Perception of Blended Timbres: Effects of Instrumentation and Sound Onset

Samantha N. Amberg

American University, Department of Psychology

University Honors, Spring 2013

Advisors: Zehra F. Peynircioǧ lu, William Brent (Audio Technology)

Abstract

Timbre, the color or quality of a tone, is the aspect of sound that distinguishes two sounds of equal loudness and pitch. The questions posed by the current study ask how instrumental timbre-blends are perceived and labeled by musicians and what influences instrument type and sound attack have on their identification. In this experiment, live recordings of cello, saxophone, violin, clarinet, and piano were blended to develop 20 distinct timbre dyad stimuli; 10 of which consisted of a single isolated tone and 10 of which consisted of the same isolated tone with the sound attack (onset) removed. Experienced musicians were presented with a sliding continuum of these timbre dyad blends and were prompted to locate the pure component timbres on the continuum. Accuracy of timbre component identification was compared across instruments and between the two experimental conditions. Main effects of blend type and attack removal were shown as well as a significant interaction effect between blend type and attack removal on identification accuracy. Secondary effects of pair asymmetry and ease of individual instrument identification were also found. The results offer a novel approach to the study of timbre by investigating its facets in a practical manner in real instruments, rather than in synthesized music. It also provides insight regarding the perception of certain instruments that would be valuable to a composer looking to achieve perceptual effects.

Perception of Blended Timbres: Effects of Instrumentation and Sound Onset

Timbre, the color or quality of sound, is the element that allows one to distinguish a clarinet from a cello when they are playing the same note at the same volume. Trained musicians and composers manipulate timbre along with pitch and loudness, to achieve their desired result which the audience is then able to enjoy. Timbre, especially in its technical and physical sense, fascinates musicians as it is vital for producing a great work.

Psychologists, particularly those interested in the perception and cognition of music, have also taken an interest in timbre and have begun to explore the ways that listeners process it. Distinct musical timbres can be recognized very quickly, outpaced only by vocal sound recognition (Augus et al., 2012). Research by Menon et al. (2002) has shown that timbre is processed bilaterally, while Brancucci and San Martini (1999) found a left ear advantage for reaction time. The neural responses to successive timbres, which have been studied using magnetoencephalography, are determined by if the two stimuli are identical or different (Seol J., Oh M., Kim J.S., Jin S., and Kim S. I., 2011). Overall, however, timbre processing is less well understood than that of other musical and auditory attributes, due partially to the difficulty identifying the underlying perceptually relevant acoustic parameters (Hajida et al., 1997; Menon et al. 2002).

Audio technology has advanced from the record to the cassette tape, all the way to modern MP3s at a very rapid pace. The 1970s were an age of acoustic exploration with the advent of frequency modulation (Chowning, 1973), granular synthesis (Roads, 1978), and re-synthesis (Grey & Moorer, 1977). As sounds have increased in complexity and decreased in their connection to physical instruments, listeners have been exposed to a great deal of unfamiliar timbres. Modern computer programs allow sounds to be precisely manipulated, blended, and synthesized to form entirely new sound combinations that would never occur naturally (Karplus & Strong, 1983; McPherson, 2010). Such artificial blends are created by digital analysis and recombination of existing instrumental recordings. These blends can serve as compositional tools, forming variations like morphing or distortion effects. Composers are utilizing progressively enhanced software to merge the spectra of two (or more) separate sounds, which contributes to finished product (Haken, Fitz, & Christensen, 2007; Ezzat, 2005; Caetano & Rodet, 2010). This ability to use timbre as a fine-tuned artistic medium has brought it to the forefront of contemporary performances and compositions by allowing musicians to tweak each recording to perfection. For example, many popular vocalists depend on synthesizers to produce harmonies and vocal feats they would not be able to produce on their own. Additionally, timbre affects perception of emotion in music (Bowman, 2011; Hailstone et. al, 2009), which is arguably the most important goal of a musical composition.

Although these complex timbre blends play a large role in contemporary music and have been studied theoretically for many years, there is a gap in research regarding actual human perception of such sounds. Research may have been deterred by the multi-dimensionality of timbre, a quality that makes it difficult to isolate as a perceptual variable. Some psychologists, however, have taken on the challenge to pursue timbre from cognitive and perceptual perspectives. For example, it has been found that exchanging the spectral energies of two tones by manipulating their centroids using computer programs, causes a corresponding switch in positions within perceptual space (Grey & Gordon, 1977, Grey,1978). In another study, participants listened to sounds of various instrumental timbres in succession (some pure parental timbres, some blended hybrids) and indicated if they believed the two sounds were from the same source or different sources. These similarity judgments were used to compare parental and hybrid sounds, which illustrated that some synthesized hybrids are perceived as a middle ground between their parental sources (McAdams et al., 1995).

Caclin and colleagues (2005) strategically altered parameters of synthesized tones, including attack time, spectral centroid, spectral fine structure, and spectral flux and asked subjects to rate their similarities. All dimensions but spectral flux (attack time, spectral centroid, and spectral fine structure) were used by participants when making dissimilarity ratings for synthetic tones, indicating their perceptual relevancy. All of these components are relevant to the current study, as ten distinct instrument dyad-blends will be compared by participants, each with unique attack times, spectral centroids, spectral fine structures, and spectral fluxes that may or may not influence blend perception. Additionally, attack time will be removed for a portion of the trials, further investigating its perceptual role in timbre. The current study also resolves the issue of utilizing completely synthetic sounds as stimuli, by blending real recordings of the five instruments.

The effect of altering the blend ratio upon timbre perception and identification of instrumental sounds is currently unknown. Listeners may be able to correctly identify each new ratio as a novel sound, easily distinguishing between the various blends. Accurate recognition and judgment of the spectral shape and temporal envelope would be key skills in tracking these minute changes. Alternatively, listeners may perceive the sound blends as falling into either parental timbre category which would indicate categorical perceptionInstruments are not often heard playing isolated single tones; they are heard playing melodies in polyphonic music. The existing research on the influence of melody on perception does not provide a unified conclusion on the topic. A study by Grey (1978) indicated that the role of melody in auditory perception may be related to the type of instrument. When participants labeled the instrumental sources of various sound samples (isolated tone, single voice melody, and multi-voice melody),clarinet and trumpet were best identified from the isolated tone condition and bassoon was best identified from the single voice melody condition. How this extends to other instruments and synthetic blends is not known, but it may be related to physical differences in individual instrument spectra. The relative perception of timbre appears to be much weaker than that of pitch, with pitch variations apparently interfering with judgment of timbre relations (Krumhansl & Iverson, 1993). Conversely, an experiment by Goad and Keefe (1992) observed no significant differences between responses for isolated tones and responses for melodies when discriminating instruments in a concert hall. The implications from Grey (1978) and Krumhansl and Iverson (1993) findings say that timbre perception is affected by melody versus pitch stimuli. However, the implications from Goad and Keefe (1992) contradict these conclusions.

Pitch variation has also been identified as affecting timbre perception and identification. Williams (2006) observed listeners' judgments of timbre invariance across three octaves using a horn and bassoon. Pitch judgments did not vary across either instrument's range and musicians were able to identify similar timbres over an impressive 2.5 octave span. In a study on how musicians use pitch and timbre to classify tones, comparisons varied in difficulty as a function of the pitch-timbre combination (Pitt, 1994). For example, the most number of errors were made when the trumpet played a lower note and the piano played a higher note, perhaps making the sound seem out of tune. Another study found that timbre can induce pitch shifts in vocal, viola, and trumpet sound samples, but that these shifts may not be perceptually salient(Vurma, Raju, & Kuuda, 2010).

Every musical instrument produces sounds with a unique timbre, a quality that the current study aims to investigate. Various psychologists have studied listeners' abilities and tactics for differentiating between such tonal qualities. Fabiani and Friberg (2011) examined the influence of timbre on the perception of instrumental dynamics by having participants rate sounds on a scale ranging from soft pianissimo to robust fortissimo. Timbre was especially pertinent in the rating of trumpet and violin dynamics, moderately important in clarinet, and the least important in piano and flute. Differences in judgment strategy varied depending on the instrument, as indicated by the different main effect sizes. However, this only applied to the perception of individual instruments, not blended or polyphonic sounds. If these methods were applied to blended or polyphonic sounds, then the judgment strategies used may be different than for solo instruments.

In another study (Kendall & Carterette, 1991), listeners evaluated several timbre attributes of wind instrument dyads and assigned descriptors to them. Trumpet, saxophone, and oboe were verbally described with notable attributes of brilliance, richness, and nasality respectively; but clarinet and piano were not associated with any distinct attributes. These results indicate that instruments perceived as having more extreme timbres should be more easily discriminated in a blend and may mask other, more balanced instruments. However, only wind instruments were studied, leaving a gap in applying results to other orchestral combinations. Another In another experiment listeners made similarity judgments for sampled from various instruments in a concert hall. Participants were better able to discriminate flute, clarinet, and trumpet than marimba and viola. This implies that instruments with a greater sustained tonal segment have increased timbre discriminability (Goad & Keefe, 1992). The central aim of the current study is to investigate the components of instrumental timbre perception in polyphonic, blended stimuli which have not been addressed by the current literature. The instruments used will be clarinet, saxophone, cello, violin and piano.

The impact of sound onset or attack removal on timbre perception is debated in psychological literature. The attack of a sound is the way it is initiated, or the first portion of the sound envelope. A study by Iverson (1995) concluded that the attributes that influence timbre perception are present in both the onsets and remainders of sounds, although he did not identify what these attributes were specifically. In layman terms, both the beginning and end of a sound hold the information necessary for timbre similarity judgments to be made. This suggests that similarity judgments of timbre may rely on global acoustic comparisons, rather than specific aspects of the attack. However, recent research found that the presence of the sound attack increased students' abilities to identify various wind instruments; the clarinet and oboe were particularly effected (Paul, 2005). In another experiment, Tardieu and McAdams (2012) used blended dyads of instruments playing both impulsive and sustained tones to examine the underlying audio descriptors. The results showed that sound onset and spectral envelope (spectral centroid and spread) were pertinent to timbre similarity judgments and slight onset differences favored sound segregation. In other words, if two instruments sounded different in the primary portion of the sound, participants were better able to distinguish them from one another. Thus, it is likely that onset characteristics and duration influence the importance of attack when identifying timbres, with distinctiveness associated with better similarity judgment accuracy. In the current study, if the sounds with attack removed are less accurately identified than those with the attack intact, then the results will indicate dependence on attack for perceptual judgments of timbre.

In order to elaborate on and unify the findings of timbre perception research, the current study incorporates instrument variety and onset presence into its methods. By including only musicians, we have attempted to control for knowledge effects, although individual specialties may impact timbre identification accuracy. Recorded sounds from an alto saxophone, violin, cello, clarinet, and piano were used to create paired blend continuums. Participants were asked to identify the pure parental timbre components located on the continuum for stimuli both with and without attack. The accuracy of these judgments probed the interaction of instrument type on blended timbre perception, as well as the role of sound attack. This study is crucial in exploring the relationship between modern music production techniques and the actual experience of the listener. It will also provide information on how different manipulations of sound (instrument pairing, attack removal) manifest in perceptual data.

Method

Participants

A total of 12 adults in the Washington, D.C. area participated in the experiment on an entirely voluntary basis. Those who had 7 or more years of experience playing an instrument were considered musicians and were eligible for participation. This limitation was imposed to ensure the participants had background knowledge of the instruments and their individual sounds upon which to base their judgments. Seven years was selected as the minimum amount of training to model after the procedures of Guclu, Sevinc, and Canbeyli (2011). No participants had any hearing difficulties. AU students who participated in the study were offered 1.5 research credits for eligible psychology courses as compensation. No compensation was offered to non-AU student participants. Participants completed an informed consent form prior to participating and received a debriefing form after the study.

Materials

Sound stimuli for the experiment were created using live recordings of musicians playing the alto saxophone, clarinet, cello, violin, and piano. Each instrument was recorded playing a single tone: A440. Ten stimuli of unique instrumental dyads were created for each pitch, pairing each instrument with every other instrument. B=0 produced a pure tone of instrument A and B=1.0 produced a pure tone of instrument B. Coefficient values between the pure timbres produced a steady graduation of timbre A/B blends that were equated for pitch and volume, forming hybridized sounds. The continuum formed a circle for each dyad, with B= 0 and B=1.0 at randomized points. The same recordings had the initial 20 ms removed to simulate attack removal. Blends were created in an identical manner, except with the attacks removed. In all, 2 sets (20 unique stimuli) were generated with these methods. The experiment was conducted on a MacBook Pro using the Pd Gem program. All stimuli were listened to using SONY headphones.

Design and Procedure

There was a single presentationset, including both the single tone with attack and single tone without attack blends. There were 120 trials in the set, comprising of three identifications of each instrument for each pair (60 with attack 60 without attack). All 120 trials were completely randomized with respect to instrument dyad, attack/no attack presentation order, and component instrument target. For example a cello-violin blend had four possible presentations a target instrument of violin, a target instrument of a cello, a target instrument of cello with the attack removed from the blend, and a target instrument of violin with the attack removed from the blend. Participants were not informed about alterations to the sound attack and the sound blending techniques. Prior to beginning the task, participants were given a trial round using trumpet and oboe blends to become accustomed to the task. Then, they were played a sample of all five pure instruments. At any point during the experiment they were able to listen to the pure instrument samples again. This was accomplished by providing a computer window with the sound files where the participants could click and listen to the desired instrument.

Participants had to slide a red dot from a randomly set point to the point on the circle continuum they believed was the best sounding target instrument. The target instrument was indicated in the upper left hand corner of the window containing the continuum (see Figure 1-Pending). At any point on the circular continuum, a sound sample of the corresponding blend ratio would be played through headphones a single time. When the participants located the point on the continuum they believed best matched the target instrument, they pressed the space bar to move to the next trial.

Results

Attack presence might contribute to instrument identification accuracy, but that effect may differ across instrument blend types. Instrument identification scores were subjected to a two-way analysis of variance having ten levels of blend type (Cello-Clarinet, Cello-Piano, Cello-Sax, Cello-Violin, Clarinet-Piano, Clarinet-Sax, Clarinet-Violin, Piano-Sax, Piano-Violin, Sax-Violin) and two levels of attack condition (Attack, No Attack). All effects were significant at the .05 significance level. The statistical results are discussed in the following section.

There was a significant main effect of instrument blend type $F(9, 63) = 6.44$, $p < .001$. Subsequent pairwise comparisons with the Bonferonni correction (.05/10) revealed that Cello-Violin (M= 25.86, SD =31.99) and Clarinet-Sax (M=16.88, SD =21.75) blends had significantly greater deviation from perfect identification than any of the other blends, with Cello-Violin being worse than Clarinet-Sax. The mean differences, p-values, and confidence intervals from the pairwise comparisons are displayed in Table 1. There was also a significant main effect of attack condition $F(1,71) = 4.44$, $p < .05$. Overall, blends with the attack removed ($M = 10.49$, $SD = 1.26$) deviated further from perfect identification than blends with the attack intact ($M = 9.62$, $SD = 1.11$). There was a significant interaction effect between blend type and attack condition $F(9,63) = 3.30$, $p <. 05$, indicating that the impact of attack removal depends on the instruments in the blend (Figure 1). The mean overall deviation from perfect identification was calculated for each instrument. Violin ($M = 10.3$, $SD = 18.64$) and clarinet ($M = 10.2$, $SD = 15.80$) had the lowest overall identification accuracy, while saxophone ($M = 7.2$, $SD = 10.06$) had the greatest ease of identification.

Additional analysis was achieved by conducting two-tailed independent t-tests. Each blend pair had two possible target instruments, one of which may be less accurately identified than the other, an effect we term pair asymmetry. A t-test was conducted for every blend pair, with one sample from target instrument ‘A’ identification and the other sample from target instrument ‘B’ identification with the Bonferonni correction (.05/10). The piano was worse identified than its blend counterpart when paired with all other instruments. In the piano-cello blend, piano ($M = 12.22$, $SD = 7.35$) was less accurately identified than cello ($M = 2.68$, $SD = 2.69$), $t(70) = 7.18$, $p < 0.0001$. In the piano-clarinet blend, piano ($M = 7.64$, $SD = 7.07$) was less accurately identified than clarinet ($M = 2.86$, $SD = 3.28$), $t(70) = 3.68$, $p < 0.0001$. Additionally, in the Piano-sax blend, piano ($M = 7.97$, $SD = 5.60$) had greater deviation from perfect identification than sax ($M = 2.55$, $SD = 2.96$), $t(70) = 5.13$, $p < 0.0001$. Lastly, when paired with violin ($M = 6.46$, $SD = 6.21$), piano ($M = 8.94$, $SD = 6.54$) was less accurately identified, $t(70) = 4.46$, $p < 0.0001$.

Clarinet ($M = 24.74$, $SD = 24.52$) was less accurately identified than saxophone at a marginally significant level ($M = 12.92$, $SD = 14.68$) when the two instruments were paired together in a blend, $t(70) = 2.48$, $p < 0.05$. Furthermore, clarinet ($M = 7.53$, $SD = 9.24$) was less accurately identified than cello ($M = 4.56$, $SD = 3.92$) at a level approaching significance, $t(70) = -1.90$, $p < 0.10$. This indicates that participants made more errors when identifying clarinet than its blend counterpart, when paired with cello and saxophone.

Discussion

We did not have any hypotheses regarding the current study. Rather, the goal was to investigate the perception of instrumental timbres in blends using a variety of instrument types. An additional aim was to clarify mixed results in the literature regarding the importance of attack in instrument identification (Iverson, 1995; Paul, 2005; Tardieu & McAdams, 2012).

It was found that blend type does have a significant impact on the perceptual identification of instruments. Violin-Cello and Clarinet-Saxophone blends were the most difficult for participants to identify, indicating that they have a perceptually unified sound and exist more congruently together than other blends. This result follows practical reasoning, since these two pairs are comprised of instruments of the same type: cello and violin are bowed strings, while clarinet and saxophone are both single-reed woodwinds. Composers may consider utilizing these blend pairs to more effectively elicit emotion from listeners. Emotional responses to music arise from complex manipulations of tension and resolution. By using more unified pairs, tension would be resolved, but by using less perceptually cohesive pairs tension could be built up.

Overall, attack removal appeared to have a detrimental effect on accuracy of instrument identification. Further analysis revealed that there was, in fact, an interaction effect between attack condition and blend type on instrument identification. This indicates that the impact of attack removal ranges depending upon the component instruments in the blend. The clarinet-saxophone blend was more accurately identified in the no-attack condition, while clarinet-violin and piano-violin blends were less accurately identified after attack removal. This result disputes the conclusion of Iverson (1995), that the attack component of a sound is not special or uniquely pertinent to perception. Alternatively, this is in agreement with the results of Tardieu and McAdams (2012), which also supported the conclusion that attack importance varies as a function of the blend components.

Saxophone was the most easily and accurately identified instrument, while clarinet and violin were the most difficult for participants to identify. This is similar to Kendall and Caterette's study (1991), in that saxophone timbre is easily discriminable and clarinet (and presumably violin) has a more indistinct timbre. Saxophone's perceptual salience makes it an ideal choice for a solo when composing music. The perceptually bland traits of clarinet and violin timbre, however, recommend them for parts in polyphonic music with many instruments playing together.

Lastly, from the data it can be concluded that there is pair asymmetry within some of the instrumental blend pairs. Piano was more difficult to perceptually isolate than any of its blend counterparts. Clarinet also proved to be more of a challenge to identify than its pair when blended with cello and saxophone. These two instruments perceptually 'pop', obscuring the other blend component. This masking effect causes the blend to sound like pure piano/clarinet when it still technically impure. Musicians and composers can use this knowledge of masking qualities and pair asymmetry to control the homogeneity of the music they are creating. Piano should only be used when the other polyphonic components are not significant, since it will 'pop' out and draw attention away from the other instruments. Alternatively, if a composer wanted clarinet to blend in with other instruments he should pair it with violin which is does not perceptually overshadow, as indicated by the lack of a pair asymmetry effect for the violin-clarinet blend.

There were several limitations in this study, although significant results were found for three main effects and two secondary effects. The study would have benefitted from higher quality instrument sample recordings. For example, standardized sample recordings of instruments from a sound bank would have offered a more uniform sound and, ultimately, cleaner blends. It would also have eliminated the risk of slight tuning differences between instruments which some participants brought to the experimenter's attention. This may have altered the tactic used to identify the instrument in the blend and shifted the focus from timbre to pitch. Perhaps it may even have influenced the ease of identification for certain instruments. Lastly, if participants had been asked to emphasize perfection and accuracy in the instructions then perhaps overall accuracy would even out among the instruments. Perhaps they were rushing through the study and not putting in adequate focus and attention, creating a false effect. Ideally, all participants would be willing to stay for as long as necessary and to work as hard as needed to achieve ideal accuracy of responses.

Much future research remains to be done by psychologists probing the perception of timbre, especially as it relates to computer created blends and attack manipulations. New studies should utilize large, diverse populations of musicians and attempt to whittle down the time required for accurate completion of the task. Perhaps uploading the program to the internet and allowing individuals to participate remotely would be an effective mode of replicating the current study. Additionally, the methodology of the present study can be extended to other applications, such as the role of pitch in timbre perception or the impact of melody on timbre identification. This could be easily achieved by switching the A440 tone for various pitches and melodies and would allow extensive practical application of results.

In total, the present study concludes that instrument type plays a central role in determining the importance of blend components and attack presence when identifying timbres. Some blends are homogenous and difficult to perceptually pick apart, while some instruments pop out so strongly that they mask out all others. Composers and musicians should consider these results when arranging musical pieces to attain the greatest audience reaction. In future research larger sample sizes, uniform instrument sound samples, and various stimuli types will allow concrete conclusions to be drawn with fewer limitations than the current experiment.

References

Caclin, A., Brattico, E., Tervaniemi, M., Natanen, R., Morlet, D., Giard, M. H., et al. (2006). Separate neural processing of timbre dimensions in auditory sensory memory. *Journal of Cognitive Neuroscience, 18* (12), 1959–1972.

Caclin, A., Giard, M. H., Smith, B. K., & McAdams, S. (2007). Interactive processing of timbre dimensions: A Garner interference study. *Journal of Brain Research, 1138*, 159–170.

Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America, 118* (1), 471–482.

Caetano, M., Rodet, X. (2010). Automatic Timbral Morphing of Musical instrument Sounds by High Level Descriptors. In Proc. *International Computer Music Association*.

Chowning, J. M. (1973). The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society, 21*, 526–534.

Ezzat, T., Meyers, E., Glass, J., and Poggio, T (2005). Morphing Spectral Envelopes Using Audio Flow. In Proc. *International Conference on Acoustics, Speech, and Signal Processing , 2005*, 1029-1032.

Fitz, K., Haken, L., and Christensen, P. (2007). Beyond tradition sampling syntehsis: real-time timbre morphing using additive synthesis. *Analysis, Synthesis, and Perception of Musical Sounds*, 122-144.

Goad, Pamela J. and Keefe, Douglas H.(1992). Timbre discrimination of musical instruments in a concert hall. *Music Perception: An Interdisciplinary Journal, 10* (1), 43-62.

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

Grey, J. M., and Gordon, J. W. (1978). Perceptual effects of spectral modifications of musical timbres. *Journal of the Acoustical Society of America*, *63*, 1493-1500.

Grey, J. M., and Moorer, J. A. (1977). Perceptual evaluations of synthesized musical instrument tones. *Journal of the Acoustical Society of America, 62*, 454-462.

Guclu, Burak, Sevinc, Emre, and Canbeyli, Resit (2011). Duration Discrimination by Musicians and Nonmusicians. *Psychological Reports*, *108*, 675-687.

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., and C. L. Krumhansl. (1993). Isolating the Dynamic Attributes of Musical Timbre. *Journal of the Acoustical Society of America, 94*, 2595-2603.

Karplus, K. and Strong, A. (1983). Digital Synthesis of Plucked-String and Drum Timbres. *Computer Music Journal, 7*(2), 43-55.

Kendall, Roger A. and Carterette, Edward C.(1991). Perceptual Scaling of Simultaneous Wind Instrument Timbres. Music Perception: An Interdisciplinary Journal, 8(4), 369-404.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 739–751.

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities and latent subject classes. *Psychological Research, 58*, 177–192.

Paul, Phyllis M. (2005). The Effect of Manipulating Sound Envelope Components on Identifying Wind Instrument Timbre. Bulletin *of the Council for Research in Music Education, 166*, 17-26.

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 976-986.

Roads, C. 1978. Automated Granular Synthesis of Sound. Computer Music Journal 2(2), 61-62.

Seol J, Oh M, Kim JS, Jin S-H, Kim SI, et al. (2011) Discrimination of Timbre in Early Auditory Responses of the Human Brain. PLoS ONE 6(9): e24959

Steele, Kenneth M. and Williams, Amber K.(2006). Is the bandwidth for timbre invariance only one octave?, *Music Perception: An Interdisciplinary Journal , 23*(3), 215-220.

Tardieu, D. and McAdams, S. (2012). Perception of dyads of impulsive and sustained instrument sounds. *Music Perception, 30* (2), 117-128.

Vurma, A., Raju, M. and Kuuda, A. (2010). Does timbre affect pitch?: Estimations by musicians and non-musicians. *Psychology of Music, 39* (3), 291-306.

Table 1

Pairwise Comparisons of Instrument Blend Types

(I) Blend	(J) Blend	*M* (I-J)	Std. Error	*p*	95% Confidence Interval Lower	95% Confidence Interval Upper
Cello-Violin	Cello-Clarinet	18.862*	3.044	.000	8.513	29.210
	Cello-Piano	17.500*	3.286	.000	6.330	28.669
	Cello-Sax	16.762*	2.795	.000	7.261	26.263
	Clarinet-Piano	20.134*	3.249	.000	9.091	31.177
	Clarinet-Sax	8.978*	2.450	.021	.649	17.307
	Clarinet-Violin	18.016*	2.943	.000	8.012	28.020
	Piano-Sax	19.739*	3.299	.000	8.524	30.955
	Piano-Violin	19.458*	3.180	.000	8.647	30.268
	Sax-Violin	18.570*	3.061	.000	8.163	28.976
Clarinet-Sax	Cello-Clarinet	9.884*	2.009	.000	3.055	16.712
	Cello-Piano	8.522*	2.461	.041	.155	16.888
	Cello-Sax	7.784*	1.597	.000	2.355	13.213
	Cello-Violin	-8.978*	2.450	.021	-17.307	-.649
	Clarinet-Piano	11.156*	2.192	.000	3.705	18.607
	Clarinet-Violin	9.038*	1.979	.001	2.310	15.767
	Piano-Sax	10.761*	1.847	.000	4.481	17.041
	Piano-Violin	10.480*	1.828	.000	4.265	16.694
	Sax-Violin	9.592*	1.608	.000	4.124	15.060

The mean difference is significant at the .05 level.*

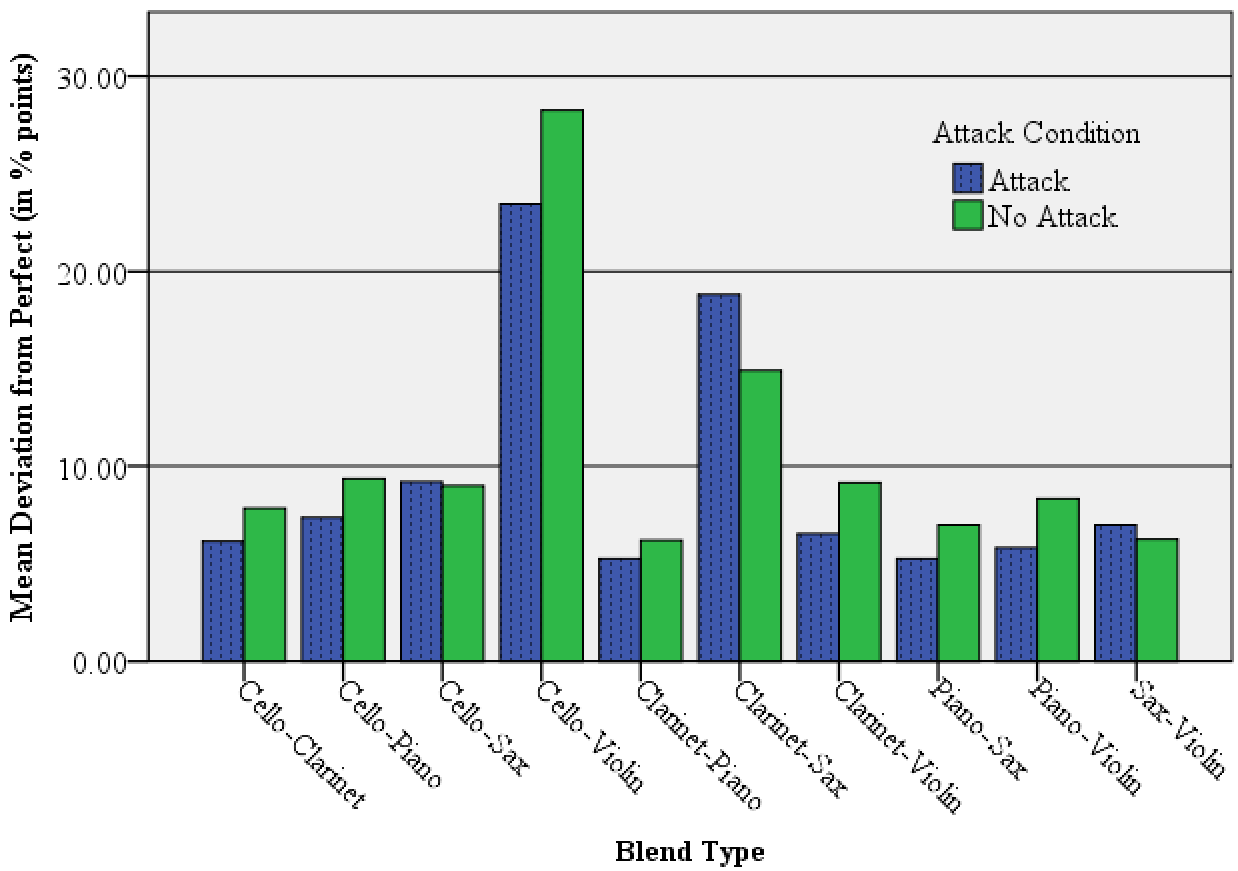

Figure 1: Mean deviation from perfect instrument identification, as a function of blend type and attack condition. This figure illustrates the differences in the mean deviations for each blend type for both attack and no attack conditions.