Identifying The Number Of Instruments In Pairs Of Simultaneously Sounding Timbre

Dora Chan and Elzbieta Slawinski
Department of Psychology, University of Calgary, 2500 University Dr. NW., Calgary Alberta, T2N IN4

ABSTRACT

Timbre as a source of variation in music has become increasingly important in music composition in recent years. The present study investigated the ability of listeners to appreciate pairs of different timbres simultaneously. The stimuli were combinations of the steady state portions (300 ms) of three instrumental timbres, including the clarinet, trombone and harp. Forty different pairs were constructed for the four stimulus conditions: three experimental conditions consisting of the instrumental timbres, and a control condition consisting of pure tones corresponding to the fundamental frequency of these timbres. Thirty undergraduates, with and without musical training, were required to listen to a total of 400 randomized tonal stimuli. Their task was to make a judgment on whether they heard one or two instrument(s) after listening to each stimulus. Overall, musicians could perceive two timbres more readily than non-musicians, thus suggesting that musical experience may enhance the perception of timbre. Performance differences between the groups across conditions are also discussed.

RÉSUMÉ

Cette etude a examine la perception de deux instruments joues en meme temps, les deux instruments etant choisis d'entre le clarinet, le trombone, et le harp. Les stimuli etait contenus dans deux portions fixes (300 ms) des timbres produits par les instruments choisi. Quarante paires etaient construites pour quatre types des conditions: trois conditions experimentales de timbre instrumental, et une condition controle de pure ton correspondant a la frequance fondamentale de ce timbre. Une trentaine d'etudiants, avec et sans formation musicale, ont ecoute a un melange de quatre cents stimuli randomises. Les participants devaient identifier si le stimulus presente etait compose d'un ou de deux instruments de musique. Nous avons trouves que les participants avec une formation musicales etaient en general plus doues. Ceci peut constituer une indication que la formation musicale ameliore l'abilite de reconnaitre des timbres. Une discussion des differences dans le groupe experimental est presente.

1. INTRODUCTION

1.1 Definitions

The concept of timbre has been a difficult topic to study in the field of musical acoustics. One major complication begins with defining the term "timbre." Many definitions of timbre have concentrated on what timbre is not, rather than what timbre is (Risset & Wessel, 1999). This often leads to vague definitions of timbre. For example, a well-known definition from the American National Standards of Institute defines timbre as: "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar" (1960, p.45). This merely explains that any acoustic attribute that does not exclusively contribute to the perception of pitch, loudness, or duration could contribute to the perception of timbre. It does not, however, describe the physical parameters that contribute to the perception of timbre. Since the acoustic basis for timbre is undefined, timbre is regarded as a multidimensional attribute which cannot be measured on a single continuum (Plomp, 1976).

Presently, there is no truly satisfying definition of timbre as a psychophysical variable. However, timbre may still be communicated to a naïve listener. Timbre is easily understood as the perceptual attribute that enables us to distinguish musical tones produced by different instruments. Therefore timbre is often defined as the quality of sound, such as how "bright" or "dark" the tone sounds. But this vague definition does not take into account the complexities of the attributes of timbre and is not scientifically useful since it does not explicate a relationship between variables (Hajda, Kendall, Carterette & Harshberger, 1997). In order to study timbre scientifically and simply without using multidimensional scalings, researchers have continued to explore the time-varying aspects of the sound envelope.

Timbre is often described in terms of the transitions of a musical tone (or more specifically, the transitions of its harmonics, or partials) over time. This can be conceptualized overall as the time for the sound to grow to full amplitude and to decay to inaudibility (Handel, 1989). The changing portions of a tone can be described by the development of the amplitude envelope over time. In the case of an isolated musical tone, the envelope can be divided into three parts: the attack (onset), steady state (sustain) and decay (offset). In the early 1960s, transients became important variables in timbre studies. Many studies attempted to isolate the salience of the attack, steady state and decay by dividing and transforming isolated musical tones.

1.2 Early Studies of Timbre Attributes

Saldanha and Corso (1964) were one of the first groups of researchers to investigate the importance of transients in the identification of musical instruments. Specifically, they evaluated the relative importance of harmonic structure, frequencies of the equally tempered scale, vibrato, transient motion, both initial (attack) and final (decay), and steady-state duration as timbre cues in the absolute judgment of musical tones. They included 10 different instruments for timbre identification. Five types of stimuli resulted from the different divisions of the attack, steady state and decay of the tones.

The results of Saldanha and Corso's (1964) study indicated that there were great differences among the different instruments in their absolute identification using auditory cues. Identification was surprisingly poor for some instruments even without alteration, which suggests that some of the information listeners normally use to identify instruments is accumulated across several tones. The result of their study indicated that important information for instrument tone identification exists in the initial part of the sound event; the greatest decline in performance occurred when the attack was removed. With the absence of the attack, the steady state portion may still produce correct identification, however, accuracy is often much lower. Eliminating the decay portion did not seem to decrease recognition. Recognition of the timbre of musical tones with vibrato was less hindered by a missing attack portion than were musical tones that lacked vibrato.

The issue of partitioning isolated tones was re-introduced by Iverson and Krumhansl (1993), who used the technique of multidimensional scaling (MDS) to map participants' judgments on the perceived similarity of tones. Similarity scaling techniques are used to determine which acoustic attributes are most salient. MDS converts similarity judgments into a spatial map where distances in the space correspond to the perceived similarity.

Iverson and Krumhansl (1993) produced three multidimensional scalings of 16 digitally recorded musical instruments from three experiments. The purpose of their study was to examine the dynamic attributes of timbre by evaluating the role of onsets in similarity judgments. There were three experimental conditions, which involved: 1) unaltered signals (complete instrumental tones), 2) constant onset transients (first 80 ms), and 3) signals with onsets (first 80 ms) removed (known as "remainders").

The results of the study by Iverson and Krumhansl (1993) have revealed further complications regarding the dynamic attributes of timbre. They concluded that salient attributes for complete tones are present at the onset. However, ratings for complete tones also corresponded to those of remainders, indicating that the salient attributes for complete tones are present in the absence of onsets as well. This finding seemingly contrasted previous research (i.e., Saldanha & Corso, 1964), however, the researchers accounted for any comparable differences in the sounds and techniques used. They employed a similarity-scaling technique therefore it might be difficult to directly compare similarity judgments with identification judgments. Perhaps identification relies on onsets, but similarity judgments have no such reliance. The issue of dynamic attributes of timbre, in particular the role of onsets, is still unsettled. The current study used only the steady portion of all musical tones. In other words, the attack and decay portions were removed.

1.3 Simultaneous Instruments

Few studies have looked at the "blending" of timbre or the simultaneous perception of instruments. The first experimental investigations on the blending of concurrent timbres were carried out by Kendall and Carterette (1991, 1993). They conducted a series of studies examining the timbres of simultaneous orchestral wind instruments. They referred to pairings as "dyads." These dyads were constructed from all possible pairings of the following instruments: alto saxophone, oboe, flute, trumpet and clarinet. Different contexts of the dyads were used including: unisons (Bb4, approximately 466 Hz); unison melodies (D5, Eb5, F5 and D5, corresponding to 587, 622, 698, and 587 Hz played successively); major thirds (Bb4 and D5); and harmonized melodies (Bb4-D5, G4-Eb5, A4-F5, Bb4-D5). In the harmonized contexts, each instrument was used as the soprano. These different contexts were subjected to similarity scaling (1991) and identification of constituent instruments and ratings of blend (1993).

In the Kendall and Carterette studies mentioned above, it was found that the configuration of similarity scalings had two interpretable dimensions. These were identified as "nasal" versus "not nasal," "rich" versus "brilliant." A third dimension was interpreted as "simple" versus "complex." The extreme of the primary dimension were oboe (nasal) and clarinet (not nasal); of the second, trumpet (brilliant) and alto saxophone (rich). Overall, they found that "nasal" combinations blended less well than "brilliant" or "rich." Although not much work has been conducted in the perceptual blending of instrumental timbre, the work that has been done by

Kendall and Carterette has revealed important implications for orchestration and composition.

On the other hand, perception of auditory patterns, where a number of complex sound events occur simultaneously or sequentially, continues to challenge empirical investigation (Singh, 1987). The perceptual attributes of sounds are often transformed when presented in the dynamic context of a sequence. It has been pointed out by Singh that "[t]he process of sequencing, often augmented by repetition, allows similarities and differences between sounds to be discovered and used as criteria in organization and categorization" (p. 886). Melodies, or specifically, passages of orchestral music, have rarely been used in psychological studies, often due to the lack of control. However, in order to fully understand listeners' perception of timbre in an ecological sense, it is practical to look beyond the single musical tones and begin looking at complex auditory patterns. Although the current study did not examine complex-tone sequences, the results of a previous study that employed a combination of timbres and musical contexts created the impetus for the current study.

In Bonfield and Slawinski's (2002) study, participants were presented with brief passages from the first movement of Stravinsky's *Ebony Concerto* and they were required to identify all the orchestral instruments presented in a 1300 ms passage of interest. There were three conditions in which the targeted passage was presented. The first condition consisted of the passage in question. The second condition included 10 s of musical material leading into the targeted passage. The third condition provided musical material before and after the targeted stimulus. The three conditions were presented repeatedly in random order. Participants were provided with a broad list of ten instruments to select from although there were four instruments in the passage of interest, including a clarinet, trombone, harp and tom-tom.

Bonfield and Slawinski (2002) found that none of the participants could correctly name all four instruments. A more surprising finding was that all the participants (both musicians and non-musicians) named piano as one of the instruments presented in the targeted passage. Thus, the researchers speculated that a certain combination of the four instruments (perhaps two particular timbres) have created spectral qualities that were similar to the spectrum of the piano, thereby creating the perception of the piano timbre.

The present study aimed to follow up on the previous phenomena of timbre identification when more than one timbre is presented simultaneously. We wanted to examine whether the combination of two instrumental timbres would perceptually sound as a single timbre different from the other two timbres. More specifically, would listeners confuse the two combined timbres as one? Three of the four timbres presented in the musical passage used in Bonfield and Slawinski's (2002) study were also examined in this study. These included the clarinet, the trombone and the harp. Three experimental conditions were constructed from the pairings of the three timbres. In each condition one timbre was maintained at Eb5 while the other timbre began one octave apart and would progressively decrease in interval until both timbres merged at Eb5 or unison. Ten different intervals were included in the study. A control condition was also included which consisted of pure tone combinations at the fundamental frequencies. Since this study was concerned with spectral or static attributes, rather than temporal or dynamic attributes of timbre, only the steady-state portion of the musical signal was used.

Musicians and non-musicians participated in this study. They were required to make a judgment on whether they heard one or two instrument(s) after listening to each combination of timbres or pure tones. A major question of this study was whether or not listeners perceive the fusion of two instrumental timbres. Particularly, this study investigated whether musical training would improve the identification of the timbres; in other words, would there be a difference between those listeners who were musically trained (for eight years or more) and those who were musically untrained. It was hypothesized that musicians would perform better (higher percentage correct) in all conditions across all signals compared to non-musicians because of their extensive experience with musical tones. These results would indicate that musical experience can influence our perception of timbre.

2. METHODOLOGY

2.1 Participants

A total of 30 University of Calgary students (9 males and 21 females) participated in this 1-hour experiment for bonus course credits. Each participant filled a self report to indicate their hearing abilities and musical training. The division of participants into groups was made on the basis of their self reports. Nine participants were naïve listeners (i.e., they had received no musical training), 11 had received some form of musical training (i.e., they had received less than 8 years of musical training). These 20 participants (mean age = 22.8, *SD* = 2.96) were considered to be the non-musicians in this study. Only 10 participants (mean age = 21.9, *SD* = 2.38) were considered to be musicians (they had received 8 or more years of musical training). All participants reported normal hearing on a self-report questionnaire. This experiment was posted on a website system that scheduled and tracked experiments for students. Each participant received one bonus course credit after participating in the experiment.

2.2 Apparatus

The stimuli in this experiment were recorded and prepared in the Faculty of Fine Arts, Department of Music in the

Electro-acoustic Laboratory at the University of Calgary. Tonal samples of three above mentioned instruments were performed by experienced university musicians who played single tones on each of the three instruments. The tones were recorded and reproduced with by a 44 100 16-bit samples per second by a DigiDesign Audiomedia digital board controlled by a Macintosh G3 computer. Various programs on a Macintosh G3 computer generated and prepared the stimuli for the study. The recorded instrumental tones were edited in Peak (Version 2.1, produced by Berkley Integrated Audio Software). Those pitches that were not played by the musicians were transposed on Sonic Worx (Version 1.0.0, produced by Prosoniq Products Software). Tones were transposed up or down two semitones, at most, in order to reduce any effects of distortions in the sound waves. The edited tones were then imported into ProTools (Version 5.1, produced by DigiDesign, Avid Technology, Inc.) where prepared samples of two different timbres were merged and converted into interleaved stereo files at a rate of 44,100 16-bit samples per second. Pure tones in this experiment were generated using SoundMaker (Version 1.01). Using this program, these tones were also mixed into interleaved stereo files at the same rate as the instrumental tones.

2.3 Stimuli

The instrumental timbres used in this study were produced by a clarinet, a trombone and a harp. Two of the three timbres were combined in the experimental conditions. The pure tones in this experiment were fundamental frequencies corresponding to the instruments.

In the program Peak, the attack and decay portions of the amplitude envelope of each instrumental tone were excised to generate steady-state signals. The operational definition of envelope constituent boundaries at present is still not operationally defined (Hajda et al., 1997); therefore, it was determined in this study that the steady-state portions of the signals were the regions where the amplitude of the envelope fluctuated the least (i.e., appeared most at a plateau). All signals were generated to be approximately 300 ms in duration, similar to previous experiments of Grey (1977)).

Ten selected notes (or frequencies) were used in this study: Eb4 (311.13 Hz), G4 (392.00 Hz), Ab4 (415.30 Hz), A4 (440.00 Hz), Bb4 (466.16 Hz), B4 (493.88 Hz), C5 (523.25 Hz), C#5 (554.37 Hz), D5 (587.33 Hz), and Eb5 (622.25 Hz) (Pierce, 1983). These 10 notes were prepared for each of the three experimental instruments. There was also a control condition in which only pure tones were used which corresponded to the fundamental frequencies of these timbres. Briefly explained, the perception of instrumental pitches is created by many harmonic (or pure tone) components. Pure tones were included in order that the fundamental frequencies (or the first harmonics) of the 10 notes were heard by the participants. Hence, there were a total of 40 different signals: 30 experimental signals produced from three instruments, and 10 control signals produced by pure tones.

In ProTools, the 40 individual signals were combined into 40 pairs of stimuli. Each trial consisted of a hybrid of two instrumental timbres. One timbre was consistently at Eb5 and was paired with another timbre at 10 different notes, starting from Eb4 (one at Eb5 and the other at Eb4) (see Figures 1 and 2), which gradually blended together as one (both at Eb5) (see Figures 3 and 4). The combinations of the three timbres yielded three experimental groups: 1) clarinet-Eb5 / trombone (at 10 different pitches); 2) harp-Eb5 / clarinet (10 pitches); and 3) trombone-Eb5 / clarinet (10 pitches). The fourth group was the control group which consisted of fundamental frequencies corresponding to the combination of timbres in the experimental condition.

The combination of the 10 notes from the four groups produced 40 different stimuli. Each stimulus was repeated 10 times in random order for a total of 400 trials. These trials were broken into five blocks with 80 trials each. Each stimulus was repeated twice in each block. The stimuli were equalized for perceived loudness and duration, in order to reduce any confounding dimensions related to the judgments on timbre (Grey, 1977).

Examples of the Eb4 and Eb5 timbral combinations are shown in the spectrograms of Figures 1 thru' 4. Figure 1 shows the spectrum of two pure tones, one at Eb4 and the other at Eb5. Figure 2 shows the spectrum of a clarinet at

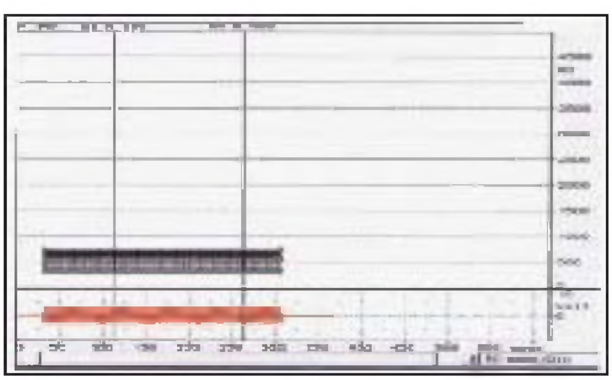

Figure 1. Spectrogram of Control Eb4 and Eb5 combinations.

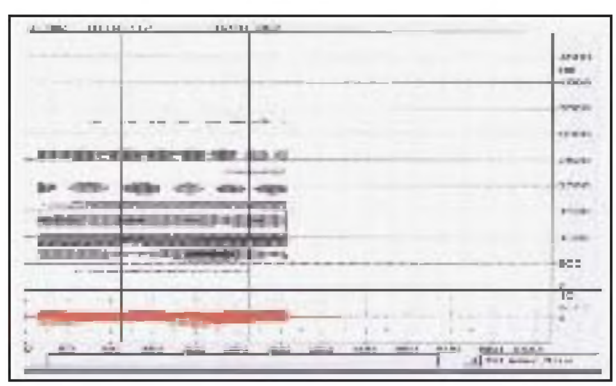

Figure 2. Spectrogram of Clarinet Eb5 and Trombone Eb4 combination.

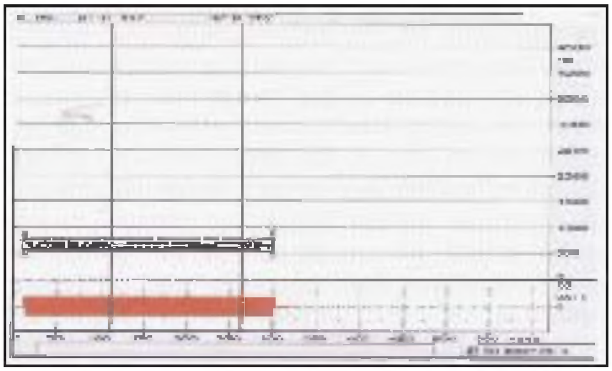

Figure 3. Spectrogram of Control Eb5 and Eb5 combination.

Eb5 combined with a trombone at Eb4. Figure 3 shows the spectrum of two pure tones at Eb5, and Figure 4 shows the spectrum of a clarinet and trombone both at Eb5. When comparing the two groups of spectra, the Control signals (Figures 1 and 3) show only the fundamental frequencies, while the Experimental signals (Figures 2 and 4) show many more partials, or mixtures of pure-tone frequencies.

2.4 Procedure

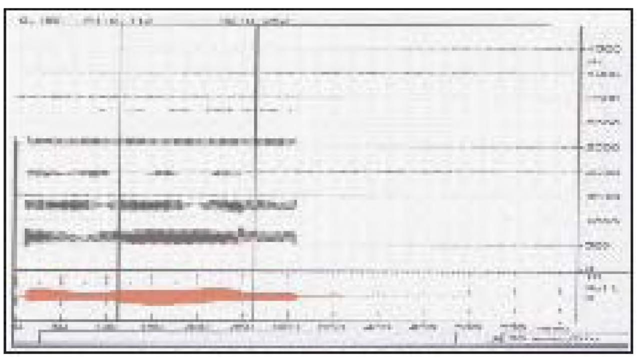

Figure 4. Spectrogram of Clarinet Eb5 and Trombone Eb5 combination.

All tests were conducted in the Speech and Audition Laboratory at the University of Calgary and each test was completed in an one-hour session. Participants were tested individually or as a group (up to six people). Participants were seated comfortably in a semi-circle facing the sound source. Each participant signed a consent form and filled out a questionnaire regarding their musical experience. Participants were verbally told about the experimental task, however, they were not told that there would be two timbres in each trial, nor were they aware of the different types of instruments involved. Before the experimental trials, there was a practice session in which participants heard 300 ms samples of a clarinet, a trombone, and a harp and samples of pure tones, each played at Eb4 and Eb5. They also heard random examples of the test stimuli (the combined signals).

For each trial, participants had to immediately offer a response to the combined timbre and decide whether they heard one or two instruments. They were required to write “1” or “2” in the numbered space provided for them on the answer sheets. There was a 2 sec break between trials for participants to make their responses. After each block of 80 trials, participants were offered the opportunity to take a short break. At the end of the experiment, participants were given a debriefing sheet containing more information about this study. The experimenter also verbally debriefed the participants.

3. RESULTS

The raw data consisted of 400 responses from each of the 30 participants. Participants’ responses of “1”s and “2”s were converted into raw percentages: “1” being an incorrect response was assigned a 0% and “2” being a correct response was assigned a 100%. With this new score, a percentage correct was calculated for each of the 40 timbres for each participant. Therefore, a set of 40 percentages for each

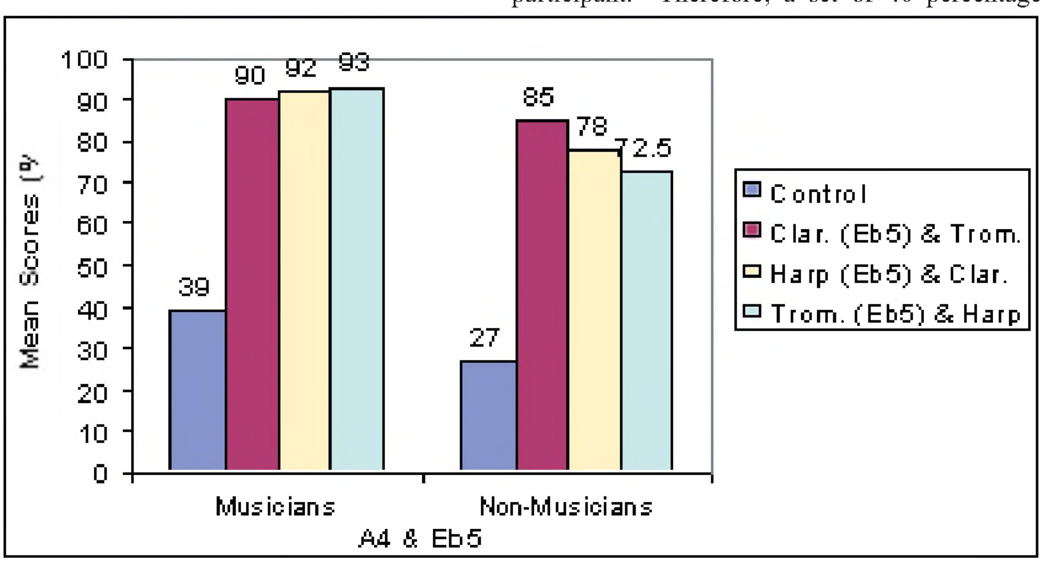

Figure 5. Mean scores for Musicians and Non-musicians for Eb4 and Eb5 combinations.

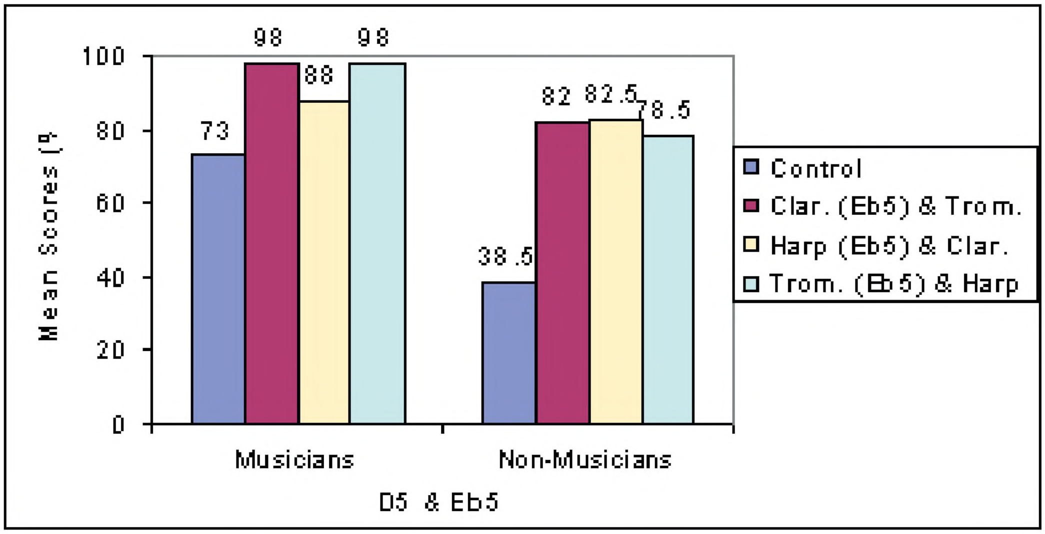

Figure 6. Mean scores for Musicians and Non-musicians for D5 and Eb5 combinations.

participant was obtained.

Mean percentages on the correct number of instruments identified was analyzed using a 10 (combinations) by 4 (conditions) by 2 (groups: musicians, non-musicians) repeated-measures analysis of variance. The omnibus ANOVA revealed that there were significant main effects of combination, $F(9,252) = 65.54$, $p < .001$ and condition, $F(3,84) = 11.78$, $p < .001$ as participants performed differently for each combination and each condition. The results of the analysis indicated that the mean values for group by combination were statistically significantly different, $F(9,252) = 3.01$, $p < .002$, as were the mean values for condition by combination, $F(27,756) = 25.40$, $p < .001$. Thus, there was a systematic difference between musicians and non-musicians, and the effect of condition and combination was different for musicians and non-musicians.

There was also a significant 3-way interaction between the three variables, $F(27,756) = 3.72$, $p < .0001$, which showed that the effect of one variable depended on the levels of two other variables. Further analyses were required to find the source of the significant interactions, however, post-hoc analyses were not conducted due to the large amount of possible comparisons; an unacceptably high error rate would result if all possible pairs of means were compared. However, since the focus of the study was on performance differences between musicians and non-musicians, overall means and standard deviations (SD) were compared and certain timbral combinations of interest were presented here; these included the Eb4, A4, D5 and Eb5 combinations (see Figures 5 to 8).

The results of the A4 and Eb5 combinations (Figure 5) show the typical findings for most of the other combinations. Generally participants performed much better in the Experimental conditions than in the Control condition. On average, both groups could perceive the two timbres, but it appeared that musicians were consistently better able to do so. For the D5 and Eb5 combinations (Figure

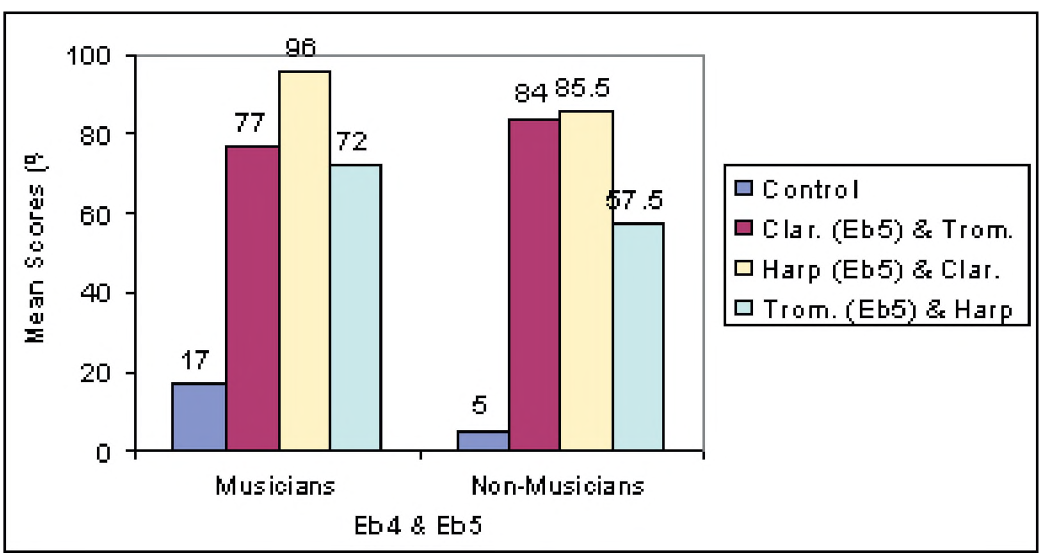

Figure 7. Mean scores for Musicians and Non-musicians for Eb4 and Eb5 combinations.

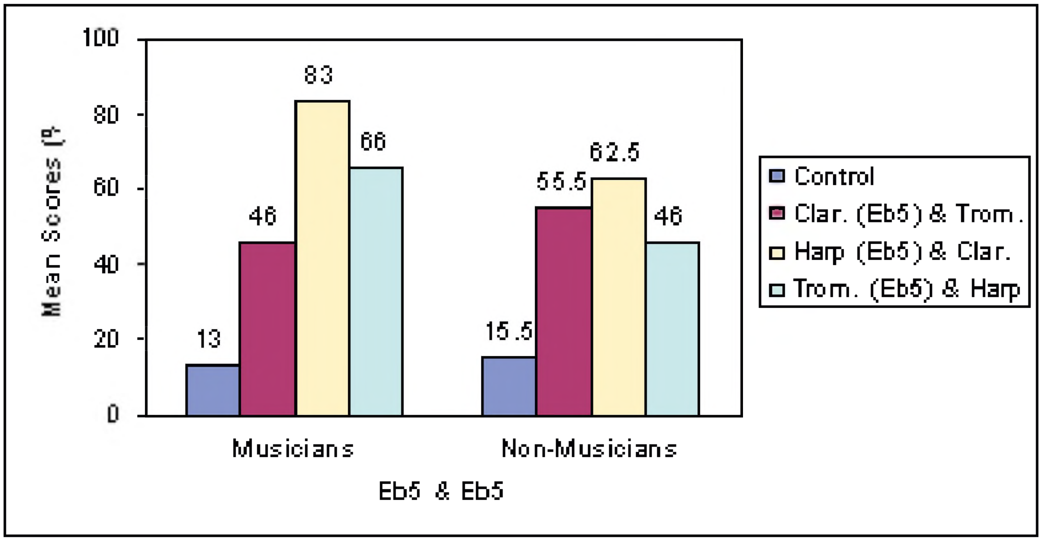

Figure 8. Mean scores for Musicians and Non-musicians for Eb5 and Eb5 combinations.

6), both groups performed well across all conditions, with musicians performing better. It was found that for the octave combinations (Eb4 and Eb5) (Figures 7 and 8), means scores for both groups were typically lower. It is interesting to note that two combinations where non-musicians performed better than musicians were the octave combinations played by the clarinet and trombone pair. The first signal was the Clarinet-Eb5/Trombone-Eb4 signal where non-musicians achieved a mean of 84% (*SD* = .230), and musicians a mean of 77% (*SD* = .295). Clarinet- Eb5/Trombone-Eb5 was the other signal where non-musicians (*m* = 55.5%; *SD* = .287) outperformed musicians (*m* = 46%; *SD* = .353).

4. DISCUSSION

In this study, the perception of timbral fusion in musicians and non-musicians was examined. The hypothesis that musicians would perform better than non-musicians across all conditions was mainly supported. Significant differences were found in this study. Overall, musicians had higher percentage scores than non-musician in identifying that there were two timbres in each stimulus. Thus it was easier for musicians, than non-musicians to hear two simultaneous sounding timbres. This may suggest that musical training or experience enhances the perception of timbre. This difference between musicians and non-musicians is in accordance with other studies (e.g., Kendall, 1986).

Pairs of instrumental timbres (Experimental Conditions) were more accurately perceived by both groups than pairs of pure tones (Control Condition). This was also found in Miller and Carterette's (1975) study where participants had to make similarity judgments between pairs of tones. In their study no differences were observed between musicians and non-musicians in the judgments of the fundamental frequencies. However, in the present study, musicians on average performed slightly better on pure tone combinations than non-musicians even though musicians typically do not have experience with sine tones. In a study (Spiegel & Watson, 1984) which involved frequency-discrimination tasks by musicians and non-musicians, single tones, including 300 ms sine-wave and square-wave tones, and complex sequential patterns of ten tones were used. In the single tone condition, musicians attained thresholds that were lower (better discrimination performance) to only one-half of the non-musicians. The other half of the non-musicians attained thresholds almost as low as musicians. The experimenters suggested that these listeners had probably gained a great degree of psychoacoustic experience and had learned quickly to discriminate single tones. Therefore, musicians tend to perform better than non-musicians on studies of the auditory system probably because of their ability to transfer their previous musical training to new tasks.

The differences in performance between the Experimental Conditions and the Control Condition in this present study can be explained by the characteristics of the waveform. All tones produced by musical instruments are not pure tones but mixtures of pure-tone frequencies or so called partials (White & White, 1980). The perception of fusion depends on the synchrony of the frequency partials in complex sounds, therefore, the fusion of two instrumental timbres is often harder to perceive by non-musicians because there is a lower probability of synchronicity between all the partials; in other words, there is a higher probability of segregation when the partial are not strictly harmonic (Handel, 1989). In contrast, a pure tone has only one harmonic (the first harmonic, or the fundamental frequency (F0)), therefore when two pure tones are combined there is a higher probability of fusion; however, this also depends on whether the two tones are harmonic or not (Handel). For example, participants reported after the experiment, that the signal of Control-D5 was usually heard as two sounds, whereas, Control-Eb4 and Eb5 were easily confused as one.

Generally, the octave combinations were more difficult for listeners to perceive as two timbres. A similar result was also found in a study by Handel, Molly and Erickson (2001). The participants in that study were unable to determine whether two different notes separated by an octave were played by an identical or a different wind instrument. The researchers concluded that “listeners can extrapolate the timbre of an instrument or voice over only a relatively short pitch range” (p. 126). When tones are separated by an octave they are considered to be musically and perceptually equivalent (they are given the same name) (Handel, 1989). Physically, the octave is the only interval in which the harmonics will coincide exactly, therefore, two notes that are separated by octaves cannot create dissonance. Numerous studies have found that octave equivalence is perceived by both experienced and inexperienced listeners, however, the percentage of accuracy ranged from 33% to 50% (Handel). It can be seen in this study that the octave combinations was often lower compared to the other pitch combinations. Surprisingly, the only combinations where non-musicians appeared to perform better than musicians were the Eb4 and Eb5 combinations (i.e., Control-Eb5, Clarinet-Eb5/Trombone-Eb4 and -Eb5). This result was unexpected; perhaps musicians have more experience with harmonic sounds (i.e., triads and chords), therefore they can fuse the harmonic partials more readily than non-musicians.

One limitation to the current study which was similar to previous studies (i.e., Grey, 1977) was the brevity of the signals (300 ms). Steady state timbres were used as the stimuli for this experiment because our goal was to examine only spectral features of timbre. However, other transient features play a crucial role in helping with the identification of an instrument (Iverson & Krumhansl, 1993), especially the attack (Saldanha & Corso, 1964). Since musical tones are in fact, not like the signals of this study, it would be practical to replicate this study by using complete tones (with attack, steady state and decay) to see whether there will be a difference in performances; perhaps there will be an increase in accuracy for both groups.

Although the use of musical passages as stimuli may provide realistic situations in which we may understand perceptual fusion (i.e., Bonfield & Slawinski, 2002), one must be cautious that such experiments may introduce many uncontrolled variables. The methodology of the current study may provide a more controlled way to study this perceptual phenomenon. However, a next step to this timbre-fusion paradigm is to include a single-tone condition. Although there were always two tones present in this study, listeners systematically reported hearing only one tone, therefore a follow-up study with a single-tone condition is necessary in order for a more comprehensive picture to emerge.

Another limitation of this study may be due to the instrumentals timbres used as the stimuli. We have attempted to equate for loudness and pitch across all signals. However, to perceptually equalize natural, albeit, brief tones is difficult because real musical signals are complex and time-variant. It is often the case that a single tone has variable pitch and loudness. For example, the harp has no true steady state; after a sharp attack, the amplitude envelope immediately decays. At no single time frame will the sound be exactly the same or equalized. Since the rationale of this study proceeded from the findings of Bonfield and Slawinski’s (2002) study, the instrumental timbres chosen were based on their previous study. Perhaps this study could be replicated by using other instrumental timbres. Participants stated that the trombone and the clarinet sounded very similar, thus they perceived the stimuli of any trombone and clarinet to be identical. But at the same time, it would be interesting to further investigate why the particular combination of clarinet and trombone played at Eb5 was harder to perceive compared to other pitch combinations.

The study of timbre as a musical attribute has received much more attention in recent decades (e.g., Saldanha & Corso, 1964; Grey, 1977; Risset & Wessel, 1999). However, we still do not fully understand its multidimensional nature. Timbral fusion or timbral combination appears to be an important area of study since most music is created by the simultaneously sounding of instruments, however, timbral fusion is still a relatively unexplored area in timbre research. In recent years, musicians and composers have taken interest in the science of music. Therefore, this study may be important in both the fields of music and psychology. The preliminary findings of this research will be useful for traditional composers and orchestrators, as well as electronic composers.

In listening to orchestral music, the perception of fusion also depends on the type of conducting. Conductors may manipulate timbre by combining and emphasizing instruments in a certain way. The fact that timbral manipulation is practiced in the real world, reminds us that, in order to understand music in an ecological sense, we must strive to understand timbre. This study, despite its limitations, has taken a step further into understanding the phenomenon of timbral fusion, as well as gaining more knowledge on the musical attribute of timbre.

5. ACKNOWLEDGMENTS

The authors would like to thank David Eagle for his assistance throughout this study, and Carrie Alain for her help with the stimuli.

6. REFERENCES

American National Standards Institute. (1960). *USA standard acoustical terminology.* New York: American National Standards Institute.

Bonfield, S. & Slawinski, E. (2002). Targeting timbre recognition abilities of musical participants: when musicians hear what isn't there. *Canadian Acoustics, 30*, 160-161.

Grey, J.M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America, 61*, 1270-1277.

Hajda, J.M., Kendall, R.A., Carterette, E.C., and Harshberger, M.L. (1997). Methodological issues in timbre research. In I. Deliége, & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 253-306). UK: Psychology Press.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events.* Massachusetts: MIT Press.

Handel, S. & Erickson, M.L. (2001). A rule of thumb: The bandwidth for timbre invariance is one octave. *Music Perception, 19,* 121-126.

Iverson, P. (1995). Auditory stream segregation by musical timbre: effects of static and dynamic acoustic attributes. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 751-763.

Iverson, P., & Krumhansl, C.L. (1993). Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America, 94*, 2595-2603.

Kendall, R.A. (1986). The role of acoustic signal partitions in listener categorization of musical phrases. *Music Perception, 4*, 185-214.

Kendall, R.A., & Carterette, E.C. (1991). Perceptual scaling of simultaneous wind instrument timbres. *Music Perception, 8*, 369-404.

Kendall, R.A., & Carterette, E.C. (1993). Identification and blend of timbres as a basis for orchestration. *Contemporary Music Review, 9*, 51-67.

Miller, J.R., & Carterette, E.C. (1975). Perceptual space for musical structures. *Journal of the Acoustical Society of America, 58*, 711-720.

Pierce, J.R. (1983). *The Science of Musical Sound.* New York: Scientific American Library.

Plomp, R. (1976). *Aspects of tone sensation: A psychophysical study.* London: Academci Press.

Risset, J-C. & Wessel D.L. (1999). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The psychology of music* (2nd ed.) (pp. 113-169). San Diego: Academic Press.

Saldanha, E.L. & Corso, J.F. (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America, 36,* 2021-2026.

Singh, P.G. (1987). Perceptual organization of complex-tone sequences: A tradeoff between pitch and timbre? *Journal of the Acoustical Society of America, 82*, 886-899.

Spiegel, A.F. & Watson, C.S. (1984). Performance on frequency-discrimination tasks by musicians and nonmusicians. *Journal of the Acoustical Society of America, 76,* 1690-1695.

White, H.E. & White, D.H. (1980). *Physics and music: the science of musical sound.* Philadelphia: Saunders College.

From the manufacturer of Scamp® Sound Masking Systems comes the revolutionary LogiSon™ Acoustic Network: the first and only sound masking, paging and music system to provide full, digital control of individual speaker settings from a central location.

Acoustic Control

At Your Fingertips™

The Network offers the highest level of component integration in the industry. Each hub contains a random masking sound generator, amplifier and even independent volume and equalizer controls for masking and paging, eliminating the need for most centralized audio equipment.

Timer and paging zones, volume and equalizer settings, and paging channel selection are configured through a centrally-located control panel or remotely with user-friendly LogiSon™ Acoustic Network Manager software.

This digital contol and accuracy, combined with small zone sizes and one-third octave equalization capabilities, allows you to custom tune the Network to suit the unique requirements of the space, increasing the masking's effectiveness.

And, as your needs change, the Network can be completely re-adjusted without rewiring or requiring access to the ceiling.

Though the LogiSon Network is typically installed above the suspended ceiling, it has been designed for visible applications and will compliment the investment you make in the professional appearance of your facility.

LogiSon™ ACOUSTIC NETWORK ©

For more information, call 1 (866) LOGISON or visit www.logison.com to request an Information Package.