Rattle 64	T						12.00	11.50
	sig						**	ns
	effect size						-0.40	-0.29
Impulse 65 + Rattle 69	T							33.00
	sig							ns
	effect size							0.00

** = Indicates the mean ranks between noise conditions is significant, p<0.05, ns = not significant.

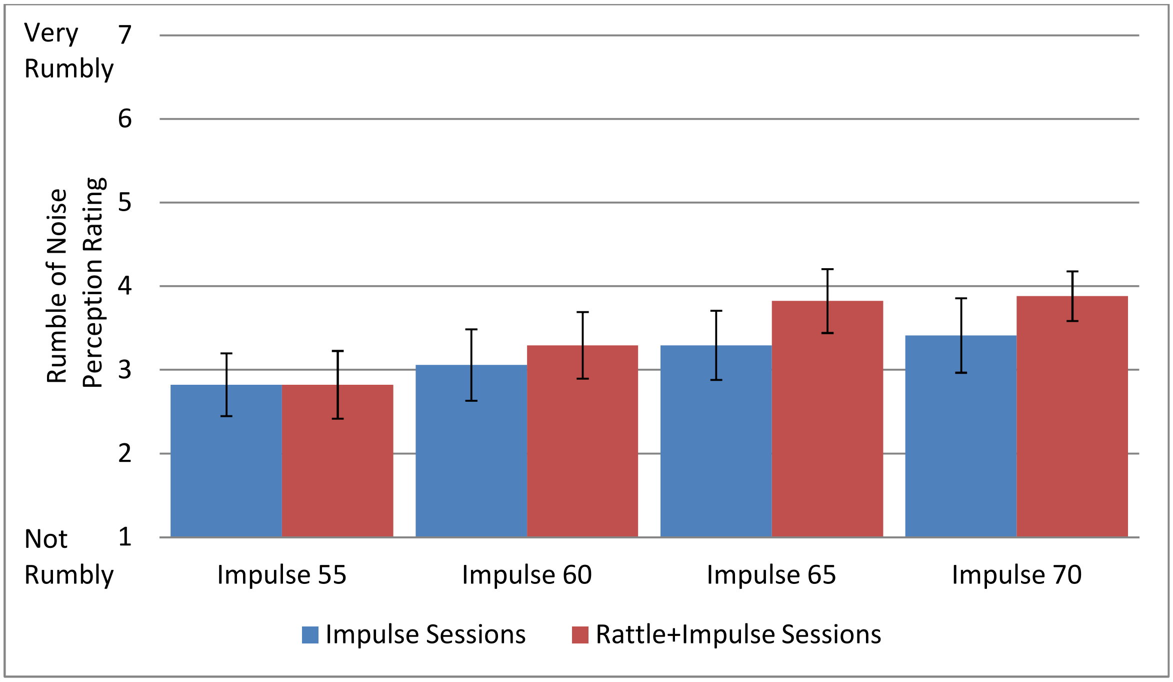


Fig. 6.26. Results of the average perception ratings of rumble of noise across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was violated so Greenhouse-Geisser estimates of sphericity were used to correct the degrees of freedom ($F(3.491, 55.862) = 3.706$, $p<0.05$, $r = 0.23$). The observed power for this test was 0.820, or 82%, according to the results reported by SPSS with $\alpha = 0.05$. This suggests the probability of the rumble perception vs. noise condition results presenting a genuine effect is good.

The Wilcoxon results found significant relationships between the loudest two impulse-plus-rattle conditions and the quietest two impulse-only conditions and quietest impulse-plus-rattle condition for the perceived rumble of noise. A significant relationship was also found between the second loudest impulse-plus-rattle and quietest impulse-only conditions. Bonferroni post hoc tests found significant relationships as shown in Table 6.8 matching two of those found in the Wilcoxon test.

Table 6.8. Bonferroni Post Hoc Tests for Rumble of Noise Ratings across Noise Conditions.

Rumble of noise ratings between sessions

	Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55							
Impulse 60							
Impulse 65							
Impulse 70							
Impulse 55 + Rattle 59						**	**
Impulse 60 + Rattle 64							
Impulse 65 + Rattle 69							

**Indicates the mean difference between noise conditions is significant, p<0.05

Test results show that rumble ratings were significantly higher for the two loudest impulse plus rattle sessions compared to the quietest impulse plus rattle session.

6.4.1.4. Annoyance to Noise across Noise Conditions

The changes in noise over time ratings were significantly affected by the different noise conditions ($\chi^2(7) = 18.00$, $p < 0.05$). In general, the perceived annoyance ratings increase as the level of impulse or impulse plus rattle increased. The average perception ratings for annoyance to noise in each noise condition are shown in Figure 6.27. The results of the Wilcoxon test are shown in Table 6.9.

Table 6.9. Wilcoxon Results between Noise Condition and Annoyance to Noise. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.05 level of significance.

Annoyance of noise ratings between sessions								
		Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55	T	*19.50*	*20.50*	*31.00*	*16.00*	*32.50*	*10.50*	*5.00*
	sig	*ns*	*ns*	*ns*	*ns*	*ns*	**	**
	effect size	*-0.06*	*-0.04*	*-0.11*	*-0.13*	*-0.09*	*-0.35*	*-0.36*
Impulse 60	T		*17.00*	*14.50*	*16.50*	*27.00*	*8.00*	*7.00*
	sig		*ns*	*ns*	*ns*	*ns*	**	**
	effect size		*-0.02*	*-0.17*	*-0.20*	*-0.09*	*-0.35*	*-0.40*
Impulse 65	T			12.00	12.00	39.00	17.50	8.00
	sig			ns	ns	ns	ns	**
	effect size			-0.22	-0.15	-0.08	-0.29	-0.43
Impulse 70	T				25.50	34.50	28.50	13.00
	sig				ns	ns	ns	ns
	effect size				-0.25	-0.06	-0.14	-0.26
Impulse 55 + Rattle 59	T					16.00	6.00	6.50
	sig					ns	**	**
	effect size					-0.26	-0.45	-0.44
Impulse 60 + Rattle 64	T						15.00	14.50
	sig						ns	ns
	effect size						-0.29	-0.29
Impulse 65 + Rattle 69	T							29.00
	sig							ns
	effect size							-0.14

** = Indicates the mean ranks between noise conditions is significant, p<0.05, ns = not significant.

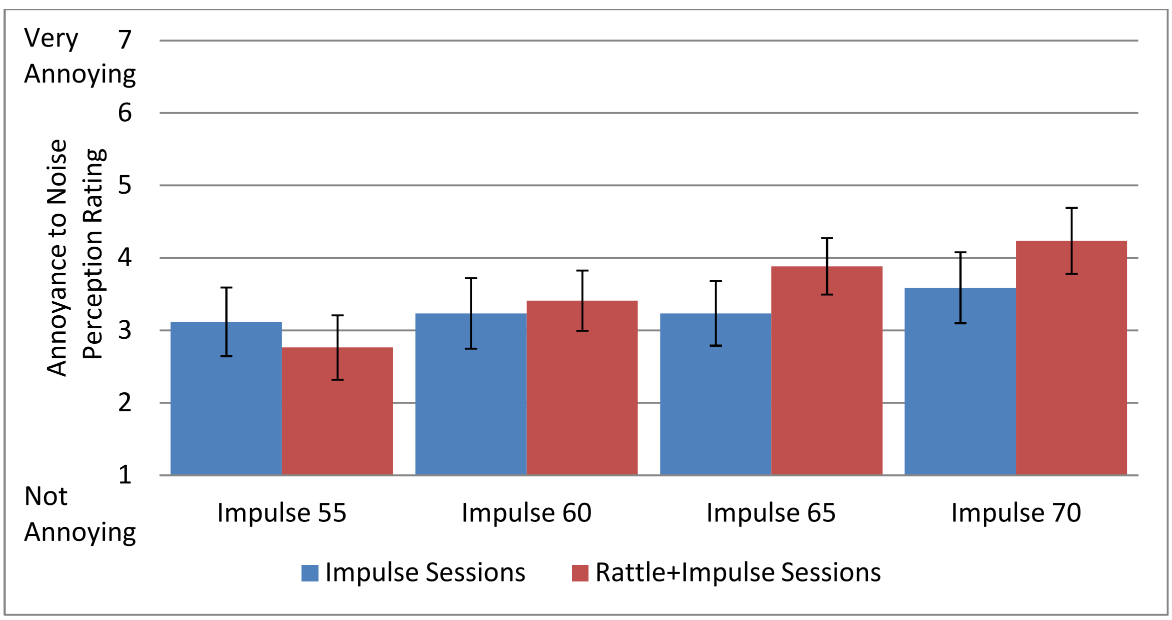


Fig. 6.27. Relationship between Noise Conditions and Annoyance to Noise Ratings. Results of the average perception ratings of annoyance to noise across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was violated so Greenhouse-Geisser so estimates of sphericity were used to correct the degrees of freedom ($F(4.214, 67.417) = 2.706$, $p<0.001$, $r = 0.20$). An effect size of 0.20 means that 20% of the change in the annoyance rating can be accounted for by the noise condition. The observed power for this test was 0.736, or 73.6%, according to the results reported by SPSS with $\alpha = 0.05$. This suggests that the probability that the annoyance perception vs. noise conditions results are exhibiting a genuine effect is very good.

The Wilcoxon results show the two loudest impulse-plus-rattle conditions rated significantly more annoying than the two quietest impulse-only and quietest impulse-plus rattle conditions. The loudest impulse-plus-rattle condition was also found to be significantly more annoying than the impulse 65 condition.

6.4.1.5. Distraction to Noise across Noise Conditions

The distraction of the noise ratings were significantly affected by the different noise conditions, $\chi^2(7) = 19.34$, $p < 0.05$. In general, the perceived distraction of the noise ratings increase as the level of impulse or impulse plus rattle increased. The average perception ratings for changes in noise over time in each noise condition are shown in Figure 6.28. The results of the Wilcoxon test are shown in Table 6.10.

Table 6.10. Wilcoxon Results between Noise Condition and Distraction to Noise. A Bonferroni correction was applied and all effects denoted with ** are significant at a 0.05 level of significance.

Distraction of noise ratings between sessions								
		Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55	T	*16.00*	*35.00*	*12.00*	*20.50*	*27.50*	*15.50*	*0.00*
	sig	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	**
	effect size	*-0.05*	*-0.05*	*-0.15*	*-0.13*	*-0.08*	*-0.21*	*-0.51*
Impulse 60	T		*33.50*	*13.50*	*22.50*	*35.50*	*25.50*	*0.00*
	sig		*ns*	*ns*	*ns*	*ns*	*ns*	**
	effect size		*-0.08*	*-0.11*	*-0.17*	*-0.05*	*-0.19*	*-0.59*
Impulse 65	T			20.00	25.50	31.50	16.00	14.50
	sig			ns	ns	ns	ns	**
	effect size			-0.14	-0.19	-0.02	-0.13	-0.38
Impulse 70	T				30.00	11.00	16.50	10.50
	sig				ns	ns	ns	**
	effect size				-0.19	-0.09	-0.04	-0.35
Impulse 55 + Rattle 59	T					28.00	22.50	3.50
	sig					ns	ns	**
	effect size					-0.15	-0.33	-0.51
Impulse 60 + Rattle 64	T						7.00	5.00
	sig						ns	ns
	effect size						-0.21	-0.31
Impulse 65 + Rattle 69	T							4.00
	sig							**
	effect size							-0.38

** = Indicates the mean ranks between noise conditions is significant, $p<0.05$, ns = not significant.

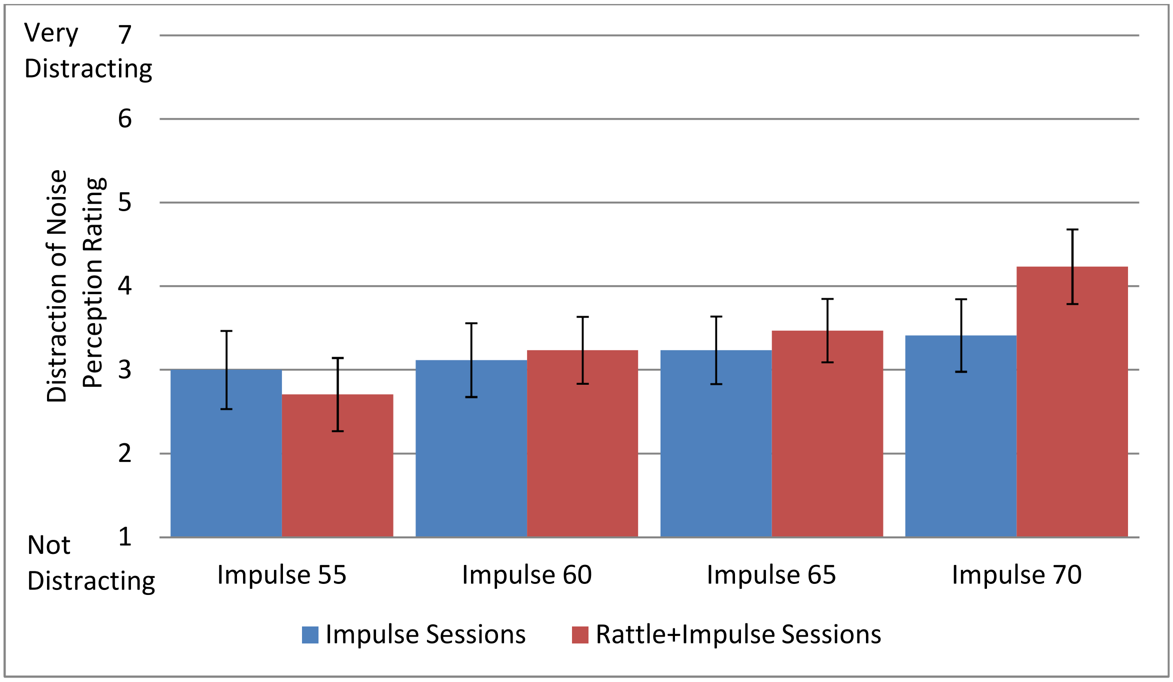


Fig. 6.28. Relationship between Noise Conditions and Distraction to Noise Ratings. Results of the average perception ratings of distraction to noise across each noise condition. Error bars represent the standard error of the mean.

A repeated measures ANOVA was additionally used for comparison. For ANOVA testing, Mauchly's sphericity was violated so Greenhouse-Geisser so estimates of sphericity were used to correct the degrees of freedom ($F(3.919, 2.709) = 2.852$, $p<0.001$, $r = 0.20$). An effect size of 0.20 means that 20% of the change in the distraction rating can be accounted for by the noise condition. The observed power for this test was .739, or 73.9%, according to the results reported by SPSS with $\alpha = 0.05$. This suggests that the probability that the distraction perception vs. noise condition results are exhibiting a genuine effect is very good.

The Wilcoxon results found the loudest impulse-plus-rattle condition to be rated significantly more distracting than every other condition except for the impulse 60 plus rattle 64 conditions.

Bonferroni post hoc tests found significant relationships as shown in Table 6.11. These significant relationships matched three of those found in the Wilcoxon test. Since the data had a non-normal distribution, the non-parametric test results are more likely to represent the accurate effects. Therefore the other significant relationships in the Bonferroni post hoc test results are presented with caution.

Table 6.11. Bonferroni Post Hoc Tests for Distraction to Noise Ratings across Noise Conditions.

Distraction of noise ratings between sessions

	Impulse 60	Impulse 65	Impulse 70	Impulse 55 + Rattle 59	Impulse 60 + Rattle 64	Impulse 65 + Rattle 69	Impulse 70 + Rattle 74
Impulse 55							**
Impulse 60							**
Impulse 65							
Impulse 70							
Impulse 55 + Rattle 59							**
Impulse 60 + Rattle 64							
Impulse 65 + Rattle 69							

**Indicates the mean difference between noise conditions is significant, $p<0.05$

Test results show that distraction ratings were significantly higher for the loudest impulse plus rattle session compared to the quietest impulse plus rattle session and the two quietest impulse only sessions.

6.4.2. Relationships of Subjective Perception Results across Noise Conditions with Gender, Age, and Noise Sensitivity as Covariates

There were very little significant relationships found between subjective perception results and noise conditions as shown in Section 6.4.1. However, additional variables were not factored into these results. Gender, age, and noise sensitivity were other independent variables collected during this study. It is necessary to look at the

effects of these variables on the relationship between subjective perception ratings to the four noise conditions to get a better understanding of everything affecting the results.

These additional independent variables are difficult to include in a non-parametric test like the ones used in this study. However, repeated measures ANOVA with covariates can still be analyzed to study these relationships with multiple independent variables. Because of the non-normal distributions of the subjective perception ratings, these results are presented with caution. Additionally, the observed power, as reported by SPSS with $\alpha = 0.05$, is reported for each repeated measures ANOVA test.

An analysis on the F-test ANOVA (repeated measures, within factors a priori) was conducted utilizing GPower 3.1 to determine a desired sample size to obtain a power of 0.80 for all 7 tested metrics; the results are shown in Table 6.12. It is estimated that a sample size of 112 subjects is necessary for the effect size, calculated with SPSS, to have a desired power of 0.80 for all tested metrics. In other words, it is necessary to test 112 subjects to be sure that these results, no statistically significant correlation between noise condition and task performance, hold true for the general population. It should be noted that a power of 0.80 was achieved for both the loudness and rumble subjective ratings, while a power of 0.73 was achieved for both the annoyance and distraction ratings. This means that the sample size used was sufficient in achieving near the desired 0.80 power for the perception metrics.

Table 6.12. Analysis of desired sample size to obtain a desired power of 0.80 for each tested variable.

Desired Power = 0.80	Effect Size, r	Desired Sample Size
Performance		
Percent Correct	0.09	112
Average Response Time	0.11	76
Perception		
Loudness	0.20	24
Changes	0.19	26
Rumble	0.23	18
Annoyance	0.20	24
Distraction	0.20	24

The SPSS outputs for each subjective perception rating with each covariate combination are shown in Tables 6.13 to 6.17. Almost all relationships remain insignificant, $p < 0.05$, with the addition of covariates. The relationship between noise condition and perceived changes in noise, which is significant without covariates, becomes insignificant with the addition of any or all analyzed covariates. However, there is one interesting outlier to note: the addition of age and noise sensitivity makes the relationship between perceived loudness of the noise and the noise condition significant.

Observed power depends on the number of independent variables and the sample size. In all cases, except for the significant relationship between perceived changes and the noise condition, the observed power is below the recommended 0.8. A larger sample size is desired to increase the observed power – increasing the probability that these tests show a genuine effect (Field and Hole 2003). Therefore, these results are presented with caution.

Table 6.13. Analysis of variance for loudness to noise ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity except when stated otherwise.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Loudness of Noise Ratings across Noise Conditions	23.853	7.000	3.408	2.671	0.014	0.143	18.695	0.886
Gender		11.739	7.000	1.677	1.436	0.199	0.087	10.054	0.583
Age		11.739	7.000	1.677	1.436	0.199	0.087	10.054	0.583
Noise Sensitivity		17.350	7.000	2.479	2.042	0.056	0.120	14.297	0.766
Gender and Noise Sensitivity		11.838	7.000	1.691	1.521	0.169	0.098	10.647	0.610
Gender and Age		2.986	7.000	0.427	0.371	0.917	0.026	2.594	0.160
Age and Noise Sensitivity		17.216	7.000	2.459	2.139	0.046	0.133	14.971	0.786
Gender, Age, and Noise Sensitivity		3.130	7.000	0.447	0.418	0.889	0.031	2.923	0.176

[a]Computed using alpha = .05

Table 6.14. Analysis of variance for change in noise over time ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity except when stated otherwise.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Changes in Noise over Time Ratings across Noise Conditions	18.110	3.622[b]	5.001	2.380	0.068	0.129	8.618	0.621
Gender		3.101	3.500[b]	0.886	0.394	0.787	0.026	1.380	0.128
Age		6.072	3.506[b]	1.732	0.787	0.524	0.050	2.758	0.222
Noise Sensitivity		3.694	3.613[b]	1.023	0.474	0.736	0.031	1.712	0.148
Gender and Noise Sensitivity		4.809	7.000	0.687	0.600	0.755	0.041	4.199	0.247
Gender and Age		3.984	3.415[b]	1.167	0.497	0.710	0.034	1.698	0.150
Age and Noise Sensitivity		6.055	3.404[b]	1.779	0.759	0.538	0.051	2.582	0.211
Gender, Age, and Noise Sensitivity		4.075	3.290[b]	1.239	0.491	0.707	0.036	1.615	0.145

[a]Computed using alpha = .05, [b]Greenhouse-Geisser correction used

Table 6.15. Analysis of variance for rumble of noise ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Rumble of Noise Ratings across Noise Conditions	19.346	3.491[b]	5.541	3.706	0.013	0.188	12.938	0.820
Gender		9.049	3.432[b]	2.637	1.762	0.159	0.105	6.045	0.463
Age		2.903	3.513[b]	0.822	0.549	0.679	0.035	1.940	0.165
Noise Sensitivity		3.737	3.372[b]	1.108	0.700	0.572	0.045	2.361	0.197
Gender and Noise Sensitivity		7.991	3.306[b]	2.417	1.537	0.214	0.099	5.082	0.398
Gender and Age		1.204	3.430[b]	0.351	0.227	0.899	0.016	0.778	0.092
Age and Noise Sensitivity		2.862	7.000	0.409	0.546	0.797	0.038	3.823	0.226
Gender, Age, and Noise Sensitivity		1.281	3.435[b]	0.373	0.244	0.888	0.018	0.839	0.095

[a]Computed using alpha = .05, [b]Greenhouse-Geisser correction used

Table 6.16. Analysis of variance for annoyance to noise ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Annoyance to Noise Ratings across Noise Conditions	25.404	4.214[b]	6.029	2.706	0.035	0.145	11.401	0.736
Gender		12.401	7.000	1.772	1.413	0.208	0.086	9.892	0.574
Age		18.714	7.000	2.673	2.070	0.053	0.121	14.493	0.773
Noise Sensitivity		4.780	4.138[b]	1.155	0.490	0.749	0.032	2.028	0.161
Gender and Noise Sensitivity		10.892	7.000	1.556	1.204	0.308	0.079	8.430	0.493
Gender and Age		2.542	7.000	0.363	0.288	0.957	0.020	2.019	0.132
Age and Noise Sensitivity		18.482	7.000	2.640	2.071	0.054	0.129	14.495	0.771
Gender, Age, and Noise Sensitivity		2.487	7.000	0.355	0.288	0.957	0.022	2.018	0.131

[a]Computed using alpha = .05, [b]Greenhouse-Geisser correction used

Table 6.17. Analysis of variance for distraction to noise ratings across noise conditions with each combination of gender, age, and noise sensitivity as covariates. All results assume for sphericity.

Covariates	Within-Subjects Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
none	Distraction of Noise Ratings across Noise Conditions	23.816	3.919[b]	6.077	2.858	0.031	0.152	11.203	0.739
Gender		7.922	3.786[b]	2.092	1.000	0.412	0.062	3.786	0.288
Age		13.988	3.805[b]	3.676	1.713	0.162	0.102	6.518	0.480
Noise Sensitivity		4.014	3.844[b]	1.044	0.468	0.752	0.030	1.797	0.150
Gender and Noise Sensitivity		12.913	3.569[b]	3.618	1.619	0.189	0.104	5.779	0.437
Gender and Age		5.267	3.613[b]	1.458	0.665	0.604	0.045	2.402	0.194
Age and Noise Sensitivity		13.760	3.617[b]	3.804	1.762	0.157	0.112	6.373	0.475
Gender, Age, and Noise Sensitivity		5.251	7.000	0.750	0.703	0.669	0.051	4.921	0.287

[a]Computed using alpha = .05, [b]Greenhouse-Geisser correction used

Chapter 7: Conclusions

This study, in two phases of testing, examined the effect of varying the time intervals of fluctuation between two background noise levels and the effect of varying levels of short broadband noise bursts both with and without a rattle element. The first phase of tests utilized two levels of background noise, one matching a room criteria rating of RC-29(H) and the other RC-47(RV). 27 subjects were exposed to these noise levels with the exposure time interval varied across four sessions with intervals of two minutes, five minutes, eight minutes, and ten minutes.

The second phase of tests utilized bursts of noise with peak A-weighted sound pressure levels (L_{Apk}) ranging from 55 to 75 dBA presented both with and without a rattle element from a separate source with L_{Apk} measurements ranging from 59 to 74 dBA. These bursts and bursts plus rattle were presented over a generated ambient background noise matching RC-29(H). 17 subjects were exposed to a number of these burst stimuli in a randomized yet controlled fashion across eight sessions.

For all sessions, task performance was measured by the total percentage of correct answers and average response time on an arithmetic task involving memorization, while subjective perception was measured from the results of a subjective questionnaire. Statistical analyses were applied to the results to further analyze the relationships.

Results for phase one found a significant relationships, $p<0.05$, between each noise condition and task performance in the form of total percentage of correct responses from both parametric and non-parametric statistical analyses. Few significant relationships were found in relation to subjective perception, except between the noise condition and the subjective perception of changes in noise.

Results for phase two found no statistically significant relationships between each noise condition and task performance, for both noise bursts alone and noise bursts plus rattle, from both parametric and non-parametric statistical analyses. A few significant relationships were found between subjective perception and performance, though, as well as between subjective perception and noise condition.

7.1. Phase One

7.1.1. Task Performance

Teichner et al. found a decrease in performance for sessions involving a single fluctuation to either an elevated or quieter level, with the decrease in performance being more pronounced the greater the difference between the original and changed levels (1963). Moorhouse found that an acceptable level for noise fluctuating every 100 milliseconds was on average 5 dB higher than a non-fluctuating of similar spectral content (2007). However, neither of these studies tested the effect of varying the rates of fluctuation on a longer time scale on performance and perception.

For performance across all test questions, a statistically significant relationship, $p<0.05$, was found between the noise conditions and the percentage of correct responses, dropping from 87% correct to 80% correct as the interval decreased from ten minutes to two minutes. This suggests that performance accuracy on an arithmetic task involving memory decreases as the interval of fluctuating noise decreases. However, response time was statistically unaffected by fluctuation rate.

We find it likely that a cut-off time interval for fluctuating noise levels becoming unacceptable is less than two minutes. However, it is unlikely that HVAC systems actually fluctuate on intervals less than two minutes. Questions were not linked to a

specific time in the noise condition. Therefore, it is not possible to extract whether the exact times when the level fluctuated had different performance results compared to the rest of the session.

7.1.2. Subjective Perception

Subjective perception results were collected from subjective questionnaire ratings on various qualities of noise: loudness of noise, changes in noise over time, rumble of noise, annoyance to noise, and distraction to noise. The only significant relationship, $p<0.05$, between subjective perception metrics and noise conditions was found in the subjective rating of changes of noise. This suggests that the subjects could distinguish between the different rates of fluctuation; however, they did not perceive any specific noise conditions as being louder, more annoying, or more distracting. Additionally, this significant relationship was found in both the parametric and non-parametric tests, so the relationships found have a good chance of being genuine effects. No significant relationships were found when accounting for additional independent variables: gender, age, and noise sensitivity.

Many subjects reported that the exact moments of time at which the noise fluctuated from one level to the next were the most distracting parts of the overall session and that they adapted pretty quickly to the new steady state level. However, a few noted that it was harder to concentrate in the elevated background noise.

Wang and Novak found that a recording of a heat pump fluctuating between two levels every thirty seconds was rated statistically significantly more distracting and annoying than other noise stimuli (2010). From perception results, we expect that a cut-

off of unacceptable time interval for fluctuating noise levels is less than two minutes, but it is unlikely that HVAC systems actually fluctuate on intervals less than two minutes.

7.2. Phase Two

7.2.1. Task Performance

No statistically significant relationships between overall performance and noise condition were found. This suggests that bursts of noise and bursts of noise with this specific rattle element had no significant effect on performance of an arithmetic task with a memory component.

No statistically significant relationships were found between performance on impulse presented questions, the same questions presented with impulse and rattle, and the same questions presented with no impulse. This suggests that noise bursts and noise bursts presented with the specific rattle stimuli utilized in this test had no significant effect on performance on the questions in which they were presented of the arithmetic task with a memory component. Ainley also found no statistically significant relationships between overall task performance in a session and the level of noise bursts with L_{Apk} ranging from 47 to 77 dBA (2012). He did, however, find a decreasing trend in percentage of correct answers in impulse-presented questions as the impulse level increased.

7.2.2. Subjective Perception

In 2011, Miller found that 83% of subject perceived booms experienced while indoors as more annoying than booms perceived while outdoors; many citing the presence of rattle as a cause of this. Also, Loubeau fond that annoyance to different types of rattles varied, even though they all had the same perceived level (PL) values (2013).

Rattles generated by larger objects were perceived as more annoying than rattles generated by smaller objects.

Subjective perception results were collected from subjective questionnaire ratings on various qualities of noise: loudness of noise, changes in noise over time, rumble of noise, annoyance to noise, and distraction to noise. The main statistically significant relationships, $p<0.05$, worth mentioning between subjective perception metrics and noise conditions were found in the subjective rating of distraction and annoyance of noise. Subjects rated the distraction of the noise burst at 70 dBA accompanied by rattle as statistically significantly more distracting than all other noise conditions except for the noise burst at 65 dBA plus rattle. Also, subjects rated the annoyance of the noise bursts at 65 and 70 dBA plus rattle as statistically significantly more annoying than the bursts at 55 and 60 dBA and the burst at 55 dBA plus rattle.

This suggests that, at least with filtered white noise bursts and this particular rattle recording, an impulse with L_{Apk} 70 dBA with accompanying rattle with L_{Apk} 74 dBA is too distracting and annoying for this type of arithmetic test. Additionally, these significant relationship were found in both the parametric and non-parametric tests. Because of this, the relationships found have a good chance of being genuine effects. No significant relationships were found when accounting for additional independent variables: gender, age, and noise sensitivity.

7.3. Future Research

In regards to phase one of this study, the difference between background noise level and the elevated level is pretty drastic compared to real-world fluctuations that would occur in the timeframe tested. Further research could test the effects on

performance and perception of even longer time scales of fluctuations. Further research could also be done to test fluctuating background noise conditions on a shorter time scale but with less drastic changes in level – maybe on the scale of 2-3 just noticeable differences (JNDs). Also, broadband noise was used for this study; noise with tonal components could also be tested.

In phase two, broadband noise bursts and a single recording of a wall hung mirror rattling (adjusted to different playback levels) were used for the testing. Results could be more pronounced if different burst signals (particularly sonic booms) as well as different rattle noises (especially generated by larger objects) were used. Also, this test was carried out in a controlled office-like setting; it would be interesting to see results of a similar test under real-world conditions as suggested by Thackray, Touchstone, and Bailey (1974).

For phase two, a sample size of 17 subjects, mostly college students, was used. Further research may consider utilizing a larger sample size, around 112 subjects, with a larger range of ages for better power for the performance results.

References

Broadbent, D. E. (1957). Effects of noises of high and low frequency on behaviour. *Ergonomics, 1*(1), 21-29. doi: 10.1080/00140135708964568

Broadbent, D. E. (1958). Effect of noise on an "intellectual" task. *The Journal of the Acoustical Society of America, 30*(9), 824. doi: 10.1121/1.1909779

Dittrich, K., & Oberfeld, D. (2009). A comparison of the temporal weighting of annoyance and loudness. *The Journal of the Acoustical Society of America, 126*(6), 3168. doi: 10.1121/1.3238233

Field, A., & Hole, G. (2003). *How to design and report experiments*. London: Sage Publications.

Frankenhaeuser, M., & Lundberg, U. (1977). The influence of cognitive set on performance and arousal under different noise loads. *Motivation and Emotion, 1*(2), 139-149. doi: 10.1007/BF00998516

Johnson, D. R., & Robinson, D. W. (1967). The subjective evaluation of sonic bangs. *Acustica, 18*(5), 241-258.

Lim, C., Kim, J., Hong, J., & Lee, S. (2008). Effect of background noise levels on community annoyance from aircraft noise. *The Journal of the Acoustical Society of America, 123*(2), 766-771.

Loubeau, A., Sullivan, B. M., Klos, J., Rathsam, J., & Gavin, J. R. (2013). *Laboratory Headphone Studies of Human Response to Low-Amplitude Sonic Booms and Rattle Heard Indoors.* (Tech. Rep. No. NASA/TM–2013-217975). NASA Langley Research Center.

Marshall, A., & Davies, P. (2010). A paired comparison experiment to examine startle evoked by low level sonic booms and other transients. *Proceedings of 20th International Congress on Acoustics, ICA 2010,* Sydney, Australia.

Marshall, A., & Davies, P. (2011). Metrics including time-varying loudness models to assess the impact of sonic booms and other transient sounds. *Noise Control Engineering Journal, 59*(6), 681-697.

Miller, D. M. (2011). *Human response to low-amplitude sonic booms.* (Unpublished Ph.D.). The Pennsylvania State University,

Moorhouse, A. T., Waddington, D. C., & Adams, M. D. (2007). The effect of fluctuations on the perception of low frequency sound. *Low Frequency Noise, Vibration and Active Control, 26*(2), 81-89.

Persson Waye, K., Bengtsson, J., Kjellberg, A., & Benton, S. (2001). Low frequency noise "pollution" interferes with performance. *Noise & Health, 4*(13), 33-49.

Rathsam, J., Loubeau, A., & Klos, J. (2012). *A study in a new test facility on indoor annoyance caused by sonic booms.* (Tech. Rep. No. NASA/TM–2012-217332). NASA Langley Research Center.

Rathsam, J., Loubeau A., Klos, J. (2013, August 26-28). *Simulator study of indoor annoyance caused by shaped sonic boom stimuli with and without rattle augmentation.* Paper presented at Noise-Con 2013, Denver, Colorado.

SanMiguel, I., Linden, D., & Escera, C. (2010). Attention capture by novel sounds: Distraction versus facilitation. *European Journal of Cognitive Psychology, 22*(4), 481-515.

Schutte, M., Marks, A., Wenning, E., & Griefahn, B. (2007). The development of the noise sensitivity questionnaire. *Noise & Health, 9*(34), 15-24.

Sullivan, B. M., Klos, J., Buehrle, R. D., McCurdy, D. A., & Haering Jr., E. A. (2010). *Human response to low-intensity sonic booms heard indoors and outdoors.* (Technical Memorandum No. NASA/TM-20).

Tafalla, R. J., & Evans, G. W. (1997). Noise, physiology, and human performance: The potential role of effort. *Journal of Occupational Health Psychology, 2*(2), 148-155. doi: 10.1037/1076-8998.2.2.148

Teichner, W. H., Arees, E., & Reilly, R. (1963). Noise and human performance, A psychophysiological approach. *Ergonomics, 6*(1), 83-97.

Thackray, R. I. (1972). Sonic boom exposure effects II.3:Startle responses. *Journal of Sound & Vibration, 20*, 519-526. doi: 10.1016/0022-460X(72)90675-X

Thackray, R. I., Touchstone, R. M., & Bailey, J. P. (1974). A comparison of the startle effects resulting from exposure to two levels of simulated sonic booms. *Journal of Sound & Vibration, 33*(4), 379-389.

Thackray, R. I., Touchstone, R. M., & Bailey, J. P. (1975). Reactions to sonic booms- A report of two studies and a general evaluation of startle effects. *Aviation, Space, and Environmental Medicine, 46*, 369-376.

Walker, E. L., & Tarte, R. D. (1963). Memory storage as a function of arousal and time with homogeneous and heterogeneous lists. *Journal of Verbal Learning and Verbal Behavior, 2*(1), 113-119.

Wang, L. M., & Novak, C. C. (2010). Human performance and perception-based evaluations of indoor noise criteria for rating mechanical system noise with time-varying fluctuations. *ASHRAE Transactions, 116*, 553-568.

Weiner, B., & Walker, E. L. (1966). Motivational factors in short-term retention. *Journal of Experimental Psychology, 71*(2), 190-193. doi: 10.1037/h0022848

Woodhead, M. M. (1958). Effects of bursts of loud noise on a continuous visual task. *British Journal of Industrial Medicine, 15*(2), 120-125.

Woodhead, M. M. (1959). Effect of brief loud noise on decision making. *The Journal of the Acoustical Society of America, 31*(10), 1329-1331.

Woodhead, M. M. (1964). The effect of bursts of noise on an arithmetic task. *The American Journal of Psychology, 77*(4), 627-633.

Zimmer, K., Ghani, J., & Ellermeier, W. (2008). The role of task interference and exposure duration in judging noise annoyance. *Journal of Sound and Vibration, 311*(3-5), 1039-1051.