# Bowed Plates and Blown Strings: Odd Combinations of Excitation Methods and Resonance Structures Impact Perception

*Erica Y. Huynh*

Master of Arts

Music Technology Area
Department of Music Research
Schulich School of Music
McGill University
Montreal, Canada

April 2019

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Arts

# Abstract

The informational value of timbre is transmitted by the mechanical properties producing a sound source. These mechanical properties apply to any object set into vibration, such as a musical instrument. The two mechanical properties of musical instruments of interest are the excitation method and the resonance structure. They are closely related given that the excitation method sets into vibration the resonance structure, which acts as a filter that amplifies, suppresses, and radiates sound components. We used Modalys, a digital physical modeling platform, to synthesize stimuli that simulate three excitation methods (bowing, blowing, striking) and three resonance structures (string, air column, plate), without the resulting sound necessarily being perceived as an existing musical instrument. We paired each excitation method with each resonance structure to produce nine excitation-resonator interactions. These interactions were either typical of acoustic musical instruments (e.g., bowed string, blown air column) or atypical (e.g., blown plate, struck air column). In two experiments, one group of listeners rated the extent to which the stimuli resembled bowing, blowing, or striking excitations (Experiment 1) and a second group rated the extent to which they resembled string, air column, and plate resonators (Experiment 2). Generally, listeners assigned the highest resemblance ratings to: (1) the excitations that actually produced the sound and (2) the resonators that actually produced the sound. These effects were strongest for stimuli that represented typical excitation-resonator interactions. However, listeners confused different excitations or resonators for one another when they heard the stimuli that represented atypical interactions. We address how perceptual data can inform physical modeling approaches, given that Modalys effectively conveyed excitations and resonators of typical interactions but not atypical interactions. Our findings emphasize that our mental models for how musical instruments are played are very specific and limited to what we perceive in the physical world. We can then infer how novel sounds in the daily environment are incorporated into our mental models.

# Résumé

Les informations relevant du timbre sont transmises par les propriétés mécaniques d'une source sonore. Ces propriétés mécaniques s'appliquent à tout objet mis en vibration, tel qu'un instrument de musique. Les deux principales propriétés mécaniques des instruments de musique pertinentes à ce projet sont la méthode d'excitation et la structure de résonance. Celles-ci sont étroitement liées étant donné que la méthode d'excitation met en vibration la structure de résonance, ce qui agit comme un filtre qui amplifie, supprime et rayonne les composants sonores. Nous avons utilisé Modalys, une plate-forme numérique de modélisation physique, afin de synthétiser des stimulus simulant trois méthodes d'excitation (frotter, souffler, frapper) et structures de résonance (corde, colonne d'air, plaque), sans que le résultat sonore ne soit nécessairement perçu comme un instrument de musique existant. Nous avons couplé chaque méthode d'excitation avec chaque structure de résonance pour produire neuf interactions excitation-résonateur. Ces interactions étaient soit typiques des instruments de musique acoustiques (ex. cordes frottées, colonne d'air soufflé) ou atypiques (ex. plaque soufflée, colonne d'air frappée). Dans deux expériences, un groupe d'auditeurs a évalué dans quelle mesure les stimulus ressemblaient à des excitations provoquées par un frottement, un soufflement ou un frappement (Expérience 1) et un second groupe a évalué dans quelle mesure les stimulus simulant la structure de résonance ressemblaient à une corde, à une colonne d'air ou à une plaque (Expérience 2). Globalement, les participants ont attribué les plus hautes valeurs aux : (1) méthodes d'excitation qui ont été utilisées pour produire le son et (2) structures de résonnance utilisées pour produire le son. Ces effets étaient les plus marqués pour les stimulus qui représentaient des interactions typiques excitation-résonateur. Cependant, les participants ont confondu certaines méthodes d'excitation ou structures de résonance lorsque celles-ci représentaient des combinaisons atypiques. Nous abordons la manière dont les informations perceptuelles peuvent informer les approches de modélisation physique, étant donné que Modalys a efficacement synthétisé les méthodes d'excitation et les structures de résonance pour les interactions types mais pas pour les interactions atypiques. Nos résultats soulignent que nos modèles mentaux de la manière dont les instruments de musique sont joués sont très spécifiques et limitées à ce que nous percevons dans le monde physique. Nous pouvons alors déduire comment de nouveaux sons de l'environnement quotidien sont incorporés dans nos modèles mentaux.

# Acknowledgements

I cannot explain how grateful I am to the wonderful individuals who have supported and guided me through this work. I owe the utmost appreciation to Stephen McAdams, who has supervised my thesis over the past two years. Before starting my Master's degree, I had little to no knowledge in physical modeling. Stephen believed that I would be able to accomplish this thesis, even when there were times (and there were many) I doubted myself. It has been an absolute pleasure to work with someone who inspires and challenges me, and I look forward to continuing our research during my PhD. I would like to thank Joël Bensoam, our collaborator from IRCAM and our Modalys guru. The stimuli of our experiments would not exist without him. I am very grateful for the time and patience he has dedicated to helping a clueless student with sound synthesis via physical modeling approaches. I hope we get to meet in person one day, so I can express my gratitude in person.

Bennett Smith, where do I begin? First of all, I cannot thank him enough for the countless hours he has dedicated to programming my experiments and helping me set them up at CIRMMT, even while I was away. His ability to be flexible and understanding is something I will always appreciate. Secondly, I am thankful for the great conversations we have shared over the past semester. I have always looked forward to them and they have certainly lightened the mood, especially during the most stressful times. Juanita Marchand Knight has been nothing short of inspirational to me. I will always appreciate our discussions, their encouragement, their support, and their perspectives on everything. Although I have not known Juanita for long, after every conversation we have had, I feel that they understand and empathize with me, and I cannot say the same for people I have known for years. My labmates (Iza Korsmit, Aurélien Antoine, Lena Heng, Behrad Madahi, Jason Noble, Kit Soden, Jina Wang, Moe Touizrar, Eddy Kazazis, Julie Delisle, Max Henry, Rachel Hottle, Yuval Adler, Tanor Bonin, Manda Fischer, and Etienne Thoret) have been nothing but supportive—academically and morally. I am lucky to be in a lab with such encouraging members. Although we all have different backgrounds, our discussions are nevertheless enriched with knowledge from an abundance of perspectives. Every single one of them is amazing and I cannot wait to see what the future holds for them.

Lastly, I would like to thank my family and friends for always being there for me, even when I have to explain what I am researching for the millionth time! To my wonderful parents, Linda Chan and Tieu Binh Huynh, words will never express how thankful I truly am. My parents have shown me nothing but unconditional love and I can only hope that they are proud of me. Not only have they raised me, but they have supported me through piano lessons; bought me a piano and a clarinet; supported my decision to study psychology and neuroscience; supported my choice to complete a Master's degree away from home; sat through endless phone conversations with me during the most stressful times; and worried sick about me during those stressful times. For all those reasons and more, I would like to dedicate my thesis to my lovely, selfless parents. Consider this a small token of my appreciation.

# Author Contributions

Under the supervision of Professor Stephen McAdams, I was responsible for coming up with the research question, designing the stimuli, designing the experiments, running the experiments, and analyzing and interpreting the data. Joël Bensoam from the Institut de Recherche et Coordination Acoustique/Musique (IRCAM) collaborated with me on the synthesis of the stimuli. Bennett Smith programmed the experiments and helped me prepare them.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

The informational value of sound is crucial for source identification. In particular, musical instruments can be identified by the informational value that emerges from the perceived timbre of a sound. Timbre refers to a plethora of auditory attributes that bear perceptually useful information about the mechanical properties of sound sources (Giordano & McAdams, 2010; McAdams & Goodchild, 2017). A large body of literature has investigated how timbre perception correlates with the acoustical features of a sound wave (Grey & Gordon, 1978; McAdams, Winsberg, Donnadieu, De Soete, & Krimphoff, 1995; Lakatos, 2000; Caclin McAdams, Smith, & Winsberg, 2005). These acoustical features do not arise spontaneously: they originate from sound sources, which have mechanical properties. It is a common misunderstanding that mechanical properties are secondary to acoustical features with respect to sound source identification, when it is in fact the primary cause of natural sound-producing events, such as those arising from musical instruments (Giordano & McAdams, 2010). Two mechanical properties of musical instruments were of interest to the current study: resonance structure and excitation method. The resonance structure is an object that vibrates and radiates sound, and the excitation method sets the resonance structure into vibration. The current study examined how interactions between excitation methods and resonance structures influence how these mechanical properties are perceived. We employed physical modeling techniques to synthesize the sounds and conducted two experiments to collect perceptual judgements based on these stimuli. This chapter reviews the framework of the relationship between sound source mechanics and timbre perception of musical instruments and impacted materials.

## 1.1 The timbre of musical instruments

Timbre has been defined as the "attribute of auditory sensation which enables a listener to judge that two nonidentical sounds, similarly presented and having the same loudness and pitch, are disimilar [sic]" by the American National Standards Institute (ANSI S1.1-1994, p. 34). This definition is misleading, as it tells us very little about timbre and describes it as what it is not, rather than what it is. McAdams (2013) explained that timbre is composed of two characteristics. The first is that timbre carries a plethora of auditory attributes that either continuously vary over the duration of a sound (e.g., brightness, attack sharpness, nasality, etc.) or are categorical (i.e., characteristic of a specific sound). The second characteristic describes timbre as the driving force of sound source identification—allowing listeners to classify sounding objects into absolute categories (McAdams, 2013).

As source identification is an automatic process, listeners have the tendency to assign a label to everyday sounds. In the case of musical instruments, when listeners hear a tone, they have a tendency to identify the instrument that produced it. It is often assumed that each musical instrument possesses its own distinct timbre. For example, individuals may claim that all notes that can be produced by a violin all share a "violin timbre." This is a common misuse of the term timbre. As described by McAdams and Goodchild (2017): "a specific clarinet played with a given fingering (pitch) at a given playing effort (dynamic) with a particular articulation and embouchure configuration produces a note that has a distinct timbre." (p. 129) Changing any of these parameters will ultimately change the timbre.

If a musical instrument produces a tone at one pitch, a different timbre can be perceived if the same instrument plays a tone at a different pitch (Krumhansl & Iverson, 1992; Handel & Erickson, 2001 2004; Marozeau, de Chevigné, McAdams, & Winsberg, 2003). Experiments by Krumhansl and Iverson (1992) explored whether pitch and timbre are perceived independently from one another. The researchers found that the timbres of musical instruments were more easily identified when there were less variances in pitch. In other words, it was easier to distinguish between timbres when sounds were produced by the same pitch; however, if two tones were produced by different pitches, it was more difficult to say whether they were produced by the same instrument.

The timbre of a musical instrument can also change when the dynamics are varied (McAdams, 2013). Previous research demonstrated the influence of dynamics on timbre perception for synthesized sounds (Haack, 1975; Melara & Marks, 1990); no studies to date have examined the influence of dynamics on timbre perception of musical instruments. Previous studies, however, examined the influence of timbre on dynamics (Fabiani & Friberg, 2011). The identification of dynamics for five musical instruments (clarinet, flute, piano, trumpet, and violin) was influenced by different sound levels and differences in timbre that resulted from the instruments being played at different dynamics. McAdams and Goodchild (2017) explained that if a French horn plays a note at softer dynamics, the resulting timbre can be perceived as darker. Consequently, we can speculate that we perceive different timbres when an instrument plays different dynamics.

Timbres produced by musical instruments are also affected by gestures that produce the sound. The physical properties of musical instruments constrain timbral features, but a variety of timbres can emerge from a single musical instrument. Halmrast, Guettler, Bader, and Godøy (2010) distinguished three methods of sound production: moving air directly, moving strings, and striking plates or membranes. The specific control of certain parameters leads to different gestures within each method of sound production. For example, when strings are moved by a bow, changes in the speed of the bow may change the timbre of the resulting sound. Changes in breath pressure can lead to timbral changes in the movement of air. Moreover, changes in the force of a hammer can produce a variety of sounds when applied to a plate or membrane. When sound is produced on a single instrument, several parameters are varied. The interactions between changes in these parameters lead to alterations in the timbre that can be anywhere from subtle to drastic (Halmrast et al., 2010). It is better to think of each musical instrument as having a constrained universe of timbres (McAdams & Goodchild, 2017). Thus, the timbre of a sound can guide the identification of a musical instrument, as long as the timbre fits within the instrument's universe of timbres.

## 1.2 Identification of musical instruments

Although a single musical instrument can create a variety of timbres by manipulating some parameters (e.g., pitch, dynamics, articulation, etc.), listeners can generally identify musical instruments well from hearing single tones. If they are unable to identify a musical instrument, they most often name another instrument that belongs to the same family (van Dinther & Patterson, 2006; Giordano & McAdams, 2010). Moreover, some musical instruments are more easily identified than others (Saldanha & Corso, 1964). Berger (1964) recorded tones of wind instruments (flute, oboe, clarinet, cornet, trumpet, alto saxophone, tenor saxophone, bassoon, French horn, and baritone) and created other versions of the tones by playing them backwards, removing the rise and decay portions, or processing them through a low-pass filter. Listeners identified the wind instruments best with unaltered recordings and next best with backward recordings (Berger, 1964). Poorer performance was associated with removal of rise and decay portions and filtered recordings. However, even with unaltered recordings of the wind instruments, identification performance depended on the instrument: it was easier for listeners to identify the oboe, clarinet, cornet, and tenor saxophone compared to other instruments (Berger, 1964). Elliott (1975) also found that identification of musical instruments—not just limited to winds—was more difficult when the attacks and releases of recorded sounds were removed. Identification performance for unaltered recordings was much better, with the exception of the cello, which was commonly mistaken for the violin (Elliott, 1975). These findings further highlight the influence of attack and release portions on the identification of musical instruments.

Srinivasan, Sullivan, and Fujinaga (2002) examined if experience with playing orchestral instruments affected musical instrument identification. The researchers tested identification performance without practice and following a short training session. Identification performance did not depend on whether there was a training session or not. However, listeners who played orchestral instruments performed better than listeners who did not. These findings suggest that long-term experience impacted listeners' mental models of musical instruments more so than short-term training. In support of Giordano and McAdams' (2010) analyses, identification of musical instruments was better for instruments that belong to different families and confusion was more common for instruments within the same family (Srinivasan et al., 2002).

## 1.3 Dissimilarity judgements of musical tones

Dissimilarity judgements involve rating pairs of sounds according to how dissimilar they are perceived to be (McAdams, 2013). The ratings are then analyzed with multidimensional scaling (MDS), which assigns distances between sounds in a model of timbre space. There is no limit as to how many dimensions an MDS model can have. Sounds judged as similar will appear closer to each other, and sounds judged as dissimilar will appear further apart in the MDS model (McAdams, 2013). The dimensions are then correlated with audio descriptors, which are acoustic parameters that can explain a portion of the variance of judgements along a given dimension (Giordano & McAdams, 2009; Peeters, Giordano, Susini, Misdariis, & McAdams, 2011; McAdams, 2013). Dissimilarity judgements are most commonly reported to be based on two audio descriptors: spectral centroid, which describes the centre of gravity of the energy distribution across frequencies and distinguishes nasal from bright sounds; and the logarithm of the attack time, which separates continuous from impulsive sounds (Grey & Gordon, 1978; McAdams, Winsberg, Donnadieu, De Soete, & Krimphoff, 1995; Lakatos, 2000; Caclin, McAdams, Smith, & Winsberg, 2005). The contribution of other audio descriptors to previous studies is less clear and varies across studies.

Grey's (1977) study involved MDS of 16 synthesized tones that were based on analyses of musical instrument tones. The MDS model had three dimensions. The first dimension was the spectral energy distribution, which corresponds to the bandwidth of the spectrum and where in the spectrum the energy is more concentrated. The second dimension was the fluctuation of the spectrum over time, which describes the synchronicity of the harmonics over time. The third dimension described the presence of low-amplitude, high-frequency energy at the attack portion. Grey (1977) proposed that the second and third dimensions might explain clustering of instrument families, distinguishing brass, woodwind, and string instruments from one another (Grey, 1977). Iverson and Krumhansl (1993) examined the role of onsets in similarity judgements of orchestral instrument tones. The tones were either unaltered or modified such that the onsets were removed or only the onsets remained. MDS of tones of just onsets or without onsets resembled MDS of unaltered tones. One dimension distinguished impulsive from continuous sounds and was revealed to be correlated with the amplitude envelope. A second dimension distinguished brighter from duller sounds and was speculated to be correlated with spectral centroid. These results imply that instrument timbres with either similar onset properties or similar spectra are perceived as more similar. Giordano and McAdams' (2010) review of previously published studies suggested a general pattern of dissimilarity judgements for instrumental timbres: tones played either by instruments of the same family or by similar excitation methods were judged as more similar. It is important to note that the properties of the acoustic waveform convey the musical instrument's mechanical properties to the listener. The waveform properties, therefore, have a secondary, communicative role between the mechanical properties (e.g., excitation method and instrument family) and timbre perception.

## 1.4 Previous work: Concentration on impacted materials

Although the primary focus of the current research will be on musical tones, the majority of studies have focused on impacted materials (i.e., nonmusical sounds). A large body of literature has examined listeners' abilities to identify the material of an impacted sound (Lutfi & Oh, 1997; Klatzky, Pai, & Krotkov, 2000; McAdams, Chaigne, & Roussarie, 2004; Giordano & McAdams, 2006; Aramaki, Besson, Kronland-Martinet, & Ystad, 2009; McAdams, Roussarie, Chaigne, & Giordano, 2010); fewer studies have examined the role of multiple actions (Warren & Verbrugge, 1984; Lemaitre & Heller, 2012; Hjortkjær & McAdams, 2016). These studies are important to consider, given that excitation methods are synonymous with actions, and materials are an aspect of resonance structures.

### 1.4.1 Material properties

Sounds produced by different materials are processed differently by listeners. Aramaki et al. (2009) examined event-related potentials (ERP) during the processing of sounds produced by metal, glass, and wood. For the duration of 150 milliseconds (ms) to 700 ms post-stimulus, metal sounds were processed differently from wood and glass sounds. Behavioural data showed that sounds produced by different materials were judged as more similar if they shared similar frequency components and decay contents (Klatzky et al., 2000). For different-sized objects made of plexiglass, wood, metal, and glass, Giordano and McAdams (2006) found that material categorization was best between gross categories of plexiglass-wood and metal-glass; performance was attributed to the damping characteristics that differed between gross categories. Categorization was worse for materials within the same gross categories, but identification depended on frequency content (Giordano & McAdams, 2006). Moreover, when identifying materials within gross categories, performance was based on the size rather than the materials of the sounds. McAdams et al. (2010) found that listeners were able to identify the material of plates made of glass and metal. Identification was independent of the material of the mallet that struck the plates. This demonstrated that listeners based their judgements on acoustical features that were relevant to the perceptual task (McAdams et al., 2010). Again, the acoustical features (e.g., damping) were derived from the mechanical properties that produced the sounds and communicated the timbre to the listener.

### 1.4.2 Action properties

Less common is research on how the actions that produce sounds affect timbre perception, even though listeners are probably more sensitive to actions rather than materials or objects that produce a sound (McAdams, 2019). An exception includes a study by Warren and Verbrugge (1984), who found that listeners had no difficulty distinguishing between breaking and bouncing glass. Moreover, listeners were able to distinguish between different speeds of rolling balls, but determining the speed depended on the size of the ball (Houben, Kohlrausch, & Hermes, 2004).

### 1.4.3 Interaction between actions and materials

More recently, research has focused on the interaction between sound-producing actions and materials. Lemaitre and Heller (2012) investigated whether sound source properties can be categorized from sound-generating events. Four different actions (scraping, rolling, hitting, and bouncing) were applied to cylinders that were made of four different materials (wood, plastic, metal, glass). By analyzing the acoustic correlates of different actions, Lemaitre and Heller identified two gross categories of actions: discrete (or impulsive) and continuous. Hitting and bouncing were classified as discrete actions, while scraping and rolling were classified as continuous actions. The attack time was the acoustic correlate that distinguished discrete from continuous actions. The materials were separated across two gross categories: one included wood and plastic and the other included glass and metal. Several acoustic correlates distinguished materials between gross categories, but distinguishing materials within gross categories was not correlated with any distinct acoustical property. In their first experiment, Lemaitre and Heller asked listeners to rate sounds based on how well they conveyed the materials and actions that produced them. Half of the time, the action or material listeners rated matched the action or material that produced the sound; the other half of the time, they did not. Lemaitre and Heller found that more confusions were made for materials within gross categories, but listeners were quite accurate at rating materials between gross categories. On the other hand, listeners were always accurate at rating actions regardless of gross category membership. The second experiment was an identification task, which also measured reaction times. Action identification was faster and more accurate than material identification. These findings demonstrate the informational value of the mechanical properties of sound sources and a greater sensitivity to actions than materials.

Hjortkjær and McAdams (2016) employed a similar stimulus design as that of Lemaitre and Heller (2012). The stimuli combined three actions (drop, rattle, and strike) with plates made of three materials (wood, metal, and glass). Participants in their study rated the dissimilarity of the stimuli. Dissimilarity ratings were explained by two dimensions, as revealed by MDS (Hjortkjær & McAdams, 2016). Dimension 1 separated materials (wood versus metal and glass) and was correlated with changes in spectral centroids. Spectral centroid is the centre of gravity of the energy distribution across frequencies and distinguishes bright from dull sounds. Dimension 2 separated the three actions and was correlated with variability in temporal centroids. Temporal centroid is the centre of gravity of the energy distribution across time, which distinguishes impulsive from continuous sounds. Consistent with Lemaitre and Heller's (2012) findings, the distinctions among materials depended on whether they belonged to the same gross category or not. However, distinctions among actions were clear regardless of gross category membership. Hjortkjær and McAdams (2016) also conducted a categorization task, which found that participants were very accurate at identifying both actions and materials that produced the sounds. However, confusions were made for materials with similar spectral centroids and actions with similar temporal centroids. Hjortkjær and McAdams' (2016) findings highlight the influence of sound source mechanics on acoustical properties, which in turn contribute to how the timbre of a sound is perceived.

## 1.5 A mental model of musical sound sources

Previous studies on identification and dissimilarity judgements of sound sources, including musical instruments, have alluded to a "mental model" of sound sources. McAdams and Goodchild (2017) argued that listeners form mental models of sound sources, even when their timbral characteristics vary with changes in pitch, dynamics, and other parameters. A mental model is acquired through exposure: listeners learn how sound sources behave in the physical world by observing or interacting with them. For example, a musician interacts with their instrument on a daily basis, allowing them to understand its techniques and restrictions of sound production. However, from passive exposure, nonmusician listeners can generally understand how a musical instrument is played, but the specific techniques or restrictions of sound production may not be as well understood. Researchers have found that listeners categorized musical sound sources very quickly, suggesting evidence for mental models of musical sound sources (Agus, Suied, Thorpe, & Pressnitzer, 2012). Musicians identified tones played by a single musical instrument across pitch differences of approximately 2.5 octaves, whereas nonmusicians were able to identify the instrument for pitch differences of approximately one octave (Steele & Williams, 2006; Handel & Erickson, 2001). Musicians acquired more exposure to the many timbral variations that a single instrument—especially their own—can produce. They also acquired more exposure to timbral variations if they have played in ensembles or orchestras. Consequently, we can hypothesize that musicians' mental models are stronger for musical sound sources than their nonmusician counterparts. Although the sounds we have synthesized are the same pitch, we are interested in whether musical exposure influences the identification of musical sound sources that vary with respect to other parameters.

## 1.6 The current study

### 1.6.1 Objectives

Previous studies have highlighted the role of sound source mechanics in timbre perception of nonmusical sounds only (i.e., impacted sound sources; Warren & Verbrugge, 1984; Lutfi & Oh, 1997; Klatzy et al., 2000; McAdams et al., 2004; Giordano & McAdams, 2006; Aramaki et al., 2009; McAdams et al., 2010; Lemaitre & Heller, 2012; Hjortkjær & McAdams, 2016). No direct evidence demonstrated that listeners use mechanical properties that produce sounds when processing musical sounds (i.e., sounds produced by musical instruments). A recent review by Giordano and McAdams' (2010) examined data from previously published studies of dissimilarity ratings and categorization of musical sounds. Listeners both rated tones as more similar and classified tones under the same category if they were played by a similar excitation method or instruments of the same family. Thus, the primary goal of this study was to provide empirical evidence for the influence of sound source mechanics on timbre perception.

### 1.6.2 Research questions

In musical instruments, excitation methods set resonance structures into vibration to produce a sound. For acoustic musical instruments, the interaction between excitation methods and resonance structures is very specific. In the case of the air column and plate, they can be blown and struck, respectively; however, not many other excitations can be applied to these resonators in the physical world. The string can be bowed, struck, and plucked; but blowing a string seems to be rarely encountered in normal musical experience. Thus, there are physical limitations for the interactions between excitation methods and resonance structures in acoustic musical instruments. As a result, our mental models of musical instruments are influenced by these physical limitations. The research questions of interest to this study were:

(1) How are excitation methods and resonance structures identified when they are combined in ways that are either typical or atypical of acoustic musical instruments?
(2) Can listeners separate information conveyed by different excitation methods from that of different resonance structures?
(3) Are musicians better at identifying and isolating different excitation methods and resonance structures than nonmusicians?
(4) How can perceptual studies inform physical modeling approaches for the stimuli of the current study?

### 1.6.3 Overview

The current study is composed of one exploratory stimulus design approach and two perceptual experiments. In order to infer how listeners distinguish between different excitation methods and resonance structures, we used physical modeling techniques to synthesize and combine three excitation methods (or excitations; e.g., bowing, blowing, and striking) and three resonance structures (or resonators; e.g., string, air column, and plate); this created a total of nine excitation-resonator interactions. Some of these interactions are considered typical of acoustic musical instruments (e.g., bowed string, blown air column, struck plate, and struck string), whereas other interactions are considered atypical of acoustic musical instruments (e.g. bowed air column, bowed plate, blown string, blown plate, and struck air column). Atypical interactions allowed for excitations and resonators to be freely associated, such that physically impossible sounds become possible with physically inspired modeling (Ystad, Aramaki, & Kronland-Martinet, 2019). Typical and atypical excitation-resonator interactions were generated with Modalys, a digital physical modeling synthesizer that allows for independent control of excitation methods and resonance structures (Dudas, 2014). To model the resonance structures, we took into consideration the assumption that a string that is excited at a short distance from the bridge and a conical air column can be modeled similarly. The string and air column modeled by Modalys employs completely different models. The string we used was fixed at both its ends. The air column was cylindrical (not conical), open at one end, and closed at the other. Consequently, this should guarantee that these two resonatros will sound different, even when the same excitation is applied to them. We used an exhaustive and controlled approach to synthesize the nine interactions. Given that the atypical interactions are rarely encountered in typical musical experiences, it was difficult to anticipate how they would sound. Moreover, physically inspired modeling of atypical interactions is quite uncommon (for notable exceptions, see Böttcher, Gelineck, & Serafin, 2007 for musical sounds; and Conan et al., 2014 for continuously impacted objects). Consequently, the exhaustive approach allowed us to choose from a variety of timbres that could be perceived as conveying the excitation method and resonance structure that produced the sounds. Following the stimulus design, we conducted two experiments. The first experiment asked participants to rate sounds based on how well they resembled three excitations (i.e., actions): bowing, blowing, and striking. The second experiment asked participants to rate sounds based on how well they resembled three resonators (i.e., objects): string, air column, and plate.

### 1.6.4 Hypotheses

Our first hypothesis is that mental models are stronger for typical excitation-resonator interactions than for atypical interactions. We expected the ratings to reflect this, such that they will be highest for excitations and resonators that actually produced them and lowest for excitations and resonators that did not produce them. For example, if a bowed string was heard, listeners will rate it very high as sounding like a bowing excitation and a string resonator, but very low as sounding like blowing and striking excitations and air column and plate resonators. Since the bowed string represents a typical excitation-resonator interaction, the resemblance to certain mechanical properties will be quite obvious. For atypical stimuli, it will depend on whether listeners can isolate the excitation methods from the resonance structure. This might be difficult and require additional processing (McAdams & Goodchild, 2017), since these interactions are not common in everyday listening.

The second hypothesis is that musicians might have stronger mental models for mechanical properties of musical instruments than nonmusicians. We might see that musician participants will have higher ratings for the resemblance of excitations and resonators of typical interactions. Since musicians have extensive experience playing their own instrument, and some musicians have experience playing in orchestras, they might rate a bowed string higher on bowing and string than nonmusicians, for example. However, musicians' mental models for typical excitation-resonator interactions may limit them when they are exposed to the atypical interactions. They may not perceive the excitations and resonators that produced the atypical interactions because they are so used to hearing the typical interactions.

Our third hypothesis is that ratings will be highest for the actual excitation methods and resonance structures that produced the sound. This hypothesis, however, depends on the efficacy of our parameter manipulations through Modalys. If Modalys can convey the excitation methods well to the listeners, then bowing, blowing, and striking ratings should be highest when bowing, blowing, and striking actually produced the sound, respectively. The type of resonator these excitations are applied to should not matter. Likewise, if resonance structures can be conveyed effectively to listeners, then they should have higher string, air column, and plate ratings for sounds produced by strings, air columns, and plates, respectively; the type of excitation should not matter. However, the effect of musical background might be unpredictable, given that musicians and nonmusicians perform differently on instrument categorization tasks (Srinivasan et al., 2002), but not dissimilarity judgements.

The fourth hypothesis is that perceived resemblance of sounds to excitation methods will be influenced by the resonance structures producing them. Similarly, when participants rate the resemblance of sounds to resonance structures, the excitation methods will have an impact. For example, if participants rate how well a blown string resembles blowing, the rating might be lower than for a blown air column, even if both sounds are produced by blowing. A blowing excitation commonly interacts with an air column resonator and our mental models would be stronger for this interaction. If blowing is applied to a string, then the string might distract participants from hearing the sound as fully blown. This would demonstrate the inability to isolate excitation methods from the resonance structures they are typically paired with. In another example, say listeners are presented with a bowed plate. Since bowing is typically applied to a string and plates are typically struck, participants might assign higher ratings of the bowed plate's resemblance to a string resonator and striking excitations, even though neither of these properties produced the sound. Again, this would demonstrate that our mental models can bias the ratings, depending on the type of excitation-resonator interaction.

# Chapter 2

# Physically Inspired Modeling of Paired Excitation Methods and Resonance Structures Using Modalys

The stimuli for these experiments simulated three excitation methods (e.g., bowing, blowing, and striking) and isolated them from the three resonance structures (e.g., string, air column, and plate) that they are typically paired with in acoustic musical instruments. Moreover, we applied each of these excitation methods to each of the resonance structures. Several physical modeling software do not separate excitation methods and resonance structures in a way that deviates from how they typically behave in an acoustic musical instrument (see Conan et al., 2014 for a notable exception with impacted materials). For example, the excitation method of blowing cannot be separated from the resonance structure of an air column. Moreover, blowing cannot be applied to a string. This is because the interaction between blowing and an air column is more physically conceivable than blowing and a string, which is considered an abstract interaction. Moreover, it is more intuitive for physical modeling software developers to simulate very specific interactions in real acoustic musical instruments. For example, Arché simulates the physics of a bow interacting with the strings of the violin, viola, and cello (Expressive E, 2019). Pianoteq implements physical modeling to simulate acoustic and electric pianos as well as chromatic percussions (Modartt, 2019).

The Musical Acoustics Team at the Institute for Research and Coordination in Acoustics/Music (IRCAM) in Paris, France have developed Modalys: a digital physical modeling synthesizer that can simulate different excitation methods and resonance structures without the resulting sound necessarily being perceived as an existing musical instrument (Eckel, Iovino, & Caussé, 1995). Modalys can also combine excitation methods and resonance structures in a way that is more abstract and less typical of acoustic musical instruments. Consequently, we have implemented an exhaustive approach to generate nine different interactions of three excitation methods and three resonance structures. The first goal of this approach was to explore the resulting sounds of the nine interactions under the manipulation of certain parameters. Two selected parameters for each of the nine interactions were manipulated. For each parameter, 20 values were tested, giving a total of 400 (20×20) stimuli for each of the nine interactions. Thus, a grand total of 3,600 stimuli (400 stimuli × 9 interactions) were generated with this approach.

The second goal of this approach was to pick three sounds (i.e., exemplars) for each of the nine interactions. They were used as the stimuli for the experiments in Chapters 3 and 4 and demonstrated the variability of each of the nine interactions that were physically modeled. We also performed an acoustical analysis of the resulting 27 stimuli (3 exemplars × 9 interactions) using the spectrogram function on Matlab (Mathworks, Natick, 2018). Additionally, three audio descriptors of the sounds were analyzed with the Timbre Toolbox (Peeters et al., 2011) and the Music Information Retrieval (MIR) Toolbox (Lartillot, Toiviainen, & Eerola, 2008): the temporal centroid, spectral centroid, and inharmonicity.

## 2.1 About Modalys

Modalys allows the user to operate as a (digital) instrument designer by employing modal synthesis to predict the behaviour of a structure in reaction to an external excitation that is applied to it (Dudas, 2014.). Consequently, the user can assemble an unlimited supply of objects into any type of instrument they desire, as long as the appropriate parameters are specified. Modal synthesis predicts the acoustical outcome of an interaction between an *exciter* and a *resonator*. The exciter refers to the source of energy: it causes the excitation. This source of energy is applied to a resonator, which is the resonance structure or vibrating object. Consequently, the sounds produced by different interactions between exciters and resonators can be estimated by Modalys; these interactions can be typical or atypical of acoustic musical instruments. Typical interactions are more physically accurate and formed with true physical modeling. Our mental models are likely stronger and more specific for these interactions, since they are more common in our everyday environment. On the other hand, atypical interactions are more abstract and formed with physically inspired modeling. These interactions are estimated by isolating an excitation from the resonator it typically interacts with and applying it to one that it does not typically interact with. Properties of an excitation are estimated by solving a time equation that predicts the temporal evolution corresponding to its movement. Properties of a resonator are estimated by computing the modes that would be present during vibration. In the case of a bowed string, Modalys estimates the properties of bowing (e.g., speed, pressure, etc.) when it is applied to a string as accurately to be as possible. The properties of the string (e.g., size, length, radius, etc.) must also be estimated as physically accurate as possible. Once the interaction is successful (i.e., the interaction sounds realistic), it is then possible to isolate the bowing excitation and apply it to another resonator, such as a plate. In turn, the properties of the plate must also be estimated to be as physically accurate as possible, likely from a striking excitation. These estimations ensure that: (1) the resonator will in fact project a sound that is as physically accurate as possible once excited and (2) the excitations will actually produce a sound once it comes in contact with the resonator. With these estimations, an excitation and resonator can be isolated from one another; then, we can estimate their interactions with other resonators and excitations.

Modalys employs a linear model to estimate the wave propagation. Linear solutions of the wave equation are written using a Green operator. Each point $i$, after discretization of the velocity (or displacement) $u$, is a sum of two terms described by the following equation (Bensoam & Roze, 2013):

$$u_i(t_i) = \tilde{u}_i(t_i)_{\to 0} + \sum_j \mathbb{G}_{ij} f_j(t_i) \qquad (1)$$

This equation computes the temporal evolution of an excitation by representing the velocity vector fields as a function of applied forces (Bensoam & Roze, 2013). The first term corresponds to the velocity (or displacement) that would occur if the system was free to vibrate at time *t*. The second term considers an external force applied at point $j$. Thus, we can obtain discrete, instantaneous representations of time $t$ at point $i$ by evaluating the state of the system $\tilde{u}_i(t_i)_{\to 0}$. This equation can also be broken down into a system of equations that represent the propagation (of the resonator) and interaction of the model:

$$\begin{cases} u_i(t_i) = \tilde{u}_i(t_i)_{\to 0} + \sum_{j=1}^{m} \mathbb{G}_{ij} f_j(t_i)\,, & i = 1, \ldots, m \\ f_i = C_{(k)}(u_i), & i = 1, \ldots, m \end{cases} \quad (2)$$

where $m$ represents the number of interactions. The first part of the system represents the wave propagation, which is computed in the form of (1). This is accomplished by using modal decomposition. The interaction is represented by $C_{(k)}$ in the second part of the system. They are typically predetermined, so it is mainly modal decomposition that must be incorporated to solve the system (Bensoam & Roze, 2013).

Modal decomposition estimates the dynamic of each mode in terms of its numerical simulation and is represented by:

$$u(x,t) = \sum_{k=1}^{K} u^{[k(t)]} e_{k(x)} \quad (3)$$

where $u(x,t)$ represents the velocity at time $t$ and position $x$. The amplitude of the $k$th mode at time $t$ is given by $u^{[k(t)]}$. Once these values are computed for each mode, the contributions of each mode are computed by the Modalys using (1) and accounting for all the interactions given in the model (2). Thus, Modalys incorporates linear equations to compute and represent the spectral (i.e., modal) properties of the resonators and the temporal evolution of the excitations. Since any vibrating object can be represented by modal decomposition, this procedure allows for atypical combinations of excitations and resonators.

## 2.2 Resonator parameters

We used Modalys to synthesize three resonators: string, air column, and plate. For the air column, Modalys simulates the modes that represent the air particles within the air column, even though the object is called a tube in Modalys. To obtain experimental control that would be necessary for the experiments in the next two chapters, the physical parameters of each resonator were kept as consistent as possible. That is, for as many controlled parameters as possible, the same string was used regardless of the exciter that was applied to it. This was also maintained for the air column and plate. We will now discuss the physical parameters that were used to construct each type of resonator. We synthesized each resonator to produce a fundamental frequency of 155 Hz, which corresponds to a pitch of E-flat-3. This was accomplished with the `set-pitch` function of Modalys. Consequently, the parameters chosen for each resonator ensured vibration at this pitch across all exciters that were applied.

The modal properties of the resonators included the number of modes, frequency loss (freq-loss) coefficient, and constant loss (const-loss) coefficient. The number of modes determines how many modes will be simulated: a higher number of modes corresponds to a greater number of frequencies that will be present when an excitation is applied to a resonator (Dudas, 2014). The freq-loss and const-loss coefficients interacted with the frequency of each mode to describe how the resonator loses energy over time. The energy loss of each mode can be described by the following equation, where *f* is the frequency of a particular mode:

$$\text{energy loss} = \text{const\_loss} + \text{freq\_loss}\left(\frac{f}{1000}\right)^2 \qquad (4)$$

When the freq-loss coefficient is nonzero, the higher modes will decay faster than the lower ones. The freq-loss and const-loss values differed depending on the excitation that was applied to the resonator. These values will be specified in Section 2.4 for each interaction. Additionally, we will specify the point of interaction (i.e., access point) between an exciter and resonator in Section 2.4.

### 2.2.1 String

We synthesized a string by using the `'bi-string` and `'mono-string` resonators in Modalys (Fig. 2.1). Both types of strings were a thin wire fixed at the endpoints. The `'bi-string` vibrated in two directions and was used for the bowing excitation. The `'mono-string` vibrated in one direction and was used for blowing and striking excitations. The parameters of both types of strings were manipulated to share as many characteristics as possible. We refer to both these types of strings as one string hereafter. The string's tension, density, radius, and number of modes were set to Modalys' default values (see Table 2.1 for the units of Modalys' parameters): tension = 100 N, density = 1000 kg/m$^3$, radius = 0.001 m, number of modes = 100. Additionally, Young's modulus (i.e., elasticity of the string) was 0.001 Pa.

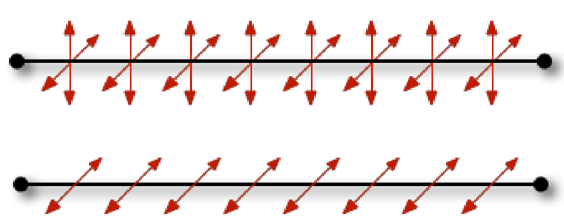

**Figure 2.1.** Diagrams of the `'bi-string` (top) and `'mono-string` (bottom) resonators of Modalys. The diagrams were retrieved from Modalys' documentation (Dudas, 2014).

**Table 2.1.** Units and their abbreviations for the parameters of Modalys.

| Parameter | Unit |
|---|---|
| **Radius** | Meters (m) |
| **Thickness** | Meters (m) |
| **Mass** | Kilograms (kg) |
| **Tension** | Newtons (N) |
| **Density** | Mass per unit volume (kg/m$^3$) |
| **Air density** | Mass per unit volume (kg/m$^3$) |
| **Air elasticity** | Area per unit mass (m$^2$/kg) |
| **Young's modulus** | Pascals (Pa = N/m$^2$) |
| **Poisson coefficient** | Ratio |
| **Stiffness** | Newtons per meter (N/m) |

### 2.2.2 Air column

The air column was synthesized by using the `'closed-open-tube` resonator in Modalys (Fig. 2.2). This resonator was defined as a cylindrical air column that is open at one end and closed at the other end. The air elasticity, air density, and radius of the air column were set to Modalys' default values for the air column: air elasticity = 7.21e-6 m$^2$/kg, air density = 1.2 kg/m$^3$, and radius = 0.01 m. The number of modes was 40, which is the value used in Example 3 ("tube, reed and hole connection"; Appendix 1), provided by Modalys. All Modalys examples that we mention were provided with the installation of Modalys.

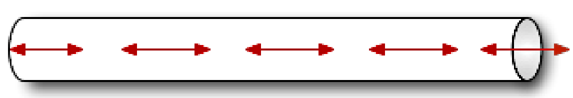

**Figure 2.2.** Diagram of the `'closed-open-tube` resonator of Modalys. The diagram was retrieved from Modalys' documentation (Dudas, 2014).

### 2.2.3 Plate

We synthesized the plate by using the `'rect-plate` resonator in Modalys (Fig. 2.3). The plate was rectangular and fixed at its edges. The plate's thickness, density, Young's modulus, Poisson coefficient, and number of modes were set to Modalys' default values for a plate made of metal: thickness = 0.01 m, density = 7800 kg/m$^3$, Young's modulus = 2e11 Pa, Poisson coefficient = 0.3, and number of modes = 80. The Poisson coefficient describes the degree to which a resonator expands in a direction that is orthogonal to that of compression. The plate's access point corresponded to x and y coordinates on the surface of the rectangular plate, such that the corners have coordinates of (0,0), (0,1), (1,0), and (1,1). These coordinate values were normalized, in length, to the size of the plate. The access point corresponded to x and y coordinates of 0.6 and 0.7, respectively.

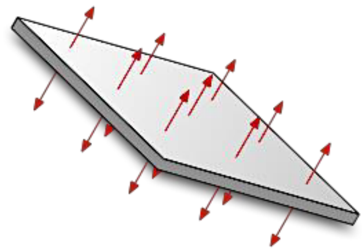

**Figure 2.3.** Diagram of the `'rect-plate` resonator of Modalys. The diagram was retrieved from Modalys' documentation (Dudas, 2014).

## 2.3 Excitation parameters

We synthesized three exciters with Modalys: bow (for bowing excitation), mouth and reed (for blowing excitation), and hammer (for striking excitation). To maintain consistency for the experiments in Chapters 3 and 4, as much as possible, the same temporal envelope of the bowing excitation was applied to each resonator. This was also maintained for blowing and striking excitations. The most variability in the temporal envelope was observed in the blowing excitation, which will be discussed in Section 2.3.2. For each type of excitation, we chose certain parameters to manipulate in an exploratory approach, since they influenced the resulting timbre significantly. These manipulations will be discussed in Section 2.4 for each excitation-resonator interaction.

### 2.3.1 Bowing

We synthesized a bow with Modalys' `'bi-two-mass` object (Fig. 2.4). The bow was made up of two masses that were connected by a spring. One of the masses accessed the resonator along the horizontal direction to control the speed of the bow; this was governed by Coulomb's Law of Friction (McMillan, 1997; Vigué, Vergez, Karkar, & Cochelin, 2017). The other mass accessed the resonator along the vertical direction to control the pressure of the bow (i.e., how much the resonator was displaced; Dudas, 2014); this was governed by the unilateral contact law, which is an example of nonlinear coupling. In the case of contact, a contact force occurs when the relative velocity vanishes; on the other hand, without contact there is no contact force and the relative velocity is no constrained. The two masses of the bow had the same parameters as those specified in Example 6 ("bowed string"; Appendix 2), provided by Modalys: each had a mass of 0.05 kg.

They were connected by a spring with a stiffness of 5e4 N/m in both the horizontal and vertical directions. The freq-loss and const-loss coefficients were 100 and 0, respectively, for both the vertical and horizontal directions of vibration.

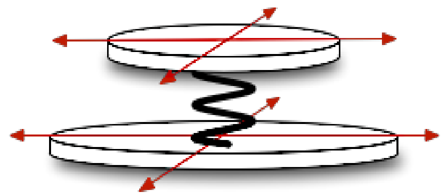

**Figure 2.4.** Diagram of the `'bi-two-mass` object of Modalys. This object was used to model a bow and hammer. The diagram was retrieved from Modalys' documentation (Dudas, 2014).

The temporal envelope for the control of the speed and pressure of the bow was consistent across all resonators to which the bowing excitation was applied. Table 2.2 summarizes the temporal envelope for the speed of the bow, the time values of which were similar to those of Example 6 (Appendix 2) provided by Modalys. The times at which the bow changed its speed along the horizontal direction were kept constant across all resonators to which the bow was applied. However, the maximum speed value was indicated by $X_1$. Manipulating this value greatly impacted the timbre of the resulting sound (as confirmed by Halmrast et al., 2010).

**Table 2.2.** Temporal envelope for the control of the speed of the bow along the horizontal direction.

| Time (s) | Speed (m/s) |
|---|---|
| **0.00** | 1 |
| **1.00** | $X_1$ |
| **9.99** | 1 |

Note: The speed of the bow changed at the listed times. The value of $X_1$ corresponded to the maximum speed and was manipulated, as discussed in Section 2.4.

The temporal envelope for the pressure of the bow is summarized in Table 2.3. The time values of the temporal envelope for the pressure of the bow were modified from those of Example 6 (Appendix 2). We applied the bowing excitation to the string and modified the time values of the temporal envelope for the pressure of the bow to make the bowing sound as realistic as possible. It was simplest to work with the string first, because a bowing excitation is typically applied to a string resonator in acoustic musical instruments. Once we obtained a realistic temporal envelope of the bow's pressure for the string resonator, we applied the same time values to the air column and plate resonators. The times at which the bow changed the displacement of the resonator along the vertical direction were consistent across all resonators to which the bow was applied. The maximum displacement value was indicated by $Y_1$, and it significantly impacts the timbre of the resulting sound (Halmrast et al., 2010).

**Table 2.3.** Temporal envelope for the control of the pressure of the bow along the vertical direction.

| Time (s) | Pressure, measured as displacement (m) |
|---|---|
| **0.00** | 0.01 |
| **0.50** | 0.00 |
| **0.70** | $Y_1$ |
| **1.10** | $Y_1$ |
| **1.70** | 0.01 |
| **5.00** | 0.01 |

Note: The displacement of the resonator changed at the listed times. The maximum displacement, $Y_1$, was manipulated, as discussed in Section 2.4.

### 2.3.2 Blowing

A single reed was synthesized with the `'single-point` object (Fig. 2.5). This object, as defined by Modalys, represents a spring-mass system and vibrates in one direction (Dudas, 2014). The reed was set into vibration by blowing. This vibration resulted in the flow of air, which then set the resonator into vibration. The parameters of the reed included the frequencies (in Hz) assigned to the modes of the vibration of the reed, bandwidths (in Hz) assigned to the loss coefficients of those modes, and the amplitudes of the modes (on a linear scale). The amplitude of the modes was set to a value of 1 and was consistent across all resonators to which the reed is applied. The frequencies and bandwidths were dynamically controlled, such that the values can be changed by sending messages through the script. The bandwidths were set to 50,000 Hz, which was the default value from Example 13 (“simple blow”; Appendix 3) provided by Modalys, and was consistent across all resonators blowing was applied to. The frequencies of the modes, however, were not consistent for all resonators. The reed only worked for a very limited range of frequencies that were assigned to the modes and this depended on the type of resonator. Consequently, the frequencies chosen were values that produced the least amount of sounds that resembled a squeak (i.e., a pitch that does not correspond to a fundamental frequency of 155 Hz). The values of the frequencies that were assigned to the modes of the virtual reed were 2,000 Hz, 1,500 Hz, and 1,000 Hz when the reed was applied to the air column, string, and plate, respectively.

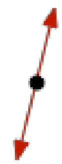

**Figure 2.5.** Diagram of the `'single-point` object of Modalys. This object was used to model a single-reed. The diagram was retrieved from Modalys' documentation.

We controlled the temporal envelope of two parameters: breath pressure and valve-zeta. Valve-zeta refers to the pressure of the lips on the reed and is often referred to as the reed opening parameter (Coyle, Guillemain, Kergomard, & Dalmont, 2015). The reed opening parameter, $\zeta$, is mathematically defined as:

$$\zeta = Z_c w H \sqrt{\frac{2}{\rho p_M}} \qquad (5)$$

where $w$ is the width of the reed and $H$ is the height of the channel; $p_M$ is the closing pressure between the mouth and the mouthpiece, $Z_c$ is the characteristic impedance at the air column's input, and $\rho$ is the air density (Coyle et al., 2015).

The temporal envelope for the breath pressure was kept as constant as possible for the resonators to which the blowing excitation was applied and is summarized in Table 2.4. The time at which the maximum breath pressure, $X_2$, was applied was different depending on the resonator. We obtained the time at which $X_2$ was applied by working with the air column resonator first. It was easiest to apply the blowing excitation to the air column, since this interaction is more typical of acoustic musical instruments. We chose a time value that corresponded to the most realistic blowing sound. Once the time at which $X_2$ occurred was obtained for the air column resonator, we attempted to apply the same time value to the string and plate resonators. However, this proved to be problematic: there was a limited range of time values at which $X_2$ can be applied to the string and plate resonators without the result sounding too abrupt or resembling a squeak. Consequently, for the string and plate resonators, $X_2$ was applied at a time value that resulted in the most realistic blowing sound. Furthermore, in the case of the plate, the higher modes resonated before the lower modes; the time at which $X_2$ was applied minimizes the difference in entry time of the higher and lower modes as much as possible. The difference in entry time of the higher and lower modes was likely attributed to the physical nature of the plate resonator. The times at which $X_2$ is applied to the air column, string, and plate were 0.150 s, 0.110 s, and 0.101 s, respectively. Manipulating $X_2$

substantially changed the resulting sound's timbre (Dalmont, Gilbert, Kergomard, & Ollivier, 2005; Halmrast et al., 2010).

**Table 2.4.** Temporal envelope for the control of the breath pressure.

| Time (s) | Breath pressure (linear scale) |
|---|---|
| **0.00** | 0 |
| **0.150[a], 0.110[b], or 0.101[c]** | $X_2$ |
| **1.70** | 0 |

Note: The breath pressure (measured on a linear scale) changed at the listed times. The maximum breath pressure, $X_2$, was manipulated, as discussed in Section 2.4.
[a] The time at which $X_2$ was applied to the air column resonator.
[b] The time at which $X_2$ was applied to the string resonator.
[c] The time at which $X_2$ was applied to the plate resonator.

The temporal envelope for the control of the reed opening parameter remained the same across all resonators that the blowing excitation was applied to and is summarized in Table 2.5. We added a temporal envelope to control the reed opening parameter so that the blowing excitation would sound more realistic. The time values of the temporal envelope were consistent regardless of the resonator to which the blowing was applied. The maximum reed opening parameter was indicated by $Y_2$. Changing its value substantially influenced the timbre of the resulting sound (Dalmont et al., 2005; Halmrast et al., 2010).

**Table 2.5.** Temporal envelope for the control of the reed opening parameter.

| Time (s) | Reed opening parameter (linear scale) |
|---|---|
| **0.00** | 0 |
| **0.10** | $Y_2$ |
| **2.20** | 0 |

Note: The reed opening parameter (measured on a linear scale) changed at the listed times. The maximum reed opening parameter, $Y_2$, was manipulated, as discussed in Section 2.4.

### 2.3.3 Striking

We synthesized a hammer by creating a `'bi-two-mass` object in Modalys (Fig. 2.4). Like the bow we synthesized in Section 2.3.1, the hammer was made up of two masses connected by a spring. One mass accessed the resonator in the horizontal direction to control the speed of the hammer, and the other one accessed the resonator in the vertical direction to control the pressure of the hammer (i.e., how much the resonator is displaced). Although a `'mono-two-mass` object would also function as a hammer, we followed the "plate and hammer" example that was provided by Modalys (Example 2b; Appendix 4), which used the `'bi-two-mass` object. The two masses of the hammer had the same parameter values as those specified in Example 2b. They are connected by a spring with a stiffness of 1.0e5 N/m in the horizontal direction and a stiffness of 1.5e5 N/m in the vertical direction. In both the horizontal and vertical directions of vibration, the freq-loss coefficient was 100, and the const-loss coefficient was 0.

Struck sounds do not correspond to an auto-oscillating coupling, unlike bowed and blown sounds. Perceptual outcomes of auto-oscillation sounds depend on the manipulation of parameters given to nonlinear coupling. For struck sounds, however, these parameters do not exist. Consequently, we chose to manipulate parameters that affect the resulting timbre significantly, such as the pressure of the hammer and/or the output positions of the sound. The temporal envelope for the control of the pressure of the hammer was consistent for all resonators to which the hammer was applied. This envelope is summarized in Table 2.6. The time values of this envelope were modified from Modalys' Example 2b (Appendix 4). We applied the striking excitation to the plate and modified the time values of the temporal envelope. Once the resulting sound was perceived to be realistic, we applied the same time values of the temporal envelope for the control of the hammer pressure to the string and air column resonators. The maximum displacement was constant at –0.001 m when the hammer is applied to the plate resonator. The value is negative because it refers to the position of the hammer "below" the plate. Because Modalys normalizes the amplitude of the output, changing the maximum pressure did not lead to significant timbral changes in the resulting sound. On the other hand, changing the maximum displacement of the hammer ($Y_3$) on the string and air column did influence the timbre of the resulting sound.

**Table 2.6.** Temporal envelope for the control of the pressure of the hammer on the resonator.

| Time (s) | Pressure, measured as displacement (m) |
|---|---|
| **0.00** | 0.10 |
| **0.05** | $Y_3$[a], -0.001[b] |
| **0.10** | 0.10 |

Note: The displacement of the resonator changed at the listed times. The maximum displacement, $Y_3$, was manipulated for string and air column resonators, as discussed in Section 2.4. The maximum displacement was kept constant for the plate.

[a] The maximum displacement, $Y_3$, was be manipulated for the string and air column resonators.

[b] The maximum displacement was -0.001 m for the plate resonator.

## 2.4 Excitation-resonator interactions: An exhaustive approach of parameter manipulation

We combined each type of resonator with each type of excitation, creating nine different interactions: bowed string, bowed air column, bowed plate, blown string, blown air column, blown plate, struck string, struck air column, and struck plate. We manipulated two parameters for each interaction, which are discussed in the sections below. We tested 20 values for each manipulated parameter, giving a total of 400 stimuli (20 values of one parameter × 20 values of another parameter) for each of the nine interactions. This exhaustive approach was used to see what combinations of parameter values resulted in sounds that are perceived to resemble the excitation and resonator of the interaction. This was especially useful for interactions that are atypical of acoustic musical instruments (e.g., bowed air column, bowed plate, blown string, blown plate, and struck air column). Given that these sounds do not exist in the physical world, it was difficult to decipher how they were supposed to sound. Moreover, if a certain combination of parameter values resulted in perceptually convincing sounds for one type of excitation-resonator interaction, it does not guarantee that the same combination of parameter values would have the same result for other interactions. Consequently, the exhaustive approach allowed us to uncover many timbral possibilities of the resulting sounds. The exhaustive approach we employed was also adapted from that of Dalmont et al. (2005), who used it to determine the oscillation and extinction regions of a clarinet. They manipulated the values of two parameters of the clarinet and plotted the outcomes of sounds from these manipulations. We borrowed this approach for our stimuli by manipulating parameters relevant to each excitation.

For each of the nine excitation-resonator interactions, the authors informally recorded the perceptual outcomes of the 400 sounds. First, whether or not there was an output was noted. Next, how well the output represented the excitation and resonator that produced them was recorded according to the categories in Figures 2.8, 2.9, 2.10, 2.13, 2.14, 2.15, 2.17, 2.19, and 2.20. Lastly, for each interaction type, three exemplars that were most effective in conveying their excitation and resonator were selected, leading to a total of 27 stimuli. We decided to choose three sounds so that the experimental task would not be too demanding or lengthy. The three chosen sounds for each interaction type were perceived to be the most variable among the sounds that also conveyed the excitation and resonator that produced them. They were also perceived to be produced by the same musical instrument. For example, a performer can play a single note of a musical instrument three different ways, and there will be variability in the timbre between each of the produced sounds depending on the exact values for the performance parameters (McAdams & Goodchild, 2017). The 27 stimuli are accessible at http://132.206.14.109/~mpcl/SoundSourceStimuli/. They will be analyzed for three audio descriptors—temporal centroid, spectral centroid, and inharmonicity—in Section 2.5.

### 2.4.1 Bowed string

The bowing excitation was applied to the string resonator at a unit of 0.1, which was normalized to its length. The output of the sound was measured and recorded at a normalized unit of 0.6 along the string. The freq-loss coefficient and const-loss coefficient of the string resonator were both 1.0.

Two parameters were manipulated for the interaction of the bowing excitation and the string resonator: the speed of the bow and displacement of the string under its resting position (i.e., bow pressure). The displacement is negative because the bow pushes down on the string. We manipulated the values for the maximum speed ($X_1$) and maximum pressure ($Y_1$) of the bow, because these parameters significantly impact the perceived timbre of the resulting sound. We used 20 values of $X_1$ between 1 m/s and 20 m/s, with increments of 1 m/s (Fig. 2.6). These were the limits of the bow speed: inputting values outside of this range resulted in sounds that did not resemble a bowed string, given the values of all of the other parameters we summarized previously. The 20 values we tested for $Y_1$ were between 0.1 mm and 2.0 mm, with increments of 0.1 mm (Fig. 2.7). These values were within the limits of the bow pressure, given the values of the other parameters we mentioned previously . Holding the values of all the other parameters constant, we plotted the reported outcomes as we perceived them for each pairing of $X_1$ and $Y_1$ for the bowed string in Figure 2.8. The values of $X_1$ and $Y_1$ for the three sounds we chose are indicated with red circles.

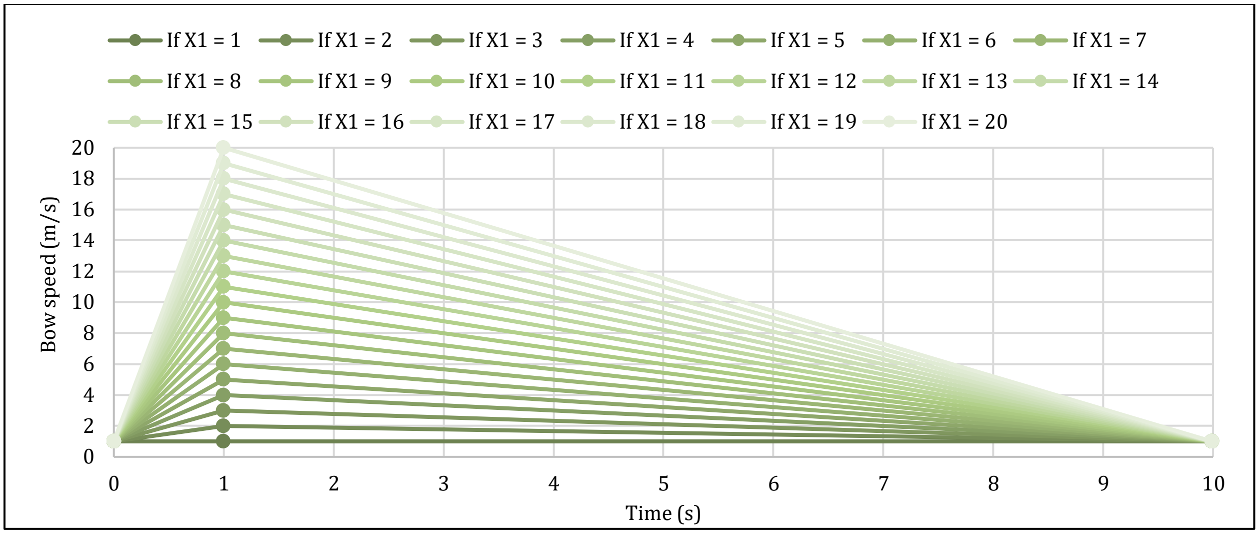


**Figure 2.6.** Twenty possible temporal envelopes for the control of the bow speed, depending on the value of $X_1$ (maximum bow speed). See Table 2.2 for the time values of the temporal envelope.

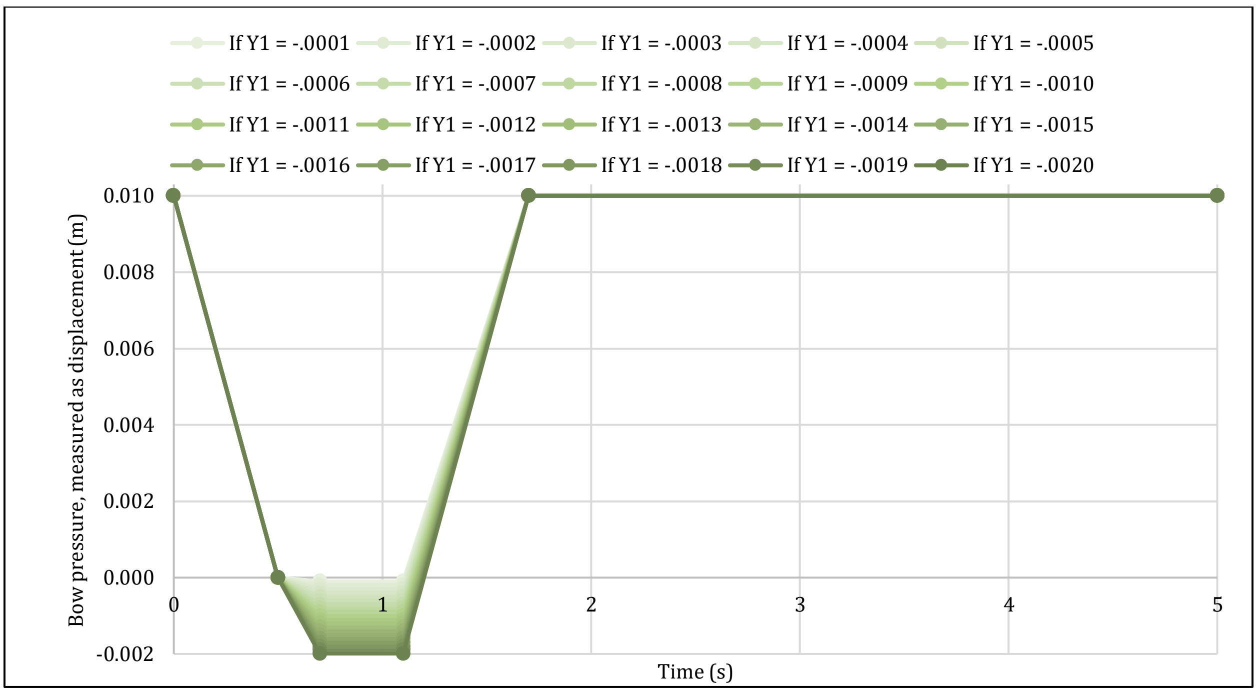


**Figure 2.7.** Twenty possible temporal envelopes for the control of the bow pressure, depending on the value of $Y_1$ (maximum displacement of the resonator). See Table 2.3 for the time values of the temporal envelope.

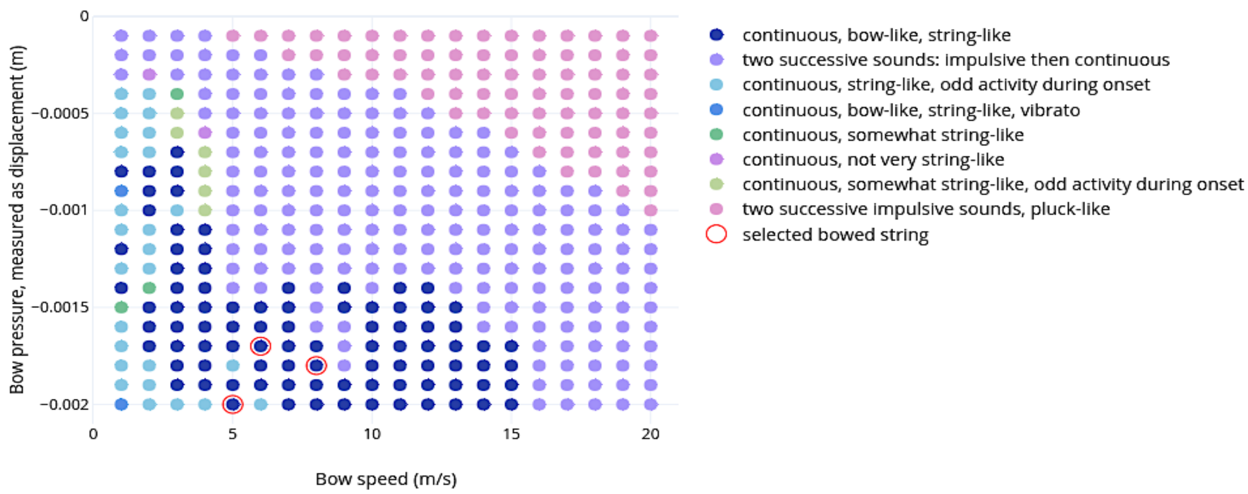


**Figure 2.8.** Reported perceptual outcomes of the 400 synthesized bowed string sounds. The maximum bow speed, $X_1$, is on the x axis; the maximum bow pressure, $Y_1$, is on the y axis. Each point refers to a synthesized bowed string, and the colour indicates the description of the sounds. The red circles indicate $X_1$ and $Y_1$ of the three selected bowed string stimuli.

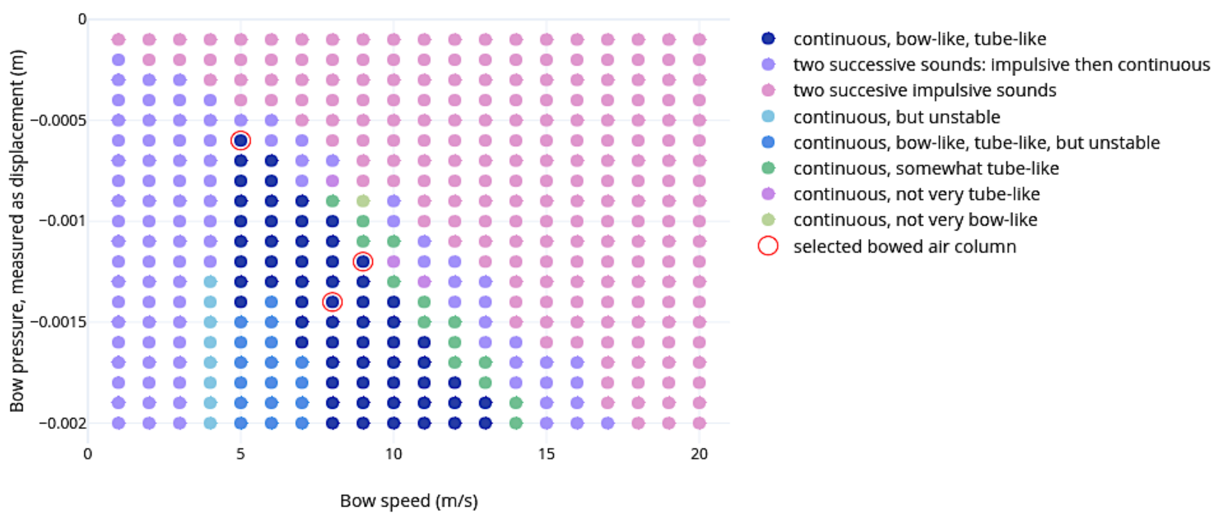


**Figure 2.9.** Reported perceptual outcomes of the 400 synthesized bowed air column sounds. The maximum bow speed, $X_1$, is on the x axis; the maximum bow pressure, $Y_1$, is on the y axis. Each point refers to a synthesized bowed air column, and the colour indicates the description of the sounds. The red circles indicate $X_1$ and $Y_1$ of the three selected bowed air column stimuli.

### 2.4.2 Bowed air column

The bowing excitation interacted with the air column resonator at a unit of 0.5 along its length. At other positions, the resulting sound was either shorter in duration or inharmonic. The output was measured and recorded at a unit of 0.6 along the air column. The freq-loss and const-loss coefficients of the air column resonator were both 1.0 when the bowing was applied to it.

We manipulated the same two parameters as we did for the bowed string: the maximum bow speed ($X_1$), and the maximum bow pressure, ($Y_1$). The bow pressure is negative because the interaction simulates a displacement of the air molecules from their resting position. Manipulating $X_1$ and $Y_1$ significantly impacted the timbre of the resulting sound. Since a bow does not typically interact with an air column in acoustic musical instruments, it was difficult to determine the potential limits of $X_1$ and $Y_1$. Consequently, we applied the same 20 values for $X_1$ and $Y_1$ as the bowed string: 1–20 m/s with increments of 1 m/s for $X_1$ (Fig. 2.6), and 0.1–2.0 mm with increments of 0.1 mm for $Y_1$ (Fig. 2.7). These values maintained consistency with the parameter manipulations of the bowed string, which is ideal for the experimental control we aim to achieve in Chapters 3 and 4. Holding constant all the values of the other parameters mentioned previously, we reported and plotted the perceptual outcomes of the pairing of $X_1$ and $Y_1$ values for the bowed air column (Fig. 2.9). $X_1$ and $Y_1$ values for the three chosen bowed air column sounds are indicated with red circles.

### 2.4.3 Bowed plate

According to Modalys, the bow is applied to the surface of the plate through an up-and-down access along the z direction (Dudas, 2014). The plate could not be bowed from the edge since all edges were fixed, and we wanted to use the same plate for all three types of excitations. Using x and y coordinates corresponding to an edge for the access point therefore resulted in an error. The output of the sound was measured and recorded at x and y coordinates of 0.6 and 0.7, respectively. The freq-loss and const-loss coefficients of the plate were 0.1 and 0.5, respectively.

The two manipulated parameters for the bowing excitation were the same as those that we manipulated for the bowed string and air column: maximum bow speed ($X_1$) and maximum bow pressure ($Y_1$, measured as displacement of the plate from its resting position). Manipulating $X_1$ and $Y_1$ significantly varied the timbre of the resulting sound. The potential limits of $X_1$ and $Y_1$ were difficult to determine because bowing is not typically applied to a plate in acoustic musical instruments. Consequently, we applied the same 20 $X_1$ and $Y_1$ values as the bowed string in Section 2.4.1 (Figs. 2.6–7): 1–20 m/s with increments of 1 m/s for $X_1$ and 0.1–2.0 mm with increments of 0.1 mm for $Y_1$. We plotted the reported perceptual outcomes of the pairing of the $X_1$ and $Y_1$ values of the bowed plate in Figure 2.10. The $X_1$ and $Y_1$ values for the three bowed plate sounds are also indicated with red circles.

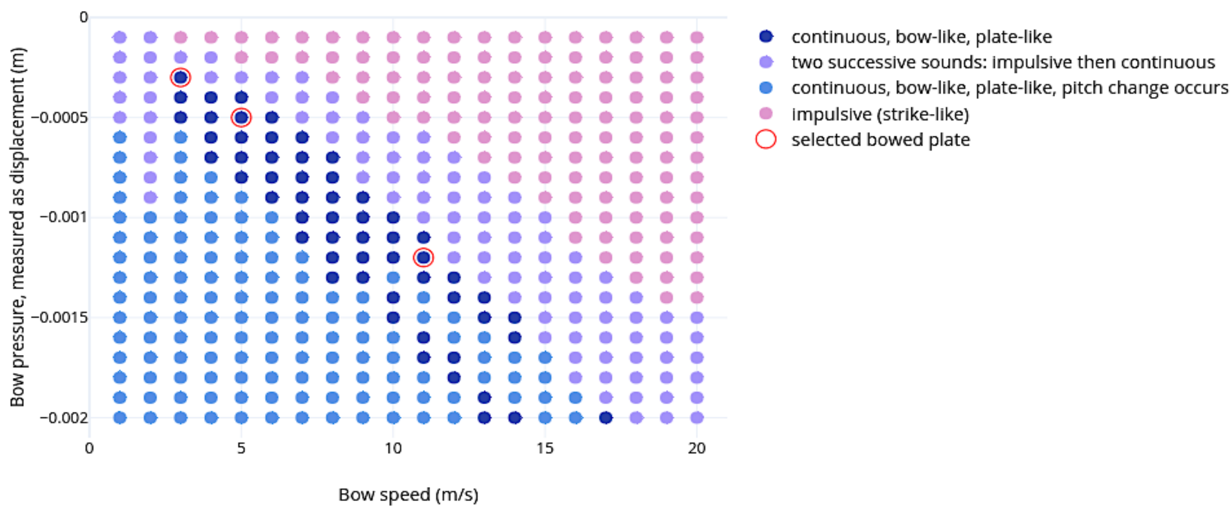


**Figure 2.10.** Reported perceptual outcomes of the 400 synthesized bowed plate sounds. The maximum bow speed, $X_1$, is on the x axis; the maximum bow pressure, $Y_1$, is on the y axis. Each point refers to a synthesized bowed plate, and the colour indicates the description of the sounds. The red circles indicate $X_1$ and $Y_1$ of the three selected bowed plate stimuli.

### 2.4.4 Blown air column

The blowing excitation was applied to the air column resonator at a position of 0.1 units, normalized to its length. This was the same access point as the one in Example 13 (Appendix 3); an access point of 0 did not produce a realistic blown air column sound. The output was measured and recorded at a unit of 0.5 along the air column. The freq-loss and const-loss coefficients of the air column were 0.8 and 1.0, respectively.

The two parameters we manipulated for the blown air column were the maximum breath pressure ($X_2$), and maximum reed opening parameter ($Y_2$), because they significantly impacted the timbre of the resulting sound. Since $X_2$ and $Y_2$ were both normalized values measured on a linear scale, we tested 20 values of both parameters from 0.05 to 1.00 with increments of 0.05 (Figs. 2.11–2.12). These were within the limits of $X_2$ and $Y_2$, given that they were both measured on a linear scale. When 0 was applied to either of these parameters, Modalys did not generate an output and there was an error. The reported perceptual outcomes of the pairing of the $X_2$ and $Y_2$ values, while holding values of all other parameters constant, are plotted in Figure 2.13. The $X_2$ and $Y_2$ values for the three chosen sounds are also summarized in the figure.

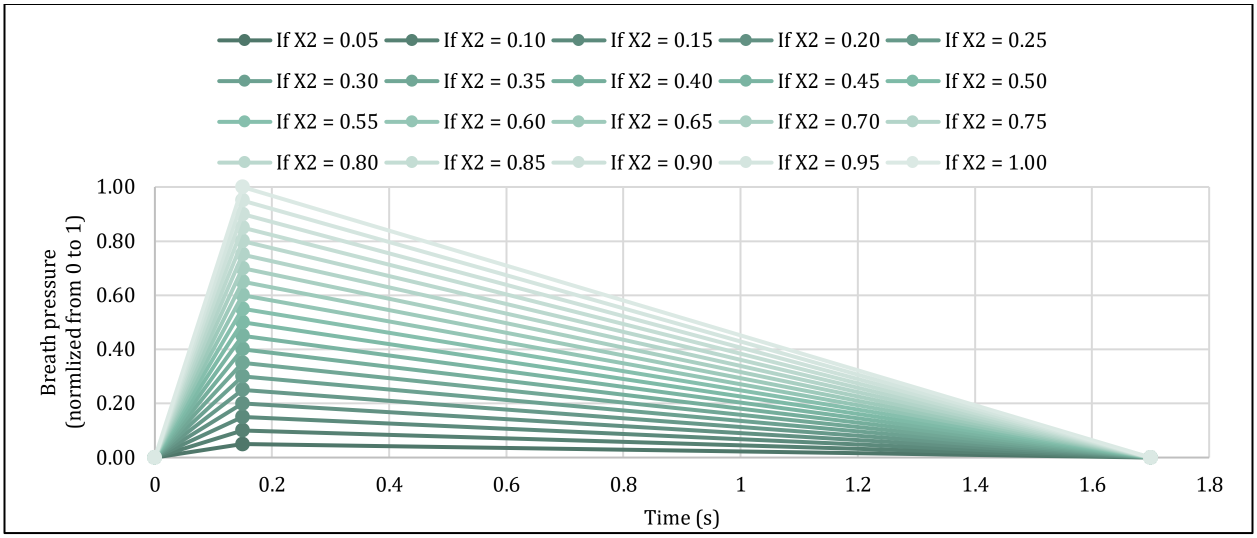


**Figure 2.11.** Twenty possible temporal envelopes for the control of the breath pressure when blowing was applied to the air column. The temporal envelopes depended on the value of $X_2$ (maximum breath pressure). The temporal envelopes were similar for the string and plate, except that the time at which $X_2$ occurred was different. See Table 2.4 for the time values of the temporal envelopes.

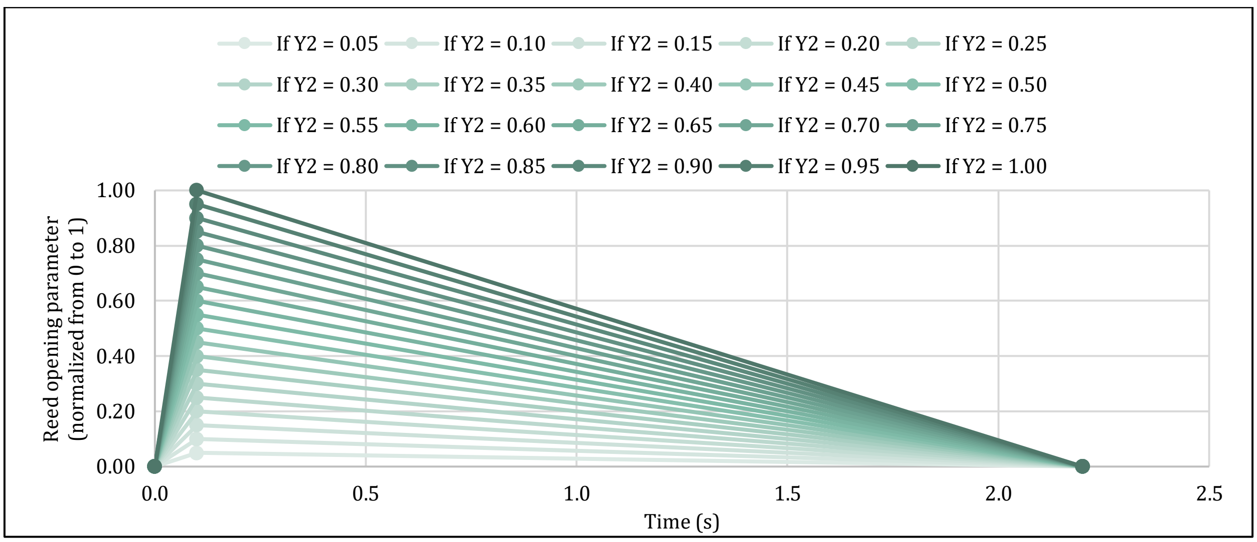


**Figure 2.12.** Twenty possible temporal envelopes for the control of the reed opening parameter, depending on the value of $Y_2$ (maximum reed opening). See Table 2.5 for the time values of the temporal envelope.

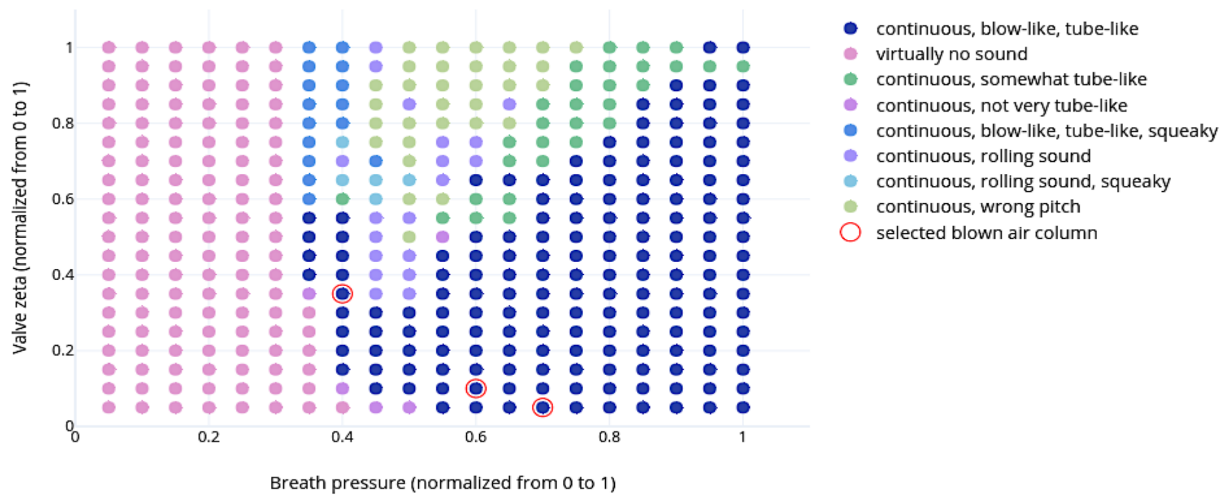


**Figure 2.13.** Reported perceptual outcomes of the 400 synthesized blown air column sounds. The maximum breath pressure, $X_2$, is on the x axis; the maximum value of the reed opening parameter, $Y_2$, is on the y axis. Each point refers to a synthesized blown air column, and the colour indicates the description of the sounds. The red circles indicate $X_2$ and $Y_2$ of the three selected blown air column stimuli.

### 2.4.5 Blown string

The blowing excitation was applied to the string resonator at a position of 0.1 units. The output was measured and recorded at a position of 0.866025 units along the string. This value corresponded to (√3)/2 and was chosen because a rational number for the position of the output produced a sound with more energy in the odd harmonics than in the even harmonics. With more energy in the odd harmonics, the sound resembles a clarinet; however, the sound we aimed to achieve was that of a string. The freq-loss and const-loss coefficients were 0.05 and 0.1, respectively.

The two parameters that were manipulated were the same as the blown air column: the maximum breath pressure ($X_2$) and maximum reed opening parameter ($Y_2$), which significantly affected the timbre of the resulting sound. As these two values were normalized, we tested 20 values of each parameter from 0.05 to 1.00 with increments of 0.05 units (Figs. 2.11–2.12). We plotted the perceptual outcomes of the 400 sounds in Figure 2.14 and indicates the $X_2$ and $Y_2$ values of the three chosen sounds with red circles.

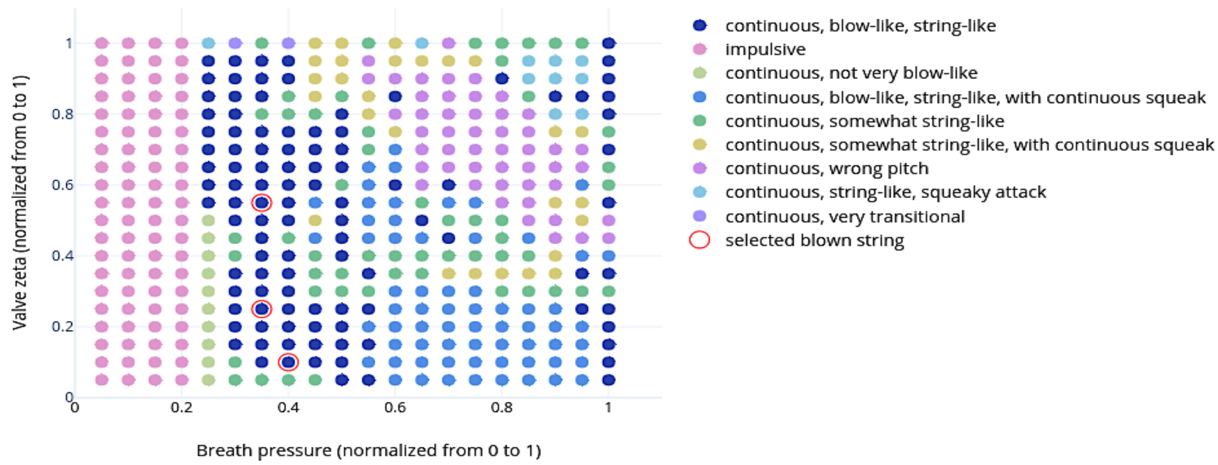


**Figure 2.14.** Reported perceptual outcomes of the 400 synthesized blown string sounds. The maximum breath pressure, $X_2$, is on the x axis; the maximum value of the reed opening parameter, $Y_2$, is on the y axis. Each point refers to a synthesized blown string, and the colour indicates the description of the sounds. The red circles indicate $X_2$ and $Y_2$ of the three selected blown string stimuli.

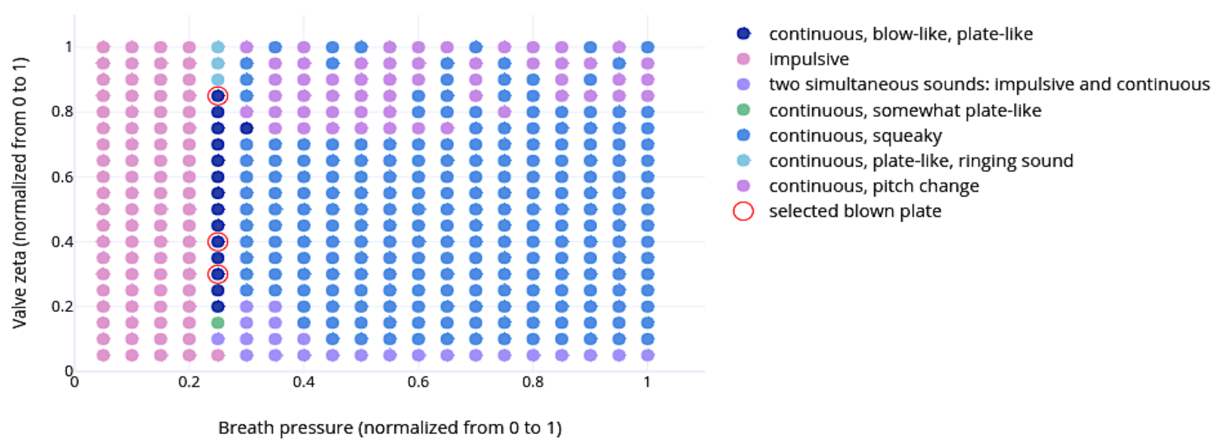


**Figure 2.15.** Reported perceptual outcomes of the 400 synthesized blown plate sounds. The maximum breath pressure, $X_2$, is on the x axis; the maximum value of the reed opening parameter, $Y_2$, is on the y axis. Each point refers to a synthesized blown plate, and the colour indicates the description of the sounds. The red circles indicate $X_2$ and $Y_2$ of the three selected blown plate stimuli.

### 2.4.6 Blown plate

The output was measured and recorded at x and y coordinates of 0.1 and 0.2, respectively, when the blowing excitation was applied to the plate resonator. The freq-loss and const-loss coefficients were 0.01 and 0.5, respectively. The two manipulated parameters were the maximum breath pressure ($X_2$) and maximum reed opening parameter ($Y_2$), which significantly impacted the timbre of the output and were consistent with manipulations of the blown air column and string. The same

20 values were tested for each parameter: 0.01–1.0 with increments of 0.05 units (Figs. 2.11–2.12). Four-hundred blown plate sounds were synthesized, and we reported on their outcome when each of the 20 $X_2$ values were paired with each of the 20 $Y_2$ values, while holding the values of the other parameters constant. The perceptual outcomes of the 400 sounds and the $X_2$ and $Y_2$ values for the three chosen blown plate sounds are plotted with red circles in Figure 2.15.

### 2.4.7 Struck plate

The freq-loss and const-loss coefficients of the plate were 0.05 and 0.6, respectively. We manipulated the x and y coordinates ($X_3$' and $Y_3$', respectively), which corresponded to where the output was measured and recorded. Changes in these values corresponded to significant changes in the timbre of the output—more so than when we changed the x and y coordinates that corresponded to the access point. This may be due to the symmetrical nature of the plate resonator. That is, the timbre resulting from one access point may be perceived as very similar to the timbre resulting from an access point that is at the opposite side (i.e., reflected) of the previously mentioned access point, if the x and y coordinates of 0.5 and 0.5 correspond to the midpoint. Consequently, the 20 values of both $X_3$' and $Y_3$' that we tested are between 0.05 units and 1.00 unit, with increments of 0.05 units. Modalys did not generate an output when $X_3$' and $Y_3$' values of 1 were entered, so we used 0.99999 instead of 1. Moreover, as mentioned in Section 2.3.3, we did not manipulate the maximum pressure of the hammer, which was measured as a displacement of the plate relative to its resting position. Changing the maximum hammer pressure did not seem to significantly impact the timbre of the output. This is because Modalys normalizes the amplitude of the output. Thus, we kept the maximum displacement at 1 mm below the resting position (Fig. 2.16). The position of the hammer was 100 mm above the plate before and after making contact with it. Contact is made at a position of 0 mm or lower. The perceptual outcomes of the 400 synthesized sounds are plotted in Figure 2.17. We also summarized the $X_3$' and $Y_3$' values of the three chosen struck plate sounds in the figure.

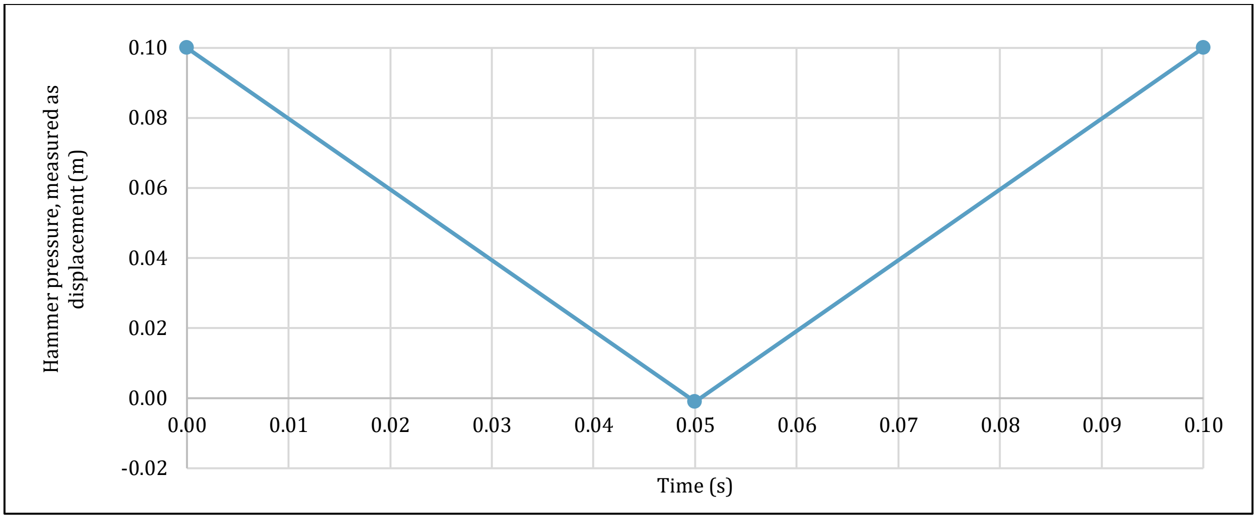


**Figure 2.16.** Temporal envelope for the control of the hammer pressure, when it was applied to the plate resonator. There was only one temporal envelope for the plate, as we did not manipulate this parameter when the hammer was applied to it. See Table 2.6 for the time values of the temporal envelopes.

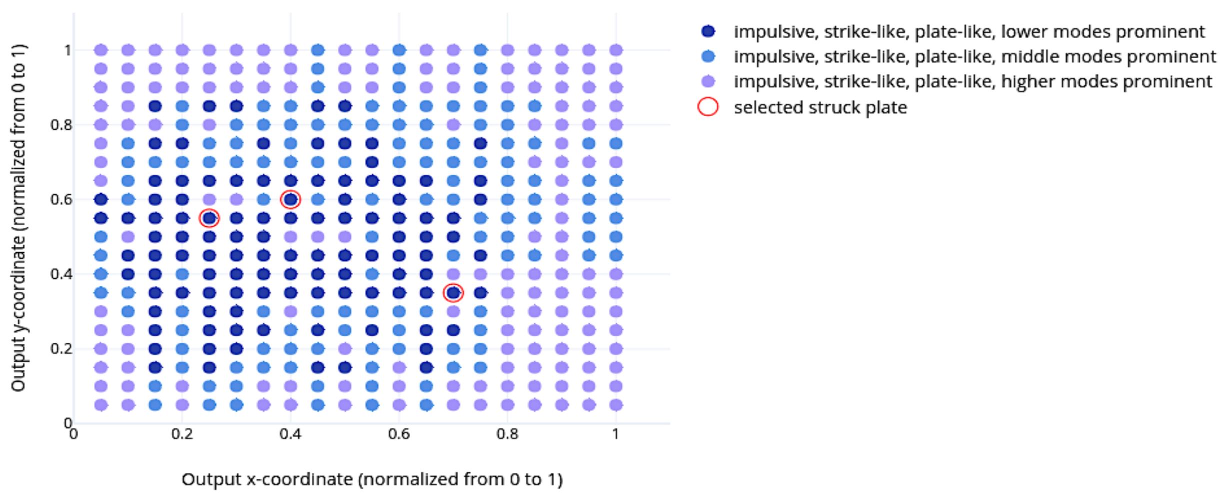


**Figure 2.17.** Reported perceptual outcomes of the 400 synthesized struck plate sounds. The x coordinate of the output on the surface of the plate, $X_3$*'*, is on the x axis; the y coordinate of the output, $Y_3$*'*, is on the y axis. Each point refers to a synthesized struck plate, and the colour indicates the description of the sounds. The red circles indicate $X_3$*'* and $Y_3$*'* of the three selected struck plate stimuli.

### 2.4.8 Struck string

The striking excitation was applied to the string resonator at a position of 0.7 units. Like the struck plate, changing the access point of the hammer onto the string did not result in significant changes in the output's timbre. This was probably due to the symmetrical nature of the string which was fixed at both ends. Applying the hammer to the string at one access point might be perceived as very similar in timbre to applying the hammer at an access point on the opposite side (i.e., reflected) if we consider 0.5 units as the midpoint. The freq-loss and const-loss coefficients of the string were 0.5 and 0.1.

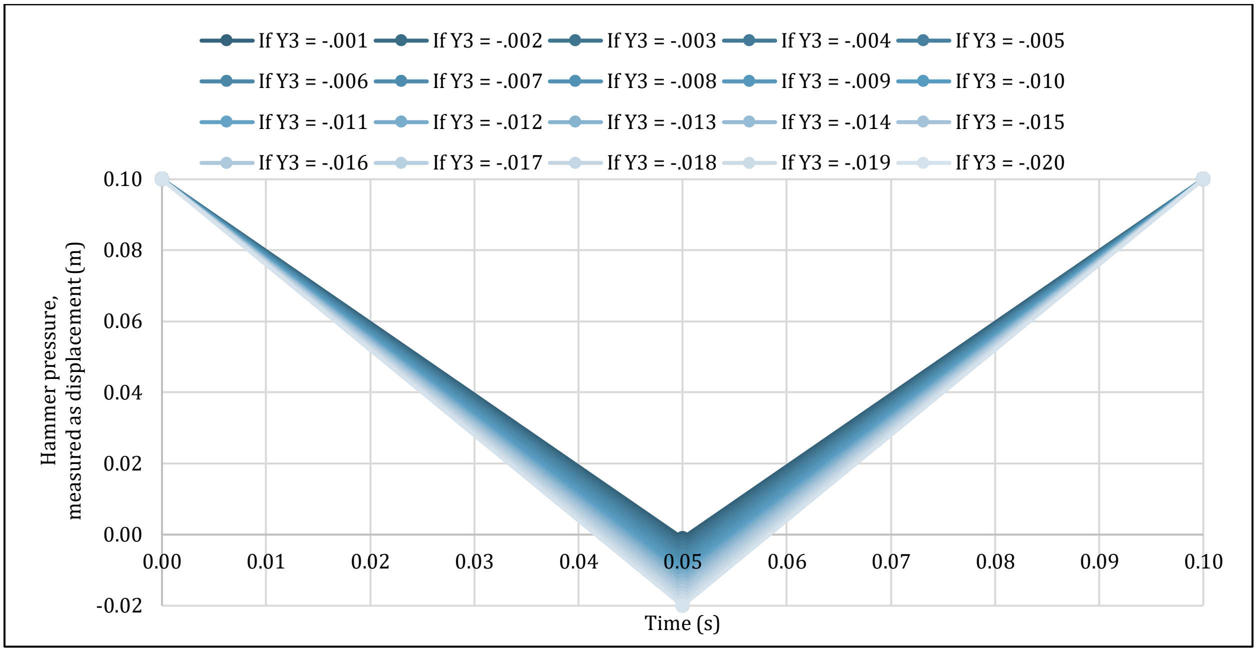


**Figure 2.18.** Possible temporal envelopes for the control of the hammer pressure, when it was applied to the string and air column resonators. There were 20 possible temporal envelopes when the hammer was applied to the string and air column, depending on the value of $Y_3$ (maximum displacement of the resonator). See Table 2.6 for the time values of the temporal envelopes.

We manipulated the values of two parameters that changed the timbre of the resulting sound significantly: the position corresponding to where the output was recorded and measured ($X_3$) and the maximum pressure of the hammer ($Y_3$). $X_3$ was measured on a linear scale and was normalized to the length of the string. $Y_3$ was measured as a displacement of the string from its resting position. We speculated that there was more variability in the resulting timbre when $Y_3$ is manipulated because a string can be more easily displaced from its resting position in comparison to a plate, given that a string is more flexible and less stiff than a metal plate. We tested 20 values of $X_3$ from 0.05 to 1.00 with increments of 0.05. Since an input of 1 for $X_3$ resulted in an error, we inputted 0.99999 instead of 1. Even by testing these $X_3$ values, the struck string was often perceived as a pluck. A pluck makes soundless contact with a string, whereas a strike sets a string into vibration upon contact; however, both excitation methods can displace a string from its resting position, which might explain the similarities in the resulting sounds of struck and plucked strings. The 20 values of $Y_3$ that we tested were between 1 mm and 20 mm, with increments of 1 mm (Fig. 2.18). These values were within the limits of the hammer pressure. The perceptual outcomes of the 400 synthesized struck string sounds are plotted in Figure 2.19. Although all of the sounds were described as "impulsive, strike-like, string-like", the three that were perceived as the most variable were chosen to be used in the experiments outlined in Chapters 3 and 4. The $X_3$ and $Y_3$ values for each of the three chosen struck string sounds are also indicated in Figure 2.19.

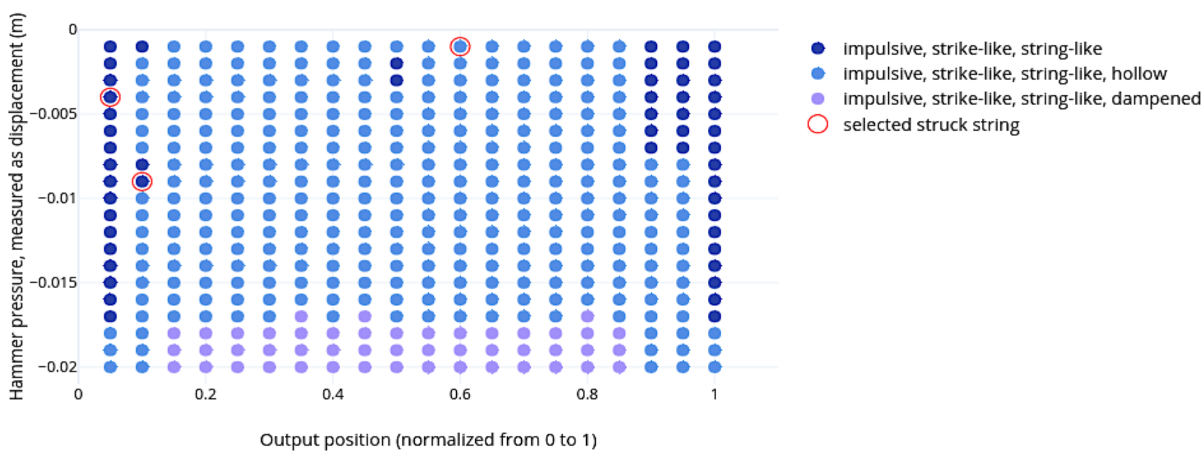


**Figure 2.19.** Reported perceptual outcomes of the 400 synthesized struck string sounds. The position of the output along the length of the string, $X_3$, is on the x axis; the maximum pressure of the hammer, $Y_3$, is on the y axis. Each point refers to a synthesized struck string, and the colour indicates the description of the sounds. The red circles indicate $X_3$ and $Y_3$ of the three selected struck string stimuli.

### 2.4.9 Struck air column

The striking excitation interacted with the air column at 0.7 units. Similar to the struck plate and struck string, changing the access point of the hammer onto the air column did not alter the timbre significantly. The freq-loss and const-loss coefficients were both 0.3 when the striking excitation was applied to the air molecules of the air column (i.e., not the tube).

Like the struck string, we manipulated the values of the position along the air column corresponding to the output ($X_3$) and the maximum hammer pressure ($Y_3$) for the struck air column. Manipulating $X_3$ significantly changed the timbre of the resulting sound as it did with the struck string. $X_3$ was measured on a linear scale and was normalized to the length of the air column. $Y_3$ was measured as the displacement of the air molecules from their resting position. We speculated that there was more variability in the resulting timbre when $Y_3$ was manipulated for the air column compared to the plate because air molecules can be easily displaced, whereas a plate made of metal is quite stiff. We tested 20 $X_3$ values from 0.05 to 1.00, with increments of 0.05. These were the same values we tested for $X_3$ of the struck string. This maintained consistency for the experimental control sought in the following two chapters. Like the struck string, inputting a value of 1 resulted in an error; instead, we applied a value of 0.99999. Given that a striking excitation does not typically interact with an air column resonator in acoustic musical instruments, it was difficult to decipher the limits of $Y_3$. Consequently, we used the same 20 values of $Y_3$ as the struck string: 1–20 mm with increments of 1 mm (Fig. 2.18). This maintained consistency with the struck string which was ideal for the experimental control sought for Chapters 3 and 4. The reported perceptual outcomes of the 400 struck air column sounds are plotted in Figure 2.20. The majority of the sounds were classified as “impulsive, strike-like, air column-like”, so we chose three that were perceived as the most variable. The three chosen struck air column sounds in Figure 2.20 are indicated with red circles.

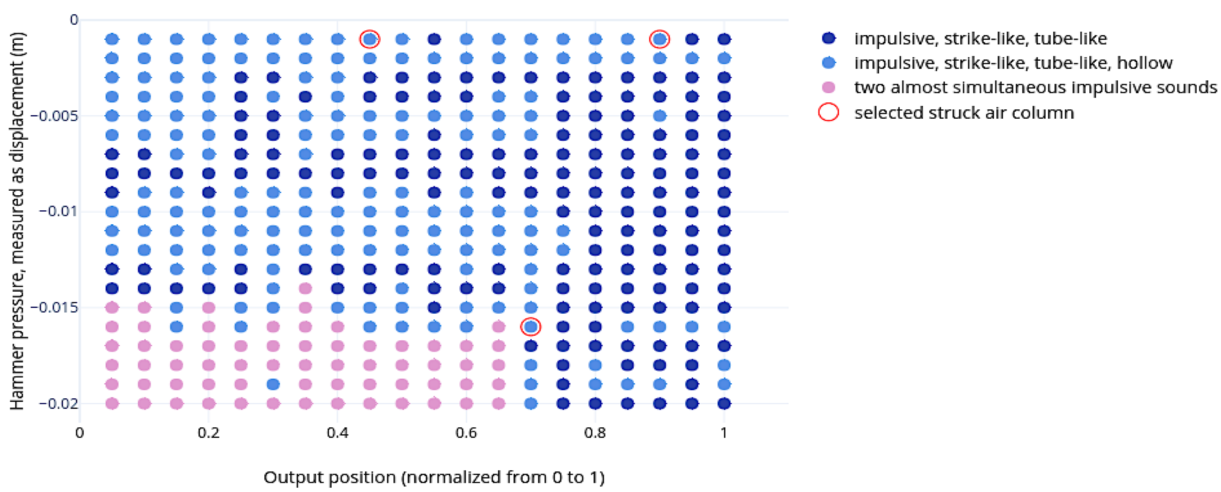


**Figure 2.20.** Reported perceptual outcomes of the 400 synthesized struck air column sounds. The position of the output along the length of the air column, $X_3$, is on the x axis; the maximum pressure of the hammer, $Y_3$, is on the y axis. Each point refers to a synthesized struck air column, and the colour indicates the description of the sounds. The red circles indicate $X_3$ and $Y_3$ of the three selected struck air column stimuli.

## 2.5 Acoustical analyses: Spectrograms and audio descriptors

As mentioned in the previous section, three sounds were chosen as exemplars for each type of excitation-resonator interaction. Since there were nine types of excitation-resonator interactions, there were a total of 27 exemplars that were used as stimuli in the experiments detailed in Chapters

3 and 4. They were acoustically analyzed in this section. We first visualized the stimuli in terms of their spectrograms using Matlab (Mathworks, Natick, 2018). We then analyzed the stimuli for three audio descriptors with the Timbre Toolbox (Peeters et al., 2011) and the Music Information Retrieval (MIR) Toolbox (Lartillot et al., 2008), which were the temporal centroid, spectral centroid, and inharmonicity.

### 2.5.1 Spectrograms of the stimuli

A spectrogram is a three-dimensional visual representation of a sound signal. Spectrograms outline the energy within narrow frequency bins over the time course of a signal. The three dimensions are: time in seconds (x axis), frequency in Hz (y axis), and energy in dB (colour). We generated the spectrograms in Matlab, which computed a series of short-term Fourier transforms (STFT). We chose a sampling rate of 44,100 Hz, segment length of 1024 samples, overlap of 512 samples, and number of discrete Fourier transforms (DFT) points of 1024. We displayed a spectrogram for one exemplar of each excitation-resonator interaction: bowed string (Fig. 2.21), bowed air column (Fig. 2.22), bowed plate (Fig. 2.23), blown string (Fig. 2.24), blown air column (Fig. 2.25), blown plate (Fig. 2.26), struck string (Fig. 2.27), struck air column (Fig. 2.28), and struck plate (Fig. 2.29). The spectrograms for the three exemplars of each interaction type were consistent, so we displayed one of each as examples. All spectrograms showed a greater concentration of energy in the lower modes than in the higher modes and a faster decay of the higher modes than their lower counterparts, as expected from the freq-loss and const-loss coefficients.

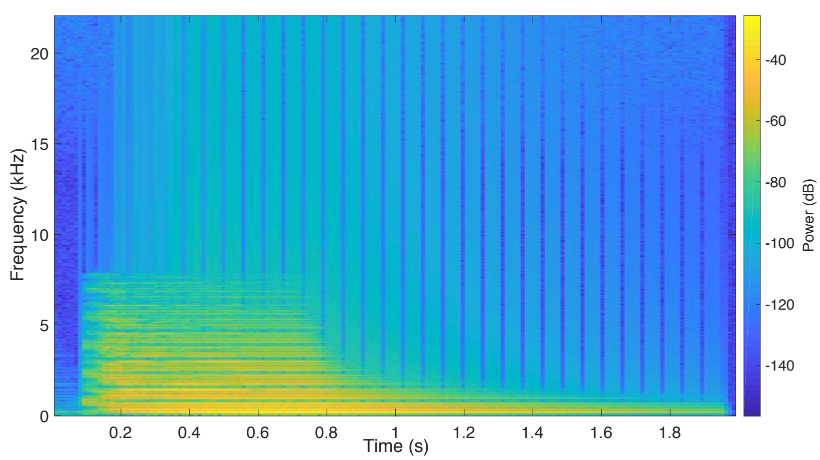


**Figure 2.21.** Spectrogram of a bowed string exemplar.

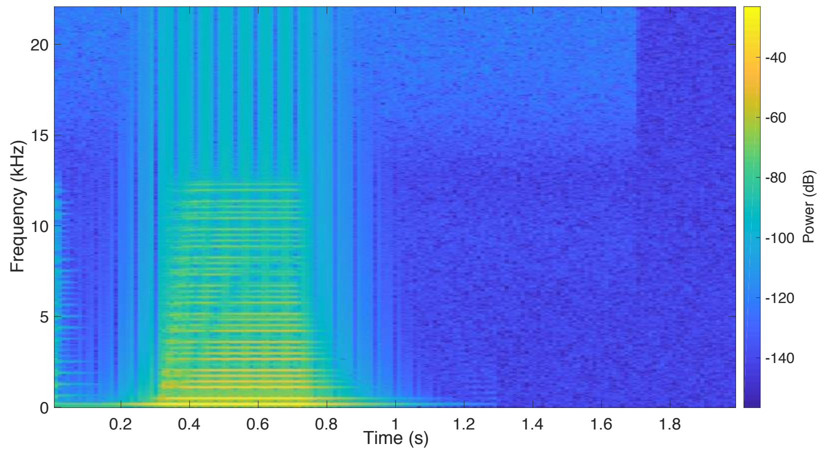


**Figure 2.22.** Spectrogram of a bowed air column exemplar.

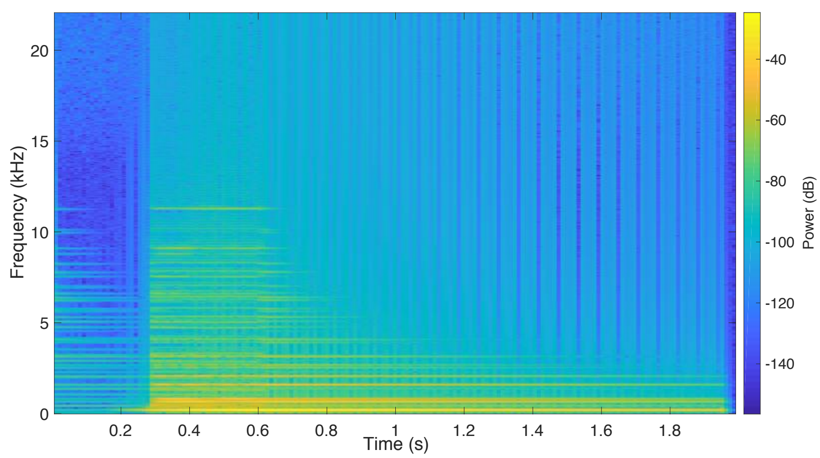


**Figure 2.23.** Spectrogram of a bowed plate exemplar.

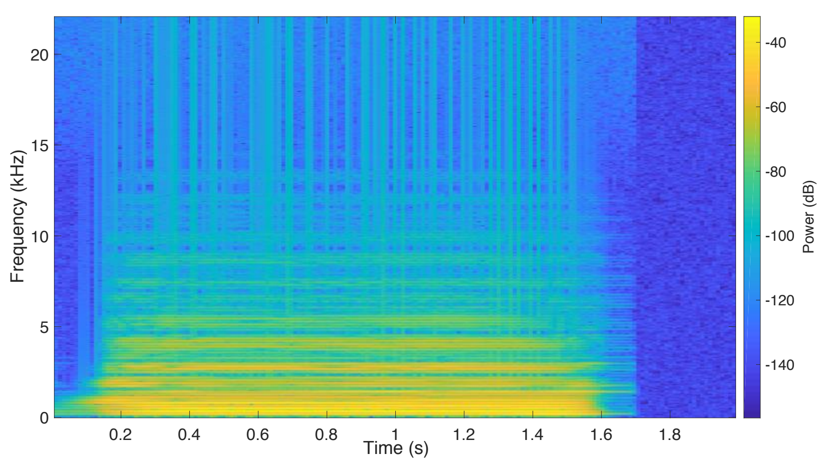


**Figure 2.24.** Spectrogram of a blown string exemplar.

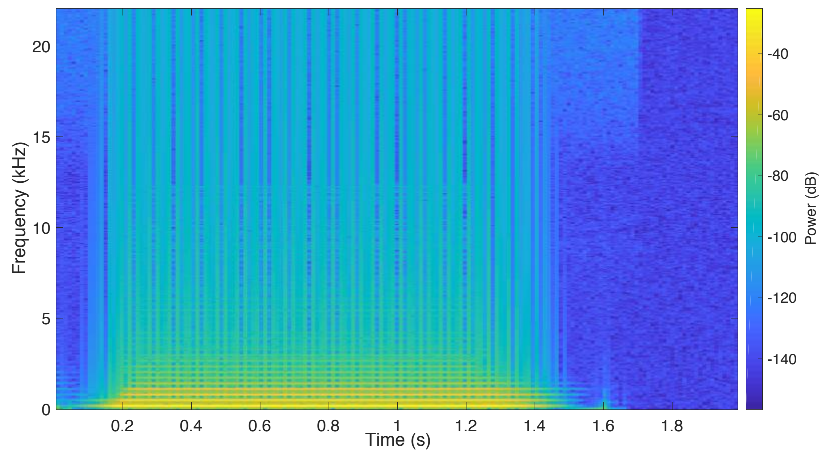


**Figure 2.25.** Spectrogram of a blown air column exemplar.

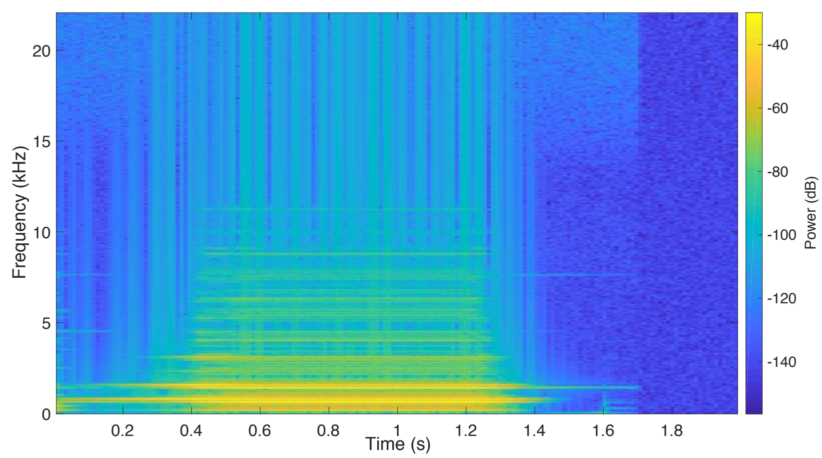


**Figure 2.26.** Spectrogram of a blown plate exemplar.

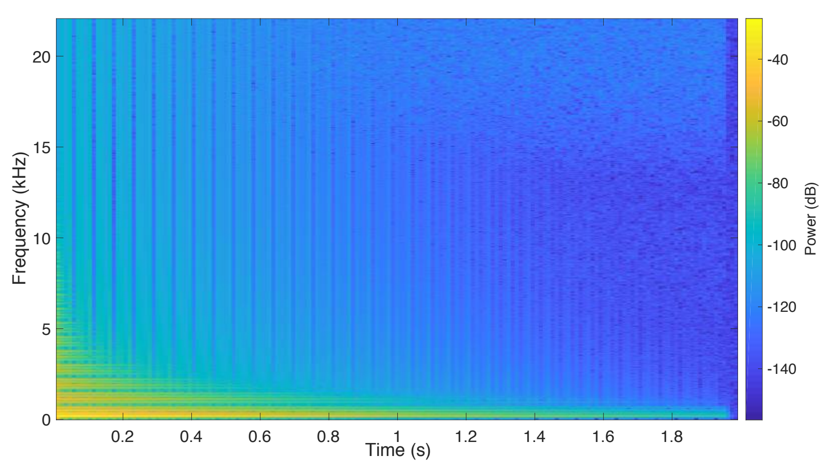


**Figure 2.27.** Spectrogram of a struck string exemplar.

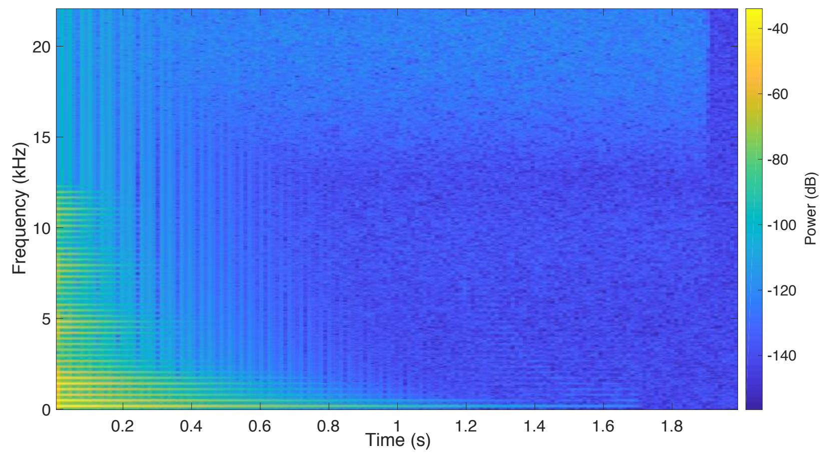


**Figure 2.28.** Spectrogram of a struck air column exemplar.

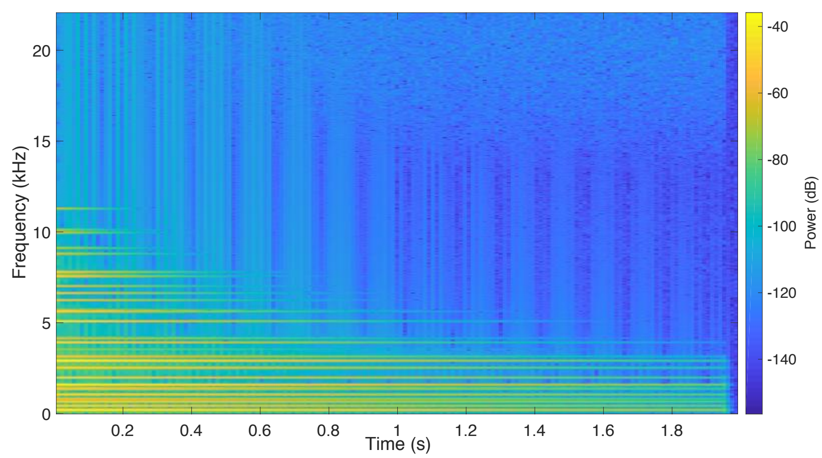


**Figure 2.29.** Spectrogram of a struck plate exemplar.

The spectrograms were very consistent in terms of the overall energy and behaviour of the frequency components for sounds that were produced by the same excitation or the same resonator. The bowed sounds had a similar shape in terms of the temporal envelope regardless of the object to which the bowing was applied. The bowed air column sounds appeared to be shorter in duration compared to the bowed string and bowed plate sounds. We speculated that this was due to the physical nature of the air column resonator. The modes representing the air molecules of the air column are more highly dampened than those representing a string or plate. Blown sounds also displayed a similar temporal envelope shape. That is, the behavior of the modes over time shared a similar shape regardless of the resonator to which the blowing was applied. The struck stimuli consistently demonstrated the impulsive nature of struck sounds. The shape of the temporal envelope was also consistent regardless of the resonator to which the striking was applied. Moreover, sounds produced by the same resonator shared similar frequency contents (i.e., spectrum), regardless of the excitation that was applied.

### 2.5.2 Audio descriptors

Spectral centroid and log attack time were typically reported as the most common audio descriptors that influenced perceptual judgements of the similarity between sounds (McAdams et al., 1995). However, previous research also found that the temporal centroid likely influences the perceived excitation of sounds (Hjortkjær & McAdams, 2016), which is more relevant to the design of our stimuli. Hjortkjær and McAdams (2016) also showed that the spectral centroid influenced the perceived materials of sounds. These audio descriptors might be relevant to our stimuli and experimental design, because participants rated how well each of the stimuli resemble bowing, blowing, and striking excitations, as well as string, air column, and plate resonators. We expected to find that the differences in temporal centroids would influence listeners' perceptions of how well the sounds resemble different excitations. Similarly, we expected that variability in spectral centroids might be associated with listeners' ratings of the resemblance to different resonators.

The Timbre Toolbox was created to extract audio descriptors from sound signals that can be useful in perceptual research (Peeters et al., 2011). We analyzed the temporal centroid of the 27 stimuli that we synthesized. The temporal centroid refers to the centre of mass of the energy distribution over time; it describes where more energy is concentrated across time. The temporal centroids computed by the MIR Toolbox (Lartillot et al., 2008) were very consistent with those of the Timbre Toolbox, so we proceeded with using the values of the Timbre Toolbox. We also analyzed the spectral centroid, which is the centre of mass of the energy distribution across frequencies; it describes whether more energy is concentrated in the high or low frequency components. This parameter was computed with the Timbre Toolbox, but some values that were computed seemed unreasonable and much too high (i.e., >7 kHz). We then computed spectral centroid with the MIR Toolbox (Lartillot et al., 2008), which outputted values that seemed to be more accurate.

We plotted the temporal centroids and spectral centroids of the 27 stimuli (Fig. 2.30). There was a distinct separation between struck (i.e., impulsive) sounds and bowed and blown (i.e., continuous) sounds in terms of their temporal centroids. The temporal centroid has been known to be one of the audio descriptors that distinguished impulsive sounds from continuous sounds (Peeters et al., 2011). The case of the bowed air column stimuli was interesting, as they had temporal centroid values that were between the struck sounds and the remainder of the bowed and blown sounds. This may likely be attributed to the dampening of modes that represent the air molecules in comparison to those that represent the string and plate. If a string or plate has been bowed, it will continue to resonate even when the bow is no longer in contact with it. In the case of the air column, the air will dampen much quicker when the bow is no longer applied to it. This is more obvious in the case of the blown air column: the air continues to resonate as long as blowing is applied to the air column. However, once the blowing stops, the sound ends shortly thereafter.

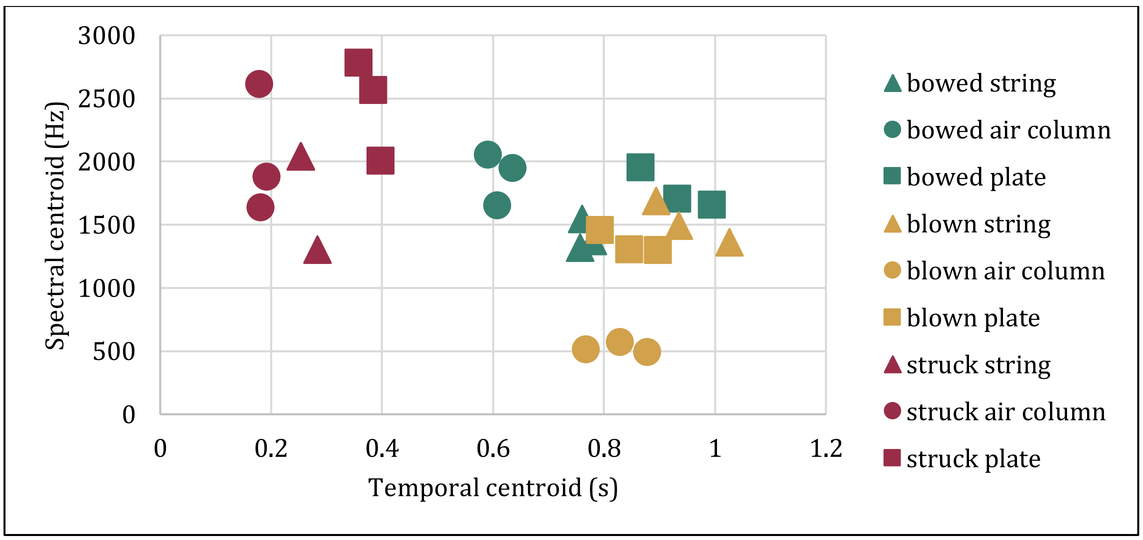


**Figure 2.30.** Temporal centroids (x axis) and spectral centroids (y axis) of the 27 stimuli. There are three points for each of the nine types of exciter-resonator interactions. The excitations of the sounds are indicated by colour, such that bowed, blown, and struck sounds are represented by green, yellow, and red, respectively. The resonators of the sounds are indicated by shape: triangles, circles, and squares represent string, air column, and plate sounds, respectively.

The patterns to describe the spectral centroids of the 27 stimuli were less clear. The string stimuli had spectral centroids between 1,200 and 2,200 Hz. Bowed and blown plates also had spectral centroids that fell in this range, but struck plates had higher values. Air column stimuli had the most variable spectral centroids: blown air columns had low spectral centroids, but those of bowed and blown air columns were much higher in comparison. Since the pattern of the stimuli's spectral centroids with respect to the resonator was vague, we speculated that this would influence how the participants perceived the air column stimuli in Experiment 2 (Chapter 4).

After listening to our 27 exemplars more carefully, we speculated that other spectral or harmonic features could distinguish the different resonators, such as inharmonicity and the odd-to-even ratio. We predicted that inharmonicity would separate plate stimuli from strings and air columns. Sounds produced by strings and air columns have a more harmonic frequency spectrum, such that the partials are integer multiples of the fundamental. Sound produced by plates would exhibit more inharmonic frequency spectra, since their partials would not be integer multiples of the fundamental. The odd-to-even ratio might distinguish string from air column stimuli. We predicted that sounds produced by air columns would have higher odd-to-even ratios than those of strings. In the case of a blown air column, the odd harmonics have greater energy than the even harmonics; this is less likely the case, however, for bowed string stimuli. We attempted to compute inharmonicity and the odd-to-even ratio with the Timbre Toolbox, but the values seemed very inaccurate (i.e., some were much too high to be considered reasonable). We then attempted to analyze inharmonicity with the MIR Toolbox; however, it did not have a function to compute the odd-to-even ratio. Figure 2.31 is a plot of the temporal centroid and inharmonicity of the 27 stimuli.

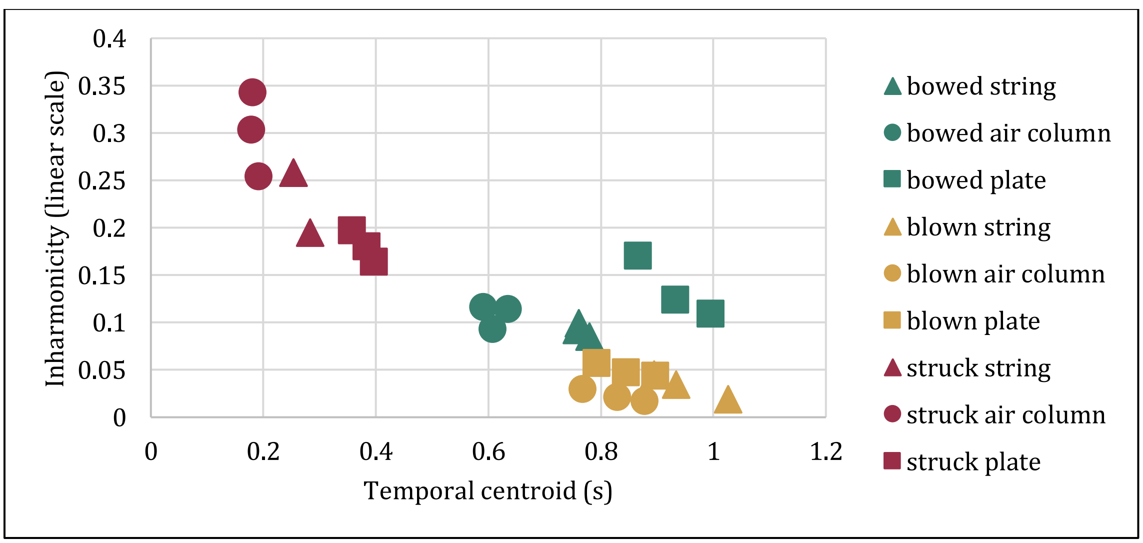


**Figure 2.31.** Temporal centroids (x axis) and inharmonicity values (y axis) of the 27 stimuli. There are three points for each of the nine types of exciter-resonator interactions. The excitations of the sounds are indicated by colour, such that bowed, blown, and struck sounds are represented by green, yellow, and red, respectively. The resonators of the sounds are indicated by shape: triangles, circles, and squares represent string, air column, and plate sounds, respectively.

Unlike our predictions, inharmonicity did not distinguish plate from string and air column stimuli. However, they did distinguish struck from continuous excitations: in general the former had greater inharmonicity values than the latter. A possible explanation is that struck sounds are inherently bound to inharmonic frequency spectra, since they typically have a strong attack and are short in duration. With a short duration, there is less time for the partials to propagate.

## 2.6 Discussion

We synthesized the stimuli with as much control as possible in terms of both the excitation parameters and resonator parameters. That is, parameters of the excitations are kept as consistent as possible regardless of the resonator that it was applied to. Similarly, we kept the parameters of the resonators as consistent as possible no matter what excitation was applied to it. The exceptions to these consistencies were due to either the physical nature of the resonator or constraints of temporal envelopes for the control of some parameters of the excitation. Moreover, we manipulated the values of two parameters for each type of excitation-resonator interaction. The values we tested were as consistent as possible when the stimuli had the excitation in common. The only exception was for the struck sounds. The struck plate had different parameter manipulations than the struck string and air column, due to the physical nature of the different resonators.

Maintaining as much consistency as possible for the stimuli had its advantages and disadvantages. The advantages contributed to the experimental control that is required for efficient experimental design. For example, we wanted to ensure that differences in how listeners perceive the stimuli were attributed to the manipulated parameters and not due to differences in parameters we did not manipulate. On the other hand, maintaining as much consistency as possible in the excitation and/or resonator parameters can be disadvantageous, as it does not represent the generalizability of how the stimuli can be produced. In other words, the stimuli were synthesized such that they were produced in very specific manners. To compensate for this disadvantage, we chose the three sounds with the most variability in terms of their perceived timbre for each of the nine excitation-resonator interactions. Even with this variability, the three chosen sounds could be perceived as being produced by the same instrument. Three sounds for each of the nine interactions were chosen to reduce experimental fatigue during the experiments to be reported in Chapters 3 and 4.

The spectrograms of the chosen 27 stimuli show consistency in the shape of the temporal envelopes for sounds that share the same excitation. Moreover, the stimuli were consistent in their frequency spectrum if they were produced by the same resonator. These consistencies highlight the controlled parameter manipulation that was applied to the excitations and resonators. Moreover, plotting the temporal centroids with spectral centroids or inharmonicity values showed some predictable patterns. Temporal centroids characterized the excitations of the stimuli as impulsive and continuous for the most part: struck sounds were impulsive and bowed and blown sounds were continuous. Variability in the spectral centroids of stimuli that were produced by the same resonator was attributed to the type of resonator: air column stimuli had the most variable spectral centroids, whereas string stimuli had the least variable spectral centroids. This was likely due to the physical nature of the resonators. Inharmonicity seemed to distinguish impulsive from continuous excitations: struck sounds had higher inharmonicity values. This was likely because struck sounds are shorter, so the partials did not have enough time to propagate.

Another way we could have synthesized the stimuli was through direct comparisons with recorded sounds—specifically for the typical excitation-resonator interactions. That is, for the bowed string, blown air column, struck string, and struck plate sounds, we could have recorded these sounds and then used Modalys to synthesize them, so that they sound as similar as possible to the recordings. At least then we would know what parameters of the excitations and resonators produced the most reasonable outputs. Then, we could isolate the parameters of the different excitations and apply them to the parameters of resonators to which they are not typically applied. This approach might have produced more reliable sounds for the atypical interactions, although it would still be difficult to decipher their accuracy, since they are completely novel and not physical by nature.

The question of interest now becomes: Do these stimuli sound like what they are supposed to sound like in terms of their excitation methods and resonance structures? In the case of typically paired excitation-resonator interactions, such as the bowed string, blown air column, struck plate, and struck string, the answer to this question might be simpler than for atypical excitation-resonator interactions, such as the bowed air column, bowed plate, blown string, blown plate, and struck air column. Stimuli of typically paired excitation-resonator interactions fit within listeners' mental models; stimuli of atypical pairings, however, do not fit within listeners' mental models of how excitations and resonators interact with each other in acoustic musical instruments. Consequently, the following two chapters addressed how well the stimuli were perceived to resemble either the excitation (Chapter 3) or resonator (Chapter 4) that produced them.

# Chapter 3

# Experiment 1: Perceived Resemblance of Excitation Methods—Bowing, Blowing, and Striking

In this experiment, participants rated how well the 27 stimuli chosen in Chapter 2 resembled three excitation methods—bowing, blowing, and striking—using a similar paradigm to that of Lemaitre and Heller (2012). We tested the hypothesis that the excitation methods (excitations) and resonance structures (resonators) that produced a sound influence how well listeners can recognize bowing, blowing, and striking excitations. Moreover, recognition depends on how well these exciter-resonator interactions align with our mental models. By grouping our participants into musicians and nonmusicians, we tested if there were differences in how these two groups rated the sounds' resemblance to the excitations based on the level of formal musical training. We examined if listeners did indeed perceive the intended excitations we synthesized with Modalys, especially for sounds that were produced by excitation-resonator interactions that are atypical of acoustic musical instruments (e.g., bowed air column, bowed plate, blown string, blown plate, struck air column). Since these sounds have rarely been synthesized, we investigated if listeners perceived them as resembling the excitation that produced them, an excitation that did not produce them, or none of the excitations we synthesized.

## 3.1 Method

The experiment was a factorial mixed-measures design with three dependent measures. Participants rated 27 stimuli according to how well each of them sounded like they were bowed, blown, or struck. For each excitation rating, there were three independent variables. The first one was a between-groups factor of musicianship with two levels: musician and nonmusician. The second one was a repeated-measures factor with three levels that corresponded to the excitation that produced the sound: bowing, blowing, or striking. The third one was a repeated-measures factor with three levels that corresponded to the resonator that produced the sound: string, air column, or plate.

### 3.1.1 Participants

Fourty-three participants (22 female, 21 male) took part in the experiment. We recruited the participants from either a mailing list or web-based advertisement certified by McGill University. All participants reported normal hearing, which was confirmed by a pure-tone audiometric test with octave-spaced frequencies from 125 Hz to 8000 Hz at a hearing threshold of 20 dB HL (ISO 398-8, 2004; Martin & Champlin, 2000). Participants were compensated for their participation with $5. The study was certified for ethical compliance by the McGill University Research Ethics Board II.

Data from two participants were discarded. One participant reported that they did not perform the task correctly, and another participant misrepresented their musical background. Of the remaining participants, there were 21 musicians (10 female, 11 male) and 20 nonmusicians (12 female, 8 male). Musicians were classified as being in at least their third year of an undergraduate music program at the Schulich School of Music of McGill University. This classification ensured that musician participants had enough formal musical training. Nonmusicians were classified as having less than five years of formal musical training. Musician participants had an average age of 25.86 years ($SD$=4.97) and nonmusician participants had an average age of 23.90 years ($SD$=3.99). Musicians and nonmusicians did not differ significantly in their ages, $t(39)=1.39$, $p=.17$. Musicians had an average of 14.38 years of formal musical training ($SD$=6.37) on at least one instrument. Nonmusician participants had an average of 1.03 years of formal musical training ($SD$=1.71). Musicians had significantly more years of formal musical training than nonmusicians, $t(39)=9.07$, $p<.001$.

We ran both Experiments 1 and 2 simultaneously with two independent groups; two months later, we ran more participants for each experiment from the other participant pool. That is, some participants who participated in Experiment 1 returned to participate in Experiment 2 and vice versa. Eight participants (3 musicians, 5 nonmusicians) who participated in this experiment had already participated in Experiment 2. The returning participants reported remembering very little of the sounds and experimental procedure of Experiment 2. The shortest time between participation in the two experiments was 69 days and the longest time was 86 days.

### 3.1.2 Stimuli

The stimuli consisted of the 27 sounds we synthesized in Chapter 2. They modeled nine types of interactions between three excitations and three resonators: bowed string, bowed air column, bowed plate, blown string, blown air column, blown plate, struck string, struck air column, and struck plate. This type of stimulus design was adapted from that of previous studies that paired different actions with different materials. Three versions of each of the nine pairings were chosen for the experiment. Each of the stimuli had a fundamental frequency of 155 Hz, corresponding to a pitch of E-flat-3.

Three additional practice stimuli were synthesized with Modalys (Dudas, 2014). We chose to synthesize the practice stimuli to convey three excitation-resonator interactions that are typical of acoustic musical instruments: bowed string, blown air column, and struck plate. These interactions were chosen to help participants understand the experimental task, since the excitations and resonators that produced the sounds were quite obvious. We modified the Modalys scripts of the bowed string, blown air column, and struck plate stimuli that we generated in Chapter 2 to synthesize the practice stimuli. The practice stimuli had a fundamental frequency of 220 Hz, corresponding to a pitch of A3.

Eleven individuals participated in a loudness-matching paradigm to normalize the loudness levels of all the stimuli, including practice stimuli. None of the participants of the loudness-matching paradigm participated in the experiment. One version of the blown air column was chosen as the standard. The task was to listen to each stimulus and adjust its level to match its loudness to that of the standard. The median decibel (dB) difference of the comparison and the standard was used as the level adjustment for each stimulus. We modified each stimulus to have a duration of 2 s: stimuli longer than 2 s were trimmed, and silence was added to the end of stimuli that were shorter than 2 s. A fade out of 50 ms was applied to the end of each stimulus to prevent auditory clicks or an abrupt cut off in the sound. These duration modifications were performed on Matlab (Mathworks, 2018). We chose 2 s as the duration because for longer stimuli, no new or additional information was conveyed.

### 3.1.3 Apparatus

The experiment was conducted in the Perceptual Testing Lab at the Center for Interdisciplinary Research in Music Media and Technology (CIRMMT) at McGill University. The experiment ran on a Mac Pro computer running OS 10.7 (Apple Computer, Inc, Cupertino, CA) and was displayed on an Apple Display 23-inch screen. The stimuli were presented over Seinnheiser HD280 Pro headphones (Sennheiser Electronic GmbH, Wedemark, Germany) and were amplified through a Grace Design m904 monitor (Grace Digital Audio, San Diego, CA). We measured the physical levels of the sounds by coupling the headphones to a Bruel and Kjær Type 4153 Artificial Ear (Bruel & Kjær, Nærum, Denmark). The sounds varied in level from 57.8 to 72.8 dB SPL. The experiment was programmed in the PsiExp computer environment (Smith, 1995).

### 3.1.4 Procedure

Terminology was defined for the participants before beginning the experiment. To simplify the definitions, we explained that for acoustic musical instruments, the pairing of an excitation method and resonance structure is required to produce a tone. We defined the excitation method as the type of *action* performed to set the resonance structure into vibration. We defined the resonance structure as an *object* that vibrates and radiates sound. During the discussion of the experimental procedure and results, we will refer to excitations as actions and resonators as objects. We then described the three actions as: bowing—the action of rubbing a bow on an object to make it vibrate; blowing—the action of blowing into a mouthpiece to make an object vibrate; and striking—the action of using a mallet to hit an object to make it vibrate. We did not define, or inform participants of, the objects that produce the sounds, because we wanted them to focus solely on the action properties.

Participants were instructed on the procedure of the experiment verbally and through written instructions. Following the instructions, they completed three practice blocks (three trials per block), which followed the same format and paradigm as the experimental blocks and allowed participants to pose questions of clarification regarding the procedure before beginning the experimental blocks. At the beginning of the experimental blocks, participants heard the full range of the stimuli to get a sense of the variability of the actions and objects producing the sounds. The full range of stimuli were presented in random order with an inter-onset interval (IOI) of 2000 ms.

The experiment had three blocks: ratings in each block concerned only one type of action. We separated the three types of ratings into three different blocks so participants were able to focus on a stimulus' resemblance to one type of action at a time. The order of the blocks was randomized for each participant. Within each block, there were 27 trials—one for each stimulus (3 excitations × 3 resonators × 3 exemplars). Thus, there was a total of 81 trials in the experiment. In each trial, participants were instructed to play the stimulus and rate its resemblance to the target action, i.e., the extent to which it sounded like the target action of that block (Fig. 3.1). They were able to play the stimulus only once per trial. We added 50 ms of silence to the beginning of the stimulus to prevent the sound of a mouse click from influencing its perception. The resemblance rating was performed on a continuous slider: the left end of the slider was labeled as "not at all," and the right end of the slider was labeled as "completely". Participants were instructed to use the full range of the slider over the course of the block. We coded ratings on the slider as values from 0 to 100 for the analyses, but the values did not appear on the interface of the experiment. The order of stimulus presentation within each block was pseudo-randomized, such that two stimuli produced by the same excitation-resonator interaction were not presented in successive trials. For example, if a given trial presented a bowed plate stimulus, the previous and following trials could not present another bowed plate stimulus. Once participants were satisfied with their rating, they clicked "next" to proceed to the following trial.

This procedure was modified from that of Lemaitre and Heller (2012): they combined four types of actions (scraping, rolling, hitting, and bouncing) with four types of materials (wood, plastic, glass, and metal). Participants in their study rated how well the sounds conveyed the actions and materials that did or did not actually generate them. Half of the ratings for the resemblance of an action or material matched the actual action or material that produced the sound. The other half of the ratings for the resemblance of an action or material did not match the action or material that produced the sound. In the current procedure, however, one-third of the action-resemblance ratings matched the action that actually produced the stimulus, and the remaining two-thirds did not. This allowed us to examine whether other actions were perceived to be conveyed by the stimuli or if none of the three actions were conveyed.

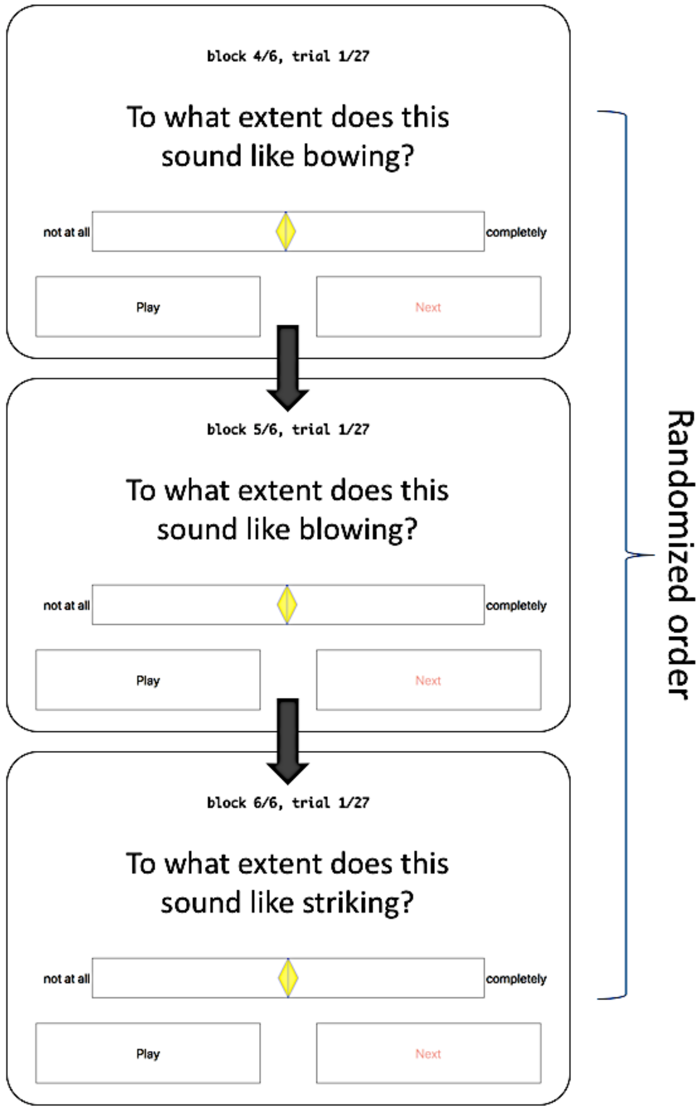


**Figure 3.1.** Summary of the experimental design and procedure of Experiment 1.

## 3.2 Results

During the experiment, participants rated how well each stimulus resembled bowing, blowing, and striking actions, giving three dependent variables. We averaged each action-resemblance rating across the three exemplars for each action-object pair for each participant. We conducted a 2×3×3 mixed-measures Multivariate Analysis of Variance (MANOVA). The three independent variables included: (1) a between-groups factor of musicianship with two levels that included musicians and nonmusicians; (2) a repeated-measures factor of action properties with three levels—bowing, blowing, and striking; and (3) a repeated-measures factor of object properties with three levels—string, air column, and plate. Univariate analyses were also computed by MANOVA, which we reported for each dependent variable.

### 3.2.1 Multivariate analyses

Box's Test of Equality of Covariance Matrices was not computed when we ran the statistical analyses on SPSS (IBM Corp, 2015). The warning indicated that "there are fewer than two nonsingular cell covariance matrices". Box's M tests the null hypothesis that the dependent variables' covariance matrices are equal across groups. Moreover, the result of Box's M indicates which multivariate test statistic—Pillai's Trace, Wilk's Lambda, Hotelling's Trace, or Roy's Largest Root—should be used in order to identify the *F* statistic. The reason Box's M was not computed is most likely because more than two levels of the independent variables shared at least 90% of the ratings on the dependent variables. Consequently, we will use Pillai's Trace, *V*, since it has been reported to be more robust to violations of assumptions (Olson, 1974).

The between-groups effect of musicianship on the bowing, blowing, and striking resemblance ratings was statistically significant, $V$=.39, $F(3,37)$=7.97, $p$<.001, $\eta_p^2$=.39. Figure 3.2 demonstrates that, overall, musicians have lower bowing ratings than nonmusicians. Musicians were less likely to perceive bowing actions than nonmusicians. This might be attributed to the relatively poor quality of the bow model, since musicians have more experience with hearing bowed strings. The within-groups effect of the actions that produce the stimuli was statistically significant, $V$=1.44, $F(6,154)$=65.15, $p$<.001, $\eta_p^2$=.72. This effect demonstrated that bowing ratings were highest for bowed stimuli, blowing ratings were highest for blown stimuli, and striking ratings were highest for struck stimuli (Fig. 3.3). However, bowing was more ambiguous in terms of the type of excitation it was perceived to be. Thus, the actions that produced the stimuli influenced how listeners rated them. The within-groups effect of the objects that produced the stimuli was statistically significant, $V$=1.55 $F(6,154)$=89.32, $p$<.001, $\eta_p^2$=.78. Bowing ratings were highest for string stimuli, blowing ratings were highest for air column stimuli, and striking ratings were highest for plate stimuli (Fig. 3.4). The object that produced the stimuli influenced participants' perceptions of the actions that produced the stimuli.

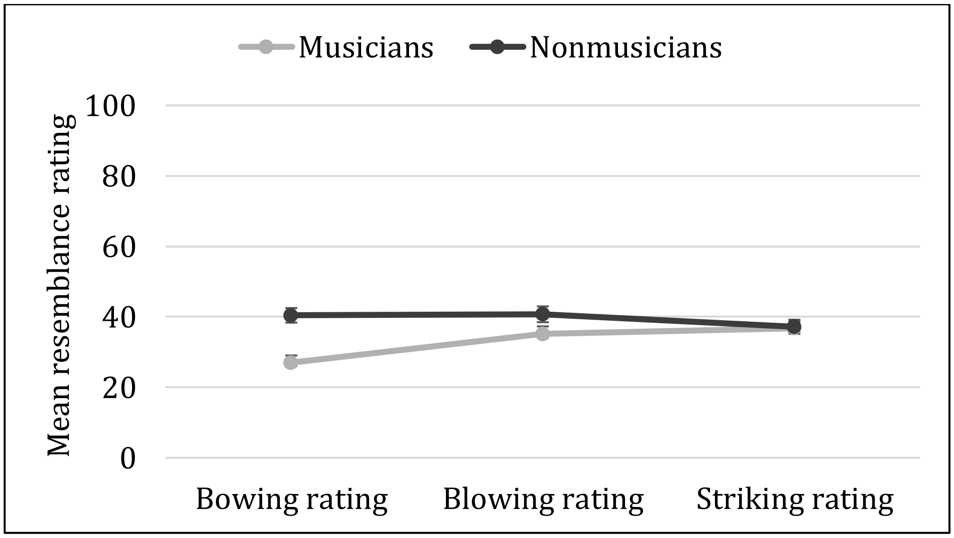


**Figure 3.2.** Musicians' and nonmusicians' mean bowing, blowing, and striking resemblance ratings. Error bars indicate standard error of the mean.

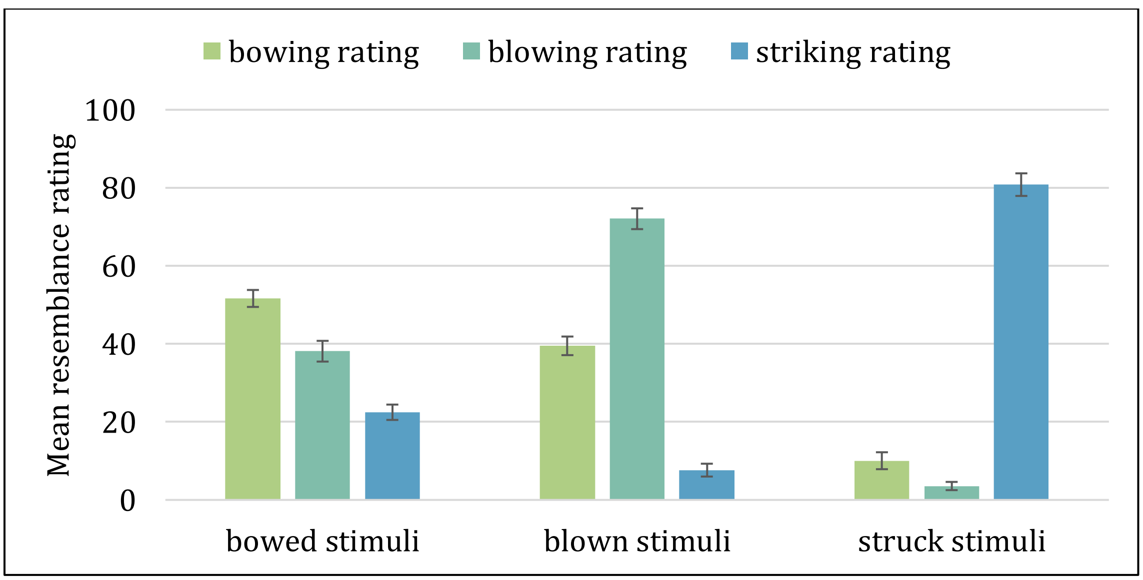


**Figure 3.3.** Mean bowing, blowing, and striking resemblance ratings based on actions. Error bars indicate standard error of the mean.

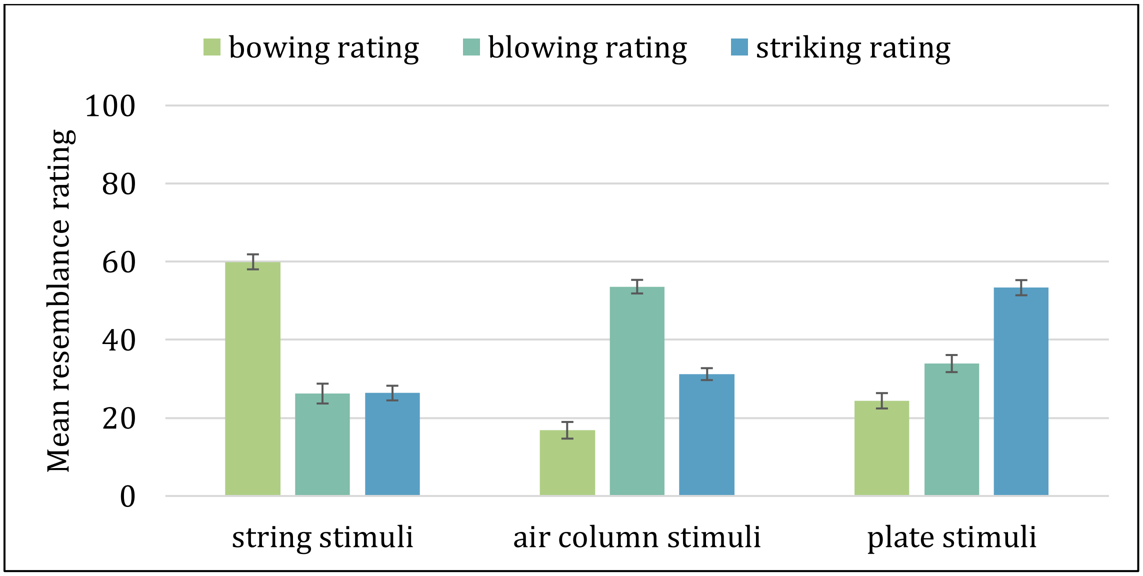


**Figure 3.4.** Mean bowing, blowing, and striking resemblance ratings based on objects. Error bars indicate standard error of the mean.

The interaction of musicianship with action properties on the three types of action resemblance ratings was not statistically significant, *V*=.13, *F*(6,154)=1.83, *p*=.097, $\eta_p^2$=.07. This means that musicians' and nonmusicians' action ratings were similar across the different action properties. The interaction between musicianship and object properties on the three types of action ratings was significant, *V*=.20, *F*(6,154)=2.78, *p*=.014, $\eta_p^2$=.10 (Fig. 3.5). Musicians and nonmusicians rated action resemblance differently depending on the types of objects that produced the sounds. The interaction of action and object properties was statistically significant, *V*=1.02, *F*(12,468)=20.14, *p*<.001, $\eta_p^2$=.34. Bowing, blowing, and striking resemblance ratings differed depending on the type of action-object interaction (Fig. 3.6). This means that participants perceived certain action-object interactions as resembling certain actions more than others. Bowed air columns were more likely perceived as blown. Bowed plates were more likely perceived as struck than bowed or blown. Blown strings were more likely perceived as bowed than blown. Struck stimuli were easily distinguished from continuous sounds. Lastly, the combined effect of musicianship, action properties, and object properties on the three action resemblance ratings was statistically significant, *V*=.15, *F*(12,468)=1.98, *p*=.024, $\eta_p^2$=.05. Thus, musicians and nonmusicians perceived the resemblance of actions differently for different action-object interactions (Fig. 3.7). Nonmusicians were more likely than musicians to perceive bowed strings as blown, blown plates as bowed, and struck strings as bowed. Moreover, nonmusicians were more confused by the type of action that produced bowed plate stimuli, whereas musicians were more likely to perceive them as struck than bowed or blown.

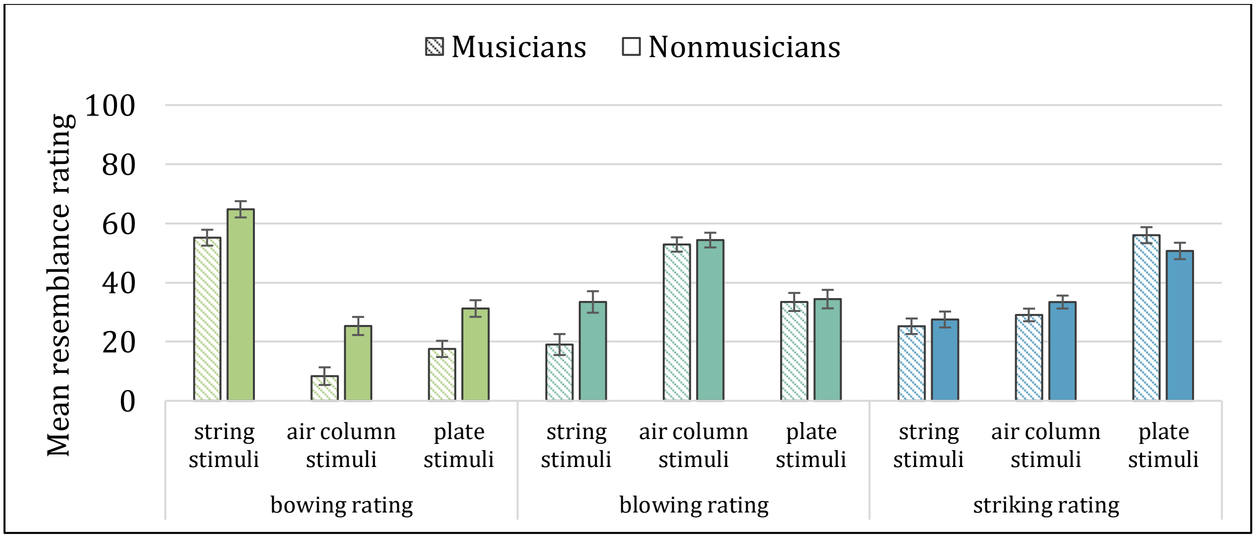


**Figure 3.5.** Musicians' and nonmusicians' mean bowing, blowing, and striking resemblance ratings for each object property. Different colours represent different action-resemblance ratings. Solid bars represent nonmusicians and patterned bars represent musicians. Error bars represent standard error of the mean.

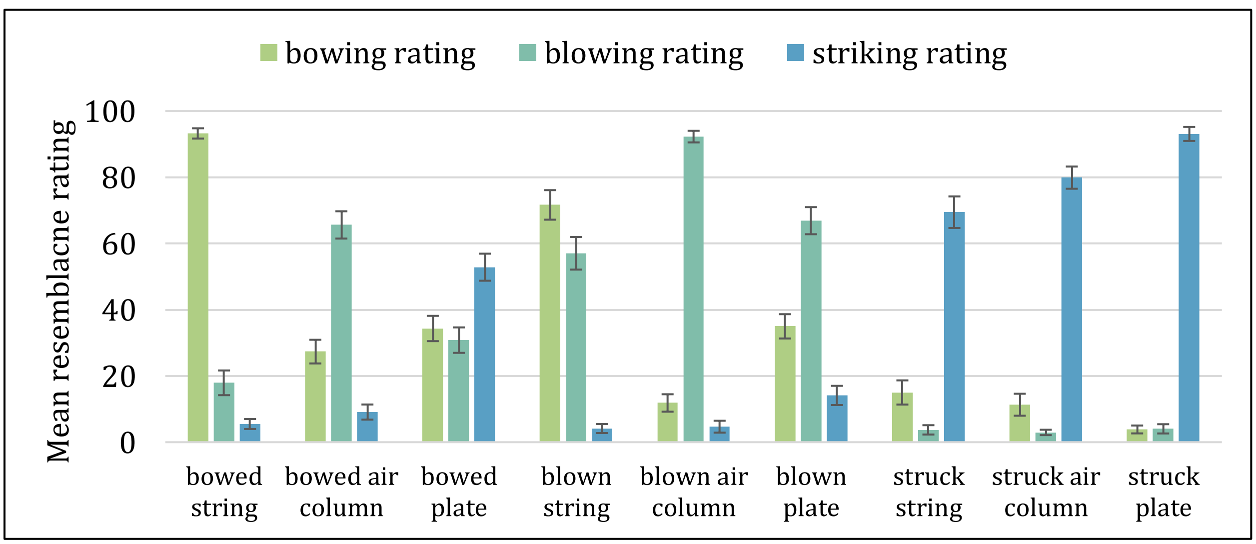


**Figure 3.6.** Mean bowing, blowing, and striking resemblance ratings for each action-object interaction. Error bars represent standard error of the mean.

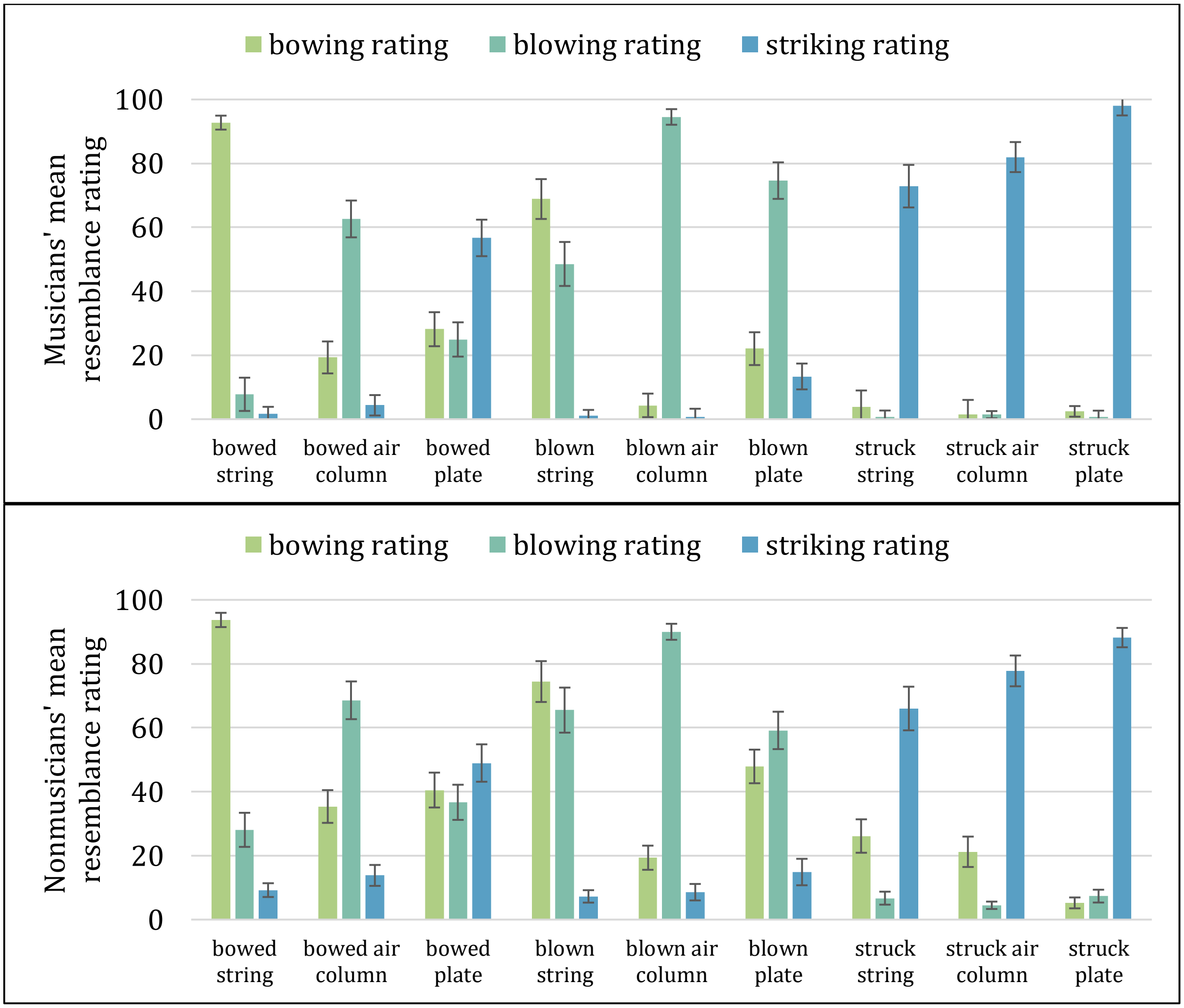


**Figure 3.7.** Mean action-resemblance ratings of musicians (top) and nonmusicians (bottom) for each action-object interaction. Error bars represent standard error of the mean.

### 3.2.2 Univariate analyses: Bowing resemblance ratings

For analyses involving repeated-measures factors, Mauchly's Test revealed that the sphericity assumption was not violated for the bowing resemblance ratings with respect to repeated-measures factors of: action properties, $p$=.524; and object properties, $p$=.244. However, the sphericity assumption was violated with respect to the interaction between action and object properties, $p$=.031. Consequently, we reported a conservative adjustment of the $F$ statistic where appropriate.

#### *3.2.2.1 Main effects*

The main effect of musicianship on the bowing resemblance ratings was statistically significant, $F(1,39)=21.50$, $p<.001$, $\eta_p^2=.36$. Musicians had a lower mean bowing rating for the stimuli compared to nonmusicians, 27.02 (*SD*=35.50) and 40.43 (*SD*=35.13), respectively.

The different types of actions that produced the stimuli had a significant effect on the bowing ratings, $F(2,78)=102.92$, $p<.001$, $\eta_p^2=.73$ (Fig. 3.8). The mean ratings were 51.64 (*SD*=36.19) for a bowing action, 39.49 (*SD*=34.94) for a blowing action, and 10.04 (*SD*=20.83) for a striking action. We used a Bonferroni-corrected post-hoc pairwise comparison to assess differences between the mean bowing ratings for the different action properties. *Z* represents the mean difference in resemblance ratings in absolute value. The difference between the mean ratings was significant for: bowing and blowing, $Z=12.15$, $p<.001$; bowing and striking, $Z=41.59$, $p<.001$; and blowing and striking, $Z=29.45$ $p<.001$. Thus, bowing was more often confused with blowing than striking. Bowing and blowing are continuous excitations, which might explain the confusion.

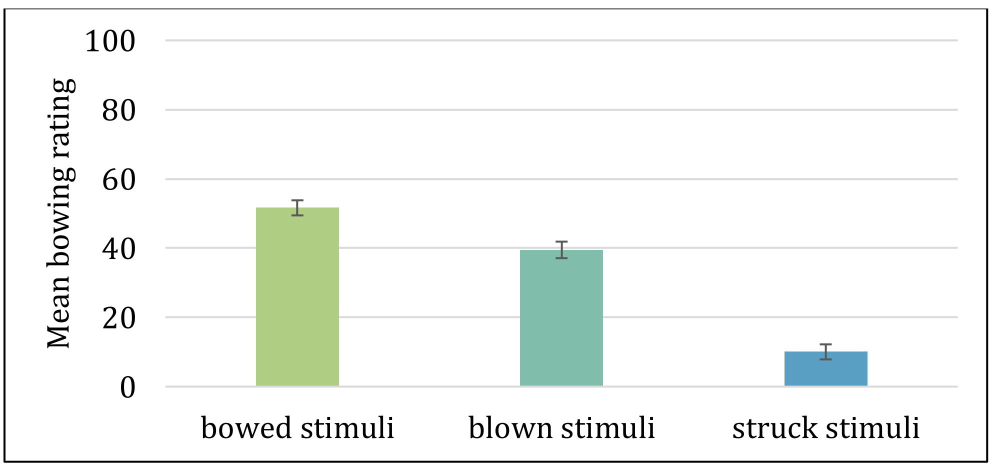


**Figure 3.8.** Main effect of action properties on bowing resemblance ratings. Error bars indicate standard error of the mean.

The univariate analysis also revealed a significant main effect of the different objects that produced the stimuli on participants' bowing ratings, $F(2,78)=180.83$, $p<.001$, $\eta_p^2=.82$ (Fig. 3.9). The mean bowing rating was 59.97 (*SD*=40.21) when a string produced the sound, 16.83 (*SD*=23.08) when an air column produced the sound, and 24.38 (*SD*=25.80) when a plate produced the sound. Pairwise comparisons revealed a significant difference between the mean bowing ratings for: strings and air columns, $Z=43.14$, $p<.001$; strings and plates, $Z=35.59$, $p<.001$; and air columns and plates, $Z$=7.56, $p$=.003. These results suggest that strings were more often perceived as bowed than air columns and plates.

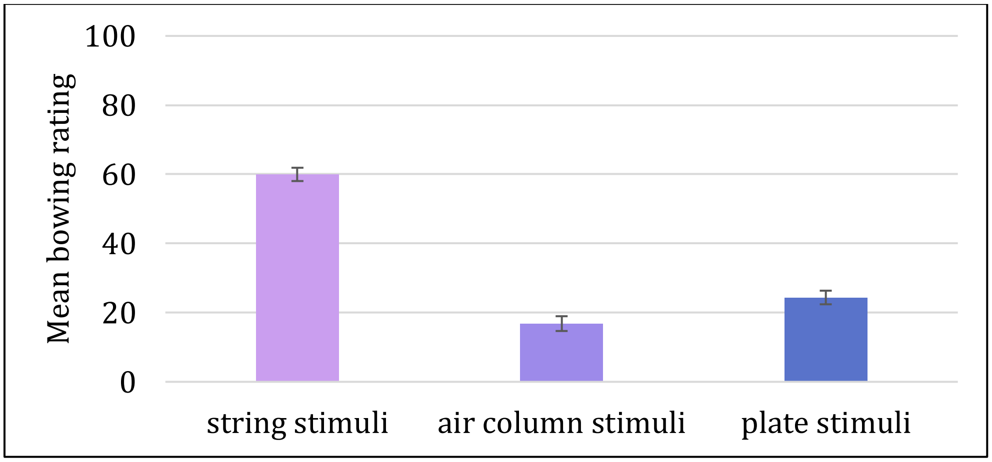


**Figure 3.9.** Main effect of object properties on bowing resemblance ratings. Error bars indicate standard error of the mean.

***3.2.2.2 Interaction effects***

The univariate analyses revealed two non-significant interactions between musicianship and the action properties of the stimuli, $F$<1, and between musicianship and object properties, $F(2,78)=1.16$, $p=.320$, $\eta_p^2=.03$. These results show that musicians and nonmusicians did not rate bowing resemblance differently depending on the action properties or the object properties.

The interaction effect between action and object properties on mean bowing ratings, however, was statistically significant, $F(3.33,129.94)=40.22$, $p<.001$, $\varepsilon=.83$, $\eta_p^2=.51$ (Fig. 3.10). Continuously excited (i.e., bowed and blown) strings were more likely to be perceived as bowed. Continuously excited plates and bowed air columns were less likely to be perceived as bowed. Blown air columns and all struck stimuli were not perceived as bowed. A nonparametric simple effects analysis is reported in Section 3.2.5 to examine the difference between mean bowing ratings for the actions across each level of the object properties and vice versa. Lastly, the three-way interaction effect of musicianship, action properties, and object properties on mean bowing ratings was statistically significant, $F(3.33,129.94)=3.60$, $p=.012$, $\varepsilon=.83$, , $\eta_p^2=.09$ (Fig. 3.11). Compared to musicians, nonmusicians' bowing ratings were higher for struck strings and air columns. Nonmusicians' ratings were also higher than those of musicians for blown plates. Moreover, nonmusicians were more likely to perceive bowed air columns as bowed than were musicians.

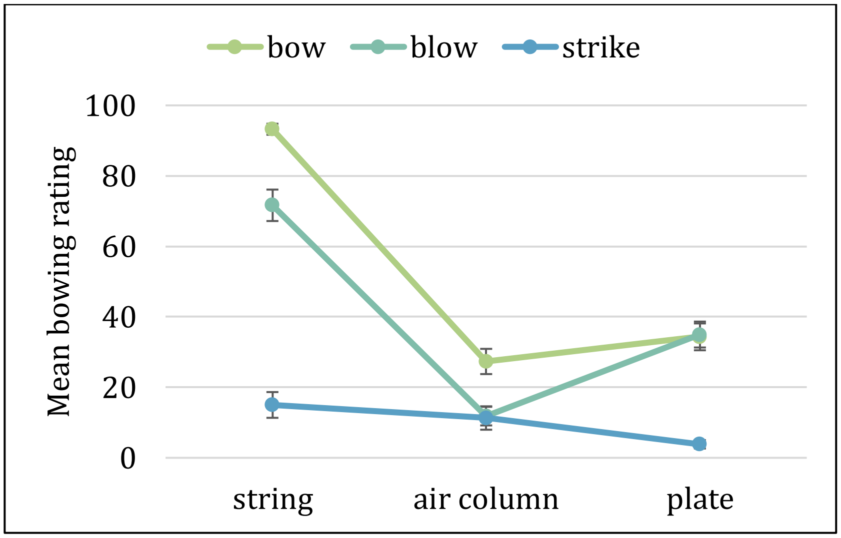


**Figure 3.10.** Interaction effect of combined actions and objects on bowing resemblance ratings. Each coloured line represents a different action and the objects are on the horizontal axis. Error bars indicate standard error of the mean.

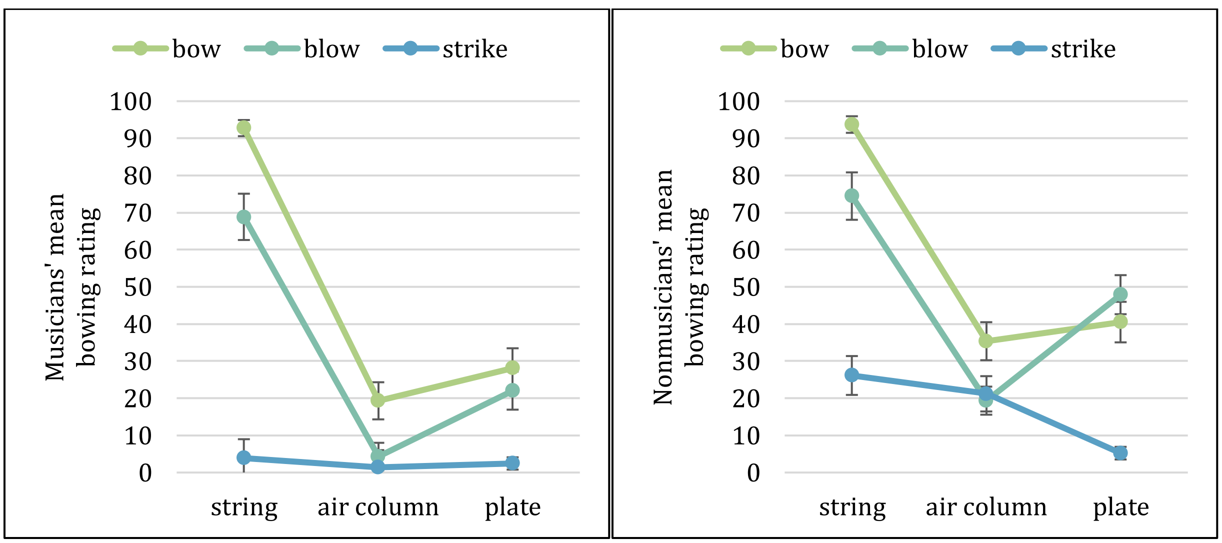


**Figure 3.11.** Interaction effect of musicians' (left) and nonmusicians' (right) mean bowing ratings for different action-object combinations. Error bars represent standard error of the mean.

### 3.2.3 Univariate analyses: Blowing resemblance ratings

Mauchly's test of sphericity was significant for the repeated-measures effects of action properties, $p$=.027, and object properties, $p$=.048, as well as the interaction between actions and objects, $p$=.046. Adjusted degrees of freedom are reported where appropriate.

#### *3.2.3.1 Main effects*

The main effect of musicianship on the blowing ratings was not significant, $F(1,39)=3.21$, $p=.081$, $\eta_p^2=.08$, meaning there was no reliable difference between musicians' and nonmusicians' mean blowing ratings. The mean blowing ratings were 35.10 (*SD*=38.51) for musicians and 40.73 (*SD*=37.25) for nonmusicians.

The main effect of action properties on the blowing ratings was statistically significant, $F(1.70,66.46)=289.30$, $p<.001$, $\varepsilon=.85$, $\eta_p^2=.88$ (Fig. 3.12). The mean blowing ratings were 38.11 (*SD*=32.54) for bowed stimuli, 72.07 (*SD*=29.10) for blown stimuli, and 3.56 (*SD*=8.26) for struck stimuli. Pairwise comparisons revealed that differences in mean blowing ratings were statistically significant for: bowing and blowing, $Z=33.96$, $p<.001$; bowing and striking, $Z=34.55$, $p<.001$; and blowing and striking, $Z=68.51$, $p<.001$. Blowing ratings were highest for blown stimuli, but blowing was occasionally confused with bowing, likely because blowing and bowing are continuous actions.

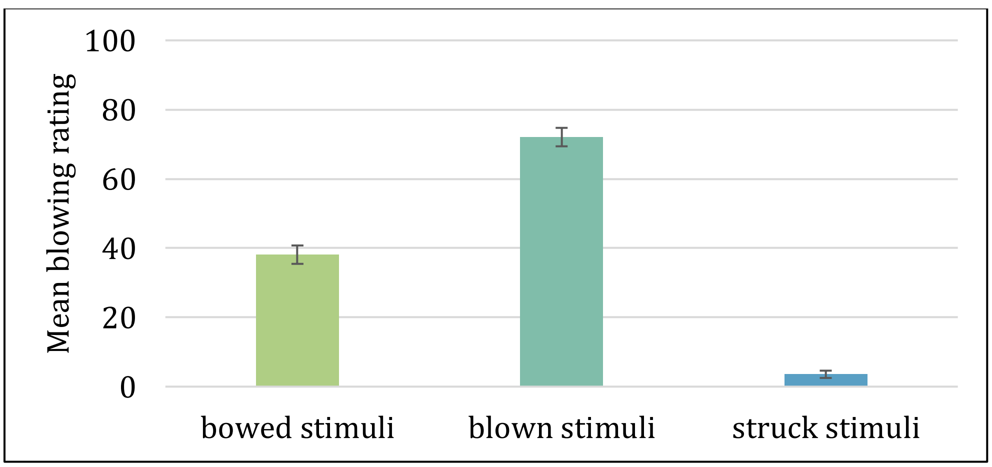


**Figure 3.12.** Main effect of action properties on blowing resemblance ratings. Error bars indicate standard error of the mean.

Univariate analyses revealed a main effect of object properties on blowing ratings, $F(1.74,67.98)=58.03$, $p<.001$, $\varepsilon=.87$, $\eta_p^2=.60$ (Fig. 3.13). The mean blowing ratings were 26.22 (*SD*=33.14), 53.61 (*SD*=41.14), and 33.91 (*SD*=33.89) when a string, air column, and plate produced the sound, respectively. Post-hoc pairwise comparisons revealed a significant difference in the mean blowing ratings for: string and air column, $Z=27.39$, $p<.001$; string and plate, $Z=7.68$, *p*=.023; and air column and plate, *Z*=19.70, *p*<.001. These results reveal that air columns bias participants into perceiving blowing more than strings and plates do.

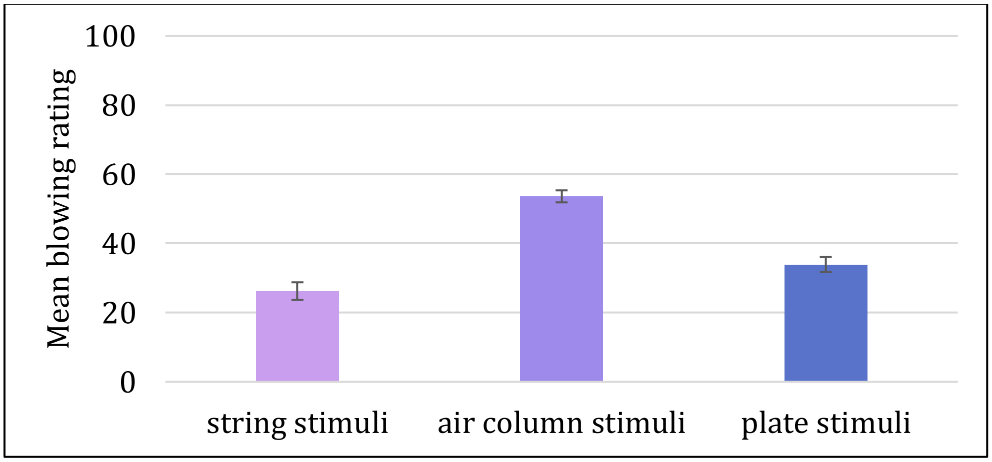


**Figure 3.13.** Main effect of object properties on blowing resemblance ratings. Error bars indicate standard error of the mean.

### *3.2.3.2 Interaction effects*

A two-way interaction effect between musicianship and action properties was not statistically significant, $F(1.70,66.46)=2.98$, $p=.066$, $\varepsilon=.85$, $\eta_p^2=.07$. Blowing ratings for different action properties were not different between musicians and nonmusicians. The two-way interaction effect between musicianship and object properties was statistically significant, $F(1.74,67.98)=4.21$, $p=.023$, $\varepsilon=.87$, $\eta_p^2=.10$ (Fig. 3.14). We conducted further analyses to test the simple effect of object properties for both musicians and nonmusicians, which were significant: $F(2,78)=42.99$, $p<.001$, and $F(2,78)=19.81$, $p<.001$, respectively. Additionally, we tested the simple effect of musicianship, which was significant for strings, $F(1,39)=8.03$, $p=.007$; but not for air columns, $F<1$, or plates, $F<1$. These results indicate that nonmusicians were more likely than musicians to perceive strings as blown.

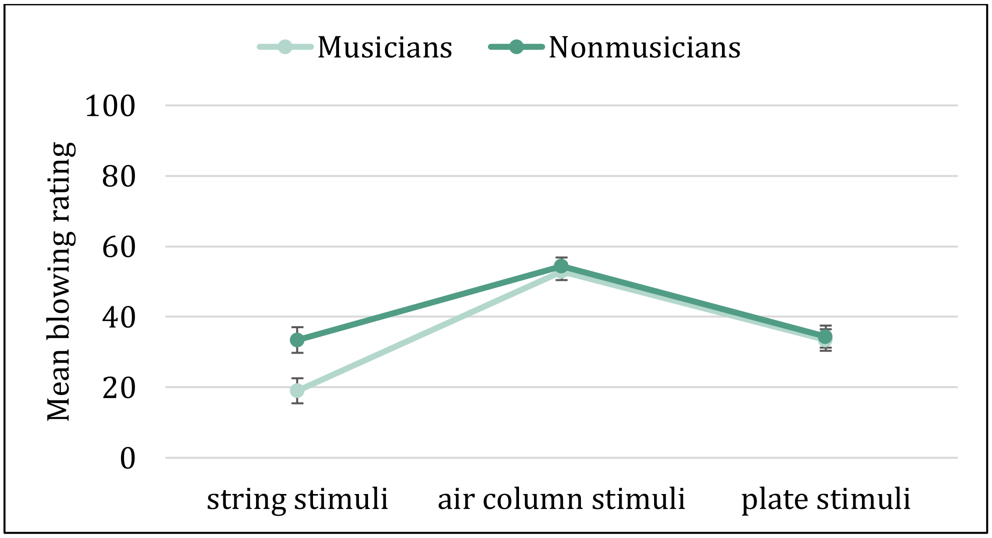


**Figure 3.14.** Interaction effect of musicianship and object properties on mean blowing resemblance ratings. Error bars represent standard error of the mean.

The interaction effect between action and object properties was significant, $F(3.45,134.39)=24.49$, $p<.001$, $\varepsilon=.86$, $\eta_p^2=.38$ (Fig. 3.15). Blown stimuli and bowed air columns were most likely perceived as blown. Bowed plate and bowed string stimuli were more likely perceived as blown than struck stimuli, which had the lowest blowing ratings. A nonparametric simple effects analysis is reported in Section 3.2.5 to observe specific differences in the ratings between different levels of action and object properties. Lastly, a significant three-way interaction was found between musicianship, action properties, and object properties, $F(3.45,134.39)=2.76$, $p=.037$, $\varepsilon=.86$, $\eta_p^2=.07$ (Fig. 3.16). We performed a simple effects analysis to test for differences between the interactions of action and object properties for each level of musicianship. For both groups, the interaction between action and object properties was significant: $F(4,156)=16.72$, $p<.001$ for musicians, and $F(4,156)=10.68$, $p<.001$ for nonmusicians. This means that both musicians' and nonmusicians' blowing ratings differed depending on the combination of actions and objects that produced the stimuli. Notably, nonmusicians' mean blowing ratings were higher for bowed string and bowed plate stimuli than those of musicians, demonstrating that nonmusicians were tricked by the continuous nature of bowing. Nonmusicians' blowing ratings were also higher than musicians' ratings for blown string stimuli, which suggests that the string might have biased musicians into thinking that the sounds were not blown.

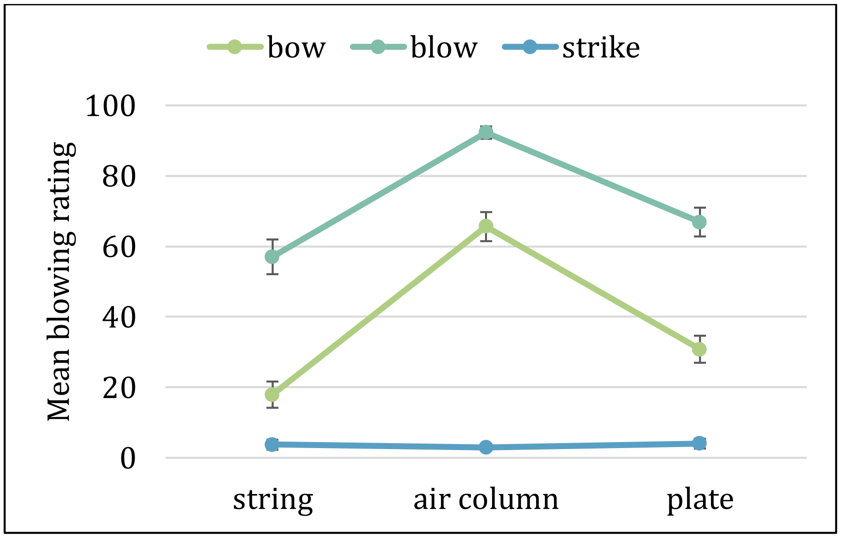


**Figure 3.15.** Interaction effect of combined actions and objects on blowing resemblance ratings. Each coloured line represents a different action and the objects are on the horizontal axis. Error bars indicate standard error of the mean.

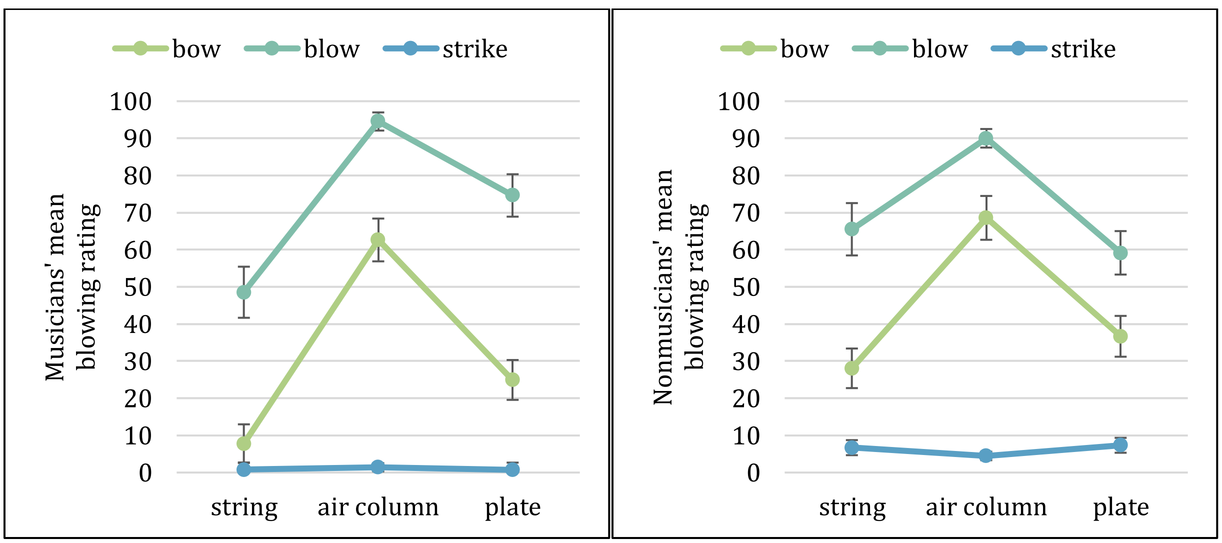


**Figure 3.16.** Interaction effect of musicians' (left) and nonmusicians' (right) mean blowing ratings for different action-object combinations. Error bars represent standard error of the mean.

### 3.2.4 Univariate analyses: Striking resemblance ratings

Univariate analyses were conducted for the striking resemblance ratings. Mauchly's test of sphericity was violated for the within-groups effects of action properties, $p<.001$, object properties, $p=.001$, and their interaction, $p<.001$.

### *3.2.4.1 Main effects*

Univariate analyses revealed that the main effect of musicianship is not significant, $F<1$. The mean striking rating of musicians was 36.75 ($SD$=40.98), and that of nonmusicians was 37.19 ($SD$=37.51), which are not reliably different from one another.

The main effect of action properties was statistically significant, $F(1.28,49.81)=329.56$, $p<.001$, $\varepsilon=.64$, $\eta_p^2=.89$ (Fig. 3.17). The mean striking ratings were 22.47 ($SD$=28.37) for bowed stimuli, 7.64 ($SD$=14.36) for blown stimuli, and 80.81 ($SD$=24.74) for struck stimuli. Pairwise comparisons further revealed a significant difference between mean striking ratings of: bowed and blown stimuli, $Z=14.83$, $p<.001$; bowed and struck stimuli, $Z=58.34$, $p<.001$; and blown and struck stimuli, $Z=73.17$, $p<.001$. Struck stimuli were commonly rated highest as resembling striking actions. Given that struck stimuli were impulsive rather than continuous, this made it easier for participants to perceive striking over bowing and blowing.

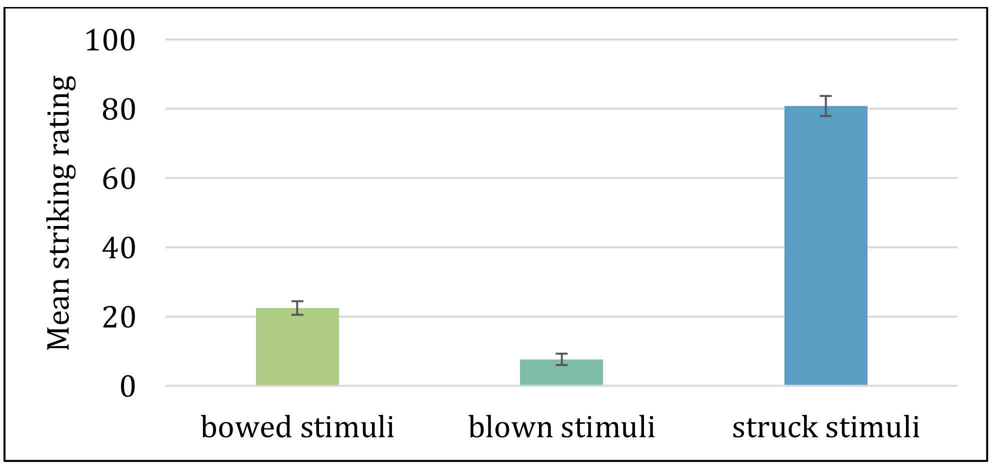


**Figure 3.17.** Main effect of action properties on striking resemblance ratings. Error bars indicate standard error of the mean.

We found a significant main effect of object properties on striking ratings, $F(1.53,59.56)=110.73$, $p<.001$, $\varepsilon=.76$, $\eta_p^2=.74$ (Fig. 3.18). The mean striking ratings were 26.36 ($SD$=36.12) for stimuli produced by a string, 31.21 ($SD$=38.41) for those produced by an air column, and 53.34 ($SD$=38.10) for those produced by a plate. Pairwise comparisons revealed significant differences for: string and air column stimuli, $Z=4.86$, $p=.002$; string and plate stimuli, $Z=26.99$, $p<.001$; and air column and plate stimuli, $Z=22.13$, $p<.001$. Plates were rated as resembling striking more often than air columns and strings, which suggest a bias to perceive plates as struck.

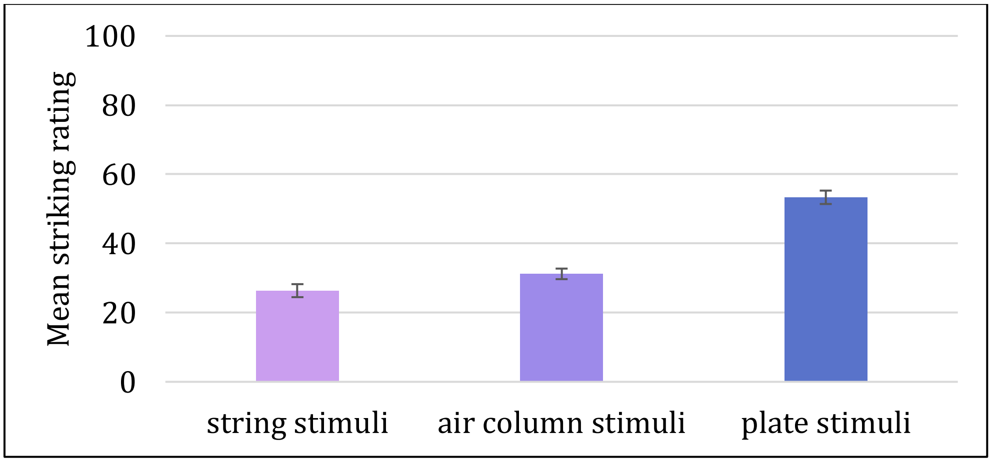


**Figure 3.18.** Main effect of object properties on striking resemblance ratings. Error bars indicate standard error of the mean.

### *3.2.4.2 Interaction effects*

Univariate analyses revealed that the interaction between musicianship and action properties on striking ratings was not statistically significant, $F(1.28,49.81)=2.31$, $p=.130$, $\varepsilon=.64$, $\eta_p^2=.06$. Musicians and nonmusicians did not rate the resemblance of a striking action differently when stimuli were produced by different actions. On the other hand, the interaction between musicianship and object properties was statistically significant, $F(1.53,5956)=3.48$, $p=.049$, $\varepsilon=.76$, $\eta_p^2=.08$ (Fig. 3.19).We found a simple effect of object properties. Both musicians and nonmusicians rated striking resemblance differently depending on the object that produced the sound, $F(2,78)=77.26$, $p<.001$, and $F(2,78)=37.91$, $p<.001$, respectively. However, the simple effect of musicianship was not significant for either of the object properties: $F<1$ for strings; $F(1,39)=2.04$, $p=.162$ for air columns; and $F(1,39)=1.89$, $p=.177$ for plates.

The two-way interaction between action and object properties was statistically significant, $F(2.59,101.16)=24.55$, $p<.001$, $\varepsilon=.65$, $\eta_p^2=.39$ (Fig. 3.20). Struck stimuli were fairly well judged as struck, with highest striking ratings for struck plates, followed by struck air columns, then struck strings. Listeners rarely rated bowed and blown stimuli as struck, with the exception of the bowed plate. The bowed plate contains an artifact that is produced by the bow initially touching the plate, which might have been perceived as a strike. A simple effects analysis is presented in Section 3.2.5

to further examine this interaction. Lastly, the three-way interaction between musicianship, action properties, and object properties was not statistically significant, *F*<1. Musicians and nonmusicians do not assign different striking ratings across the different action-object interactions.

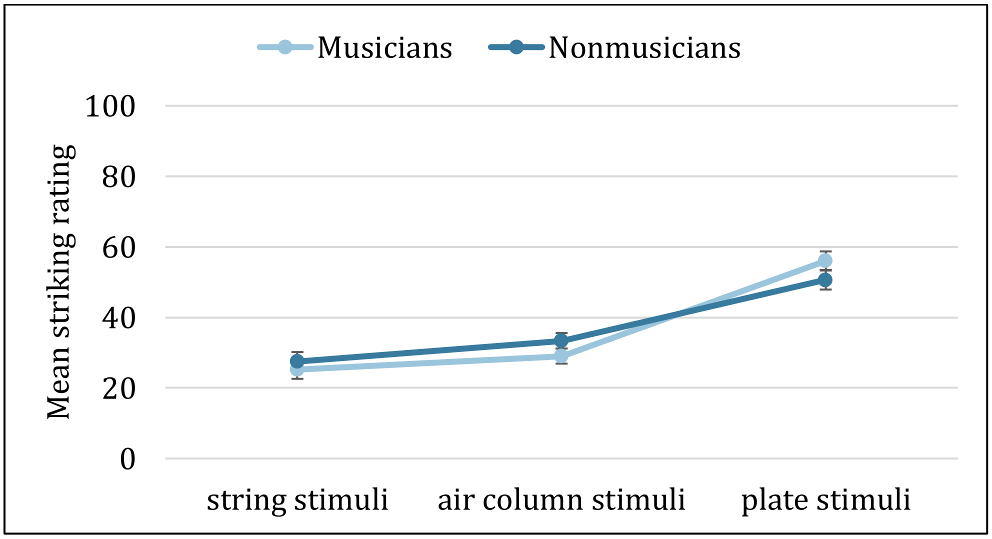


**Figure 3.19.** Interaction effect of musicianship and object properties on mean striking resemblance ratings. Error bars represent standard error of the mean.

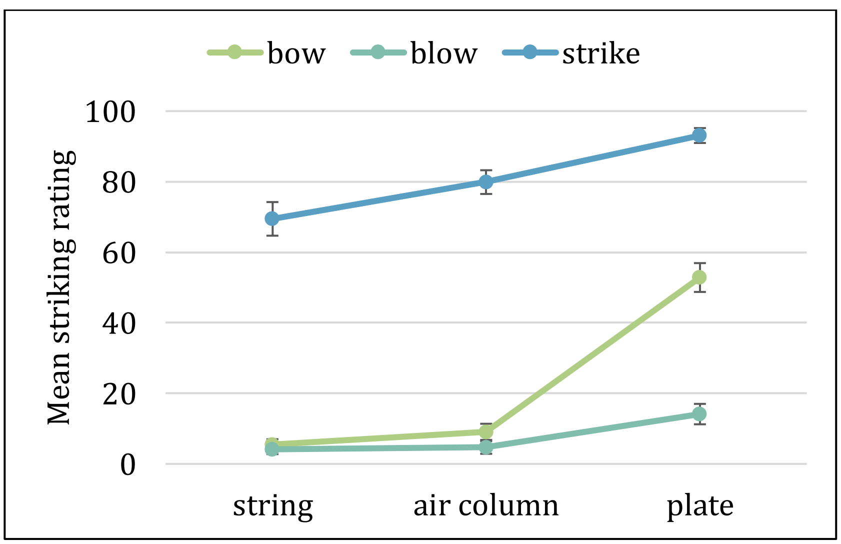


**Figure 3.20.** Interaction effect of combined actions and objects on striking resemblance ratings. Each coloured line represents a different action and the objects are on the horizontal axis. Error bars indicate standard error of the mean.

### 3.2.5 Simple effects analyses: Interaction between actions and objects

When conducting the univariate analyses, Levene's test of homogeneity of variance was significant for multiple bowing, blowing, and striking resemblance ratings of the stimuli. This means that the variances were significantly different. Consequently, we assessed the normality of the residuals of the MANOVA (Kozak & Piepho, 2018). The Shapiro-Wilk test of normality was significant, $p<.001$, meaning that the normality assumption was violated for the residuals. Friedman's test is the nonparametric equivalent of a one-way repeated-measures ANOVA. We assessed the significant interaction effects of action and object properties for the three action-resemblance ratings with Friedman's test to examine simple effects of: (1) action properties across each type of object and (2) object properties across each type of action. This resulted in six computations of Friedman's test for each type of action resemblance rating; since there were three action-resemblance ratings, a total of 18 tests were performed. With multiple analyses, there is a greater risk of type I errors. Thus, we applied the Bonferroni-Holm correction to control the family-wise error rate for multiple hypothesis tests (Abdi, 2010).

#### *3.2.5.1 Bowing resemblance ratings*

Six Friedman's tests were conducted to analyze two simple effects of the interaction between action and object properties on bowing ratings (Fig. 3.10). The first was the simple effect of actions at each type of object on bowing resemblance ratings. The test revealed a significant simple effect of action properties when the object that produced the stimuli was a string, $\chi^2(2, N=41)=58.50$, $p<.001$. Bowing ratings were highest for the bowed string, but listeners often confused the blown string for bowing as well. This suggests that continuous actions were confused with one another. A low bowing rating for struck string stimuli supports this speculation, since striking actions are impulsive. Friedman's test was significant for the simple effect of action properties when the object was an air column, $\chi^2(2, N=41)=21.75$, $p=.001$. Participants were more likely to perceive bowing when they heard bowed air columns compared to blown and struck air columns; however, the mean bowing rating was still quite low for bowed air columns. Participants might have been limited by their mental models, which are weak for bowed air columns. It is also a possibility that Modalys was not as effective at conveying bowing when it was applied to an air column. The simple effect of action properties for plate stimuli was statistically significant, $\chi^2(2, N=41)=45.65$, $p<.001$, as revealed by Friedman's test. Bowing ratings were almost the same for bowed and blown plates and very low for struck plates. Struck plates are impulsive and represent a typical action-object interaction, so it might have been obvious that these sounds were not bowed. On the other hand, the similarity between bowing ratings for bowed and blown plates indicate that participants confused continuous actions, especially when both of them were applied to an object they do not typically interact with.

The other simple effect of interest was that of objects at each type of action on bowing ratings. For bowed stimuli, there was a significant simple effect of object properties on bowing ratings, $\chi^2(2, N=41)=60.09$, $p<.001$. Bowing ratings were highest for bowed strings, which is expected since they represent a typical action-object interaction. Bowing ratings were much lower for bowed air columns and plates, suggesting that these sounds did not match listeners' mental models or Modalys did not convey bowing effectively. The simple effect of object properties was statistically significant, $\chi^2(2, N=41)=58.08$, $p<.001$, for blown stimuli. Ratings were lowest for blown air columns and plates. This implies that listeners were able to pick apart the blowing action and distinguish it from bowing for these particular interactions. However, their ratings were also highest for blown strings: the string object biased listeners into perceiving that the blown strings were bowed. The simple effect was non-significant for object properties when the action was striking, $\chi^2(2, N=41)=6.32$, $p=.041$. Bowing ratings for struck stimuli were very low, indicating that participants were able to distinguish struck from bowed sounds. Among these low ratings, however, they were highest for struck strings: the string may have mildly biased participants into perceiving struck strings as bowed, at least more so than struck air columns and plates.

#### *3.2.5.2 Blowing resemblance ratings*

We conducted simple effects analyses to further investigate the interaction effect between action and object properties on blowing ratings (Fig. 3.15). We first conducted Friedman's tests of the simple effects of actions at each type of object property on blowing ratings. Friedman's test revealed a significant simple effect of action properties for the string on blowing ratings, $\chi^2(2, N=41)=59.89$, $p<.001$. Blowing ratings were highest for blown strings and lowest for struck strings. Bowed strings were more likely to be perceived as blown than struck strings, but blowing ratings were low overall for these stimuli. The simple effect of action properties for stimuli produced by an air column was statistically significant, $\chi^2(2, N=41)=76.78$, $p<.001$. Blown air columns were most likely perceived as blown, as expected. Struck air columns were not perceived as blown, likely due to the impulsive nature of these sounds. Blowing ratings for bowed air column stimuli were closer to those of their blown counterparts than of their struck counterparts, suggesting that the air column with a continuous action biased participants into hearing bowed air columns as blown. When the object was a plate, there was a significant simple effect of action properties on the blowing ratings, $\chi^2(2, N=41)=69.35$, $p<.001$. Blown plates were most likely perceived as blown and struck plates were least likely perceived as blown. Bowed plate stimuli, however, were more likely perceived as blown than struck plate stimuli. This was likely due to the continuous nature of bowed sounds.

Next, we conducted simple effects analyses of objects at each type of action on the blowing ratings. Friedman's test revealed a significant simple effect of object properties when stimuli were produced by bowing, $\chi^2(2, N=41)=48.04$, $p<.001$. Bowed air columns were most likely perceived as blown, but bowed plates and strings were less likely perceived as blown. There seems to be a bias that bowed stimuli were more likely perceived as blown, as they are both continuous actions. A significant simple effect of objects was revealed when stimuli are blown, $\chi^2(2, N=41)=43.92$, $p<.001$. Blown air columns were most likely perceived as blown, followed by blown plate stimuli, then blown string stimuli. Globally, listeners were able to perceive blown actions, with a clear bias for air columns. The simple effect of object properties for struck stimuli was not statistically significant, $\chi^2(2, N=41)=.53$, $p=.782$, meaning that the mean blowing ratings were similar for all struck stimuli. Struck sounds had very low blowing ratings, implying that listeners were able to make a distinction between continuous and impulsive sounds.

#### *3.2.5.3 Striking resemblance ratings*

To explain the interaction effect between combined actions and objects on striking ratings, we conducted simple effects analyses (Fig. 3.20). A simple effects analysis was conducted for action properties at each type of object on striking ratings. Friedman's test revealed that for stimuli produced by a string, there was a significant simple effect of action properties on the striking ratings, $\chi^2(2, N=41)=57.37$, $p<.001$. Struck strings were most likely perceived as struck and continuously excited strings were not perceived as struck. A significant main effect of action properties was revealed when the stimuli were produced by an air column, $\chi^2(2, N=41)=52.76$, $p<.001$. Struck air columns had the highest striking ratings. Continuously excited air columns had very low striking ratings. Listeners were able to distinguish between impulsive and continuous actions. Friedman's test revealed a significant simple effect of action properties for plate stimuli, $\chi^2(2, N=41)=74.62, p<.001$. Struck plates had the highest striking ratings and blown plates had the lowest striking ratings. Bowed plates were occasionally perceived as struck, likely due to the impulsive artifact during the beginning of the sound.

The simple effect of objects at each type of action was analyzed for the striking ratings. Friedman's test revealed that the simple effect of objects was significant for bowed stimuli, $\chi^2(2, N=41)=56.06, p<.001$. Bowed strings and air columns were rarely perceived as struck. Bowed plates were occasionally perceived as struck. Again, this might be attributed to the artifact produced during the beginning of the sound. Friedman's test revealed a significant simple effect of object properties on blown stimuli, $\chi^2(2, N=41)=21.71, p<.001$. None of the blown stimuli were perceived as struck: impulsive sounds were easily distinguished from continuous sounds. The simple effect of object properties for struck stimuli was significant, $\chi^2(2, N=41)=29.78$. Globally, struck stimuli were likely perceived as struck. This highlights the obvious distinction between impulsive and continuous sounds.

## 3.3 Discussion

We initially hypothesized that combinations between different excitations and resonators would influence how listeners perceived the actions that produced the sounds. This was examined by having participants listen to 27 stimuli composed of nine excitation-resonator combinations and rate each of them based on how well they resemble three given actions: bowing, blowing, and striking.

We found main effects of action properties and object properties for each action-resemblance rating. Listeners' bowing ratings differed depending on the types of actions and objects that produced the sound. The same phenomenon applies to the blowing and striking ratings. For the main effect of action properties, mean resemblance ratings were highest when the action matched the one that participants were rating (Figs. 3.8 for bowing, 3.12 for blowing, and 3.17 for striking ratings). This finding confirms our sensitivity to the actions of sounds (Lemaitre & Heller, 2012). Moreover, the main effect of object properties on action-resemblance ratings was more telling of listeners' perceptions of the sounds. Bowing ratings were highest for string stimuli (Fig. 3.9), blowing ratings were highest for air column stimuli (Fig. 3.13), and striking ratings were highest for plate stimuli (Fig. 3.18). These findings imply that listeners' mental models for frequently encountered excitation-resonator interactions are quite strong, since the action-resemblance ratings were highest for objects that they are more typically paired with in acoustic musical instruments. Moreover, since the atypical interactions are not physical by nature, they may not have been well synthesized, even by a system such as Modalys.

A significant interaction between the action properties and object properties of the stimuli was also found in the current study. This allowed us to investigate more specific differences in mean action-resemblance ratings across actions at each type of object and across objects at each type of action. Here, we can examine if there were any confusions that were due to specific action or object properties. Bowing resemblance ratings (Fig. 3.10) were highest for the bowed string stimuli. This was not surprising because these stimuli are typical of excitation-resonator interactions of acoustical musical instruments. Thus, it might have been obvious to participants that the sounds were indeed produced by bowing. Mean bowing ratings were also lowest for blown air columns and struck plates, probably because these sounds also represented typical excitation-resonator interactions, which made it obvious to participants that these stimuli were not produced by bowing. Interestingly, bowing ratings were higher for the blown strings than for the bowed air columns and bowed plates. Although all these sounds represented atypical excitation-resonator combinations, listeners may have been biased to rate blown strings as bowed, which is the interaction they usually encounter.

Blowing resemblance ratings (Fig. 3.15) were highest for blown air columns, as expected. These sounds represented a typical excitation-resonator interaction, which highly influenced listeners to perceive them as blown. All struck sounds had the lowest blowing ratings; their impulsive nature may have biased listeners to perceive them as not being blown. Participants' blowing ratings were quite similar for blown strings, blown plates, and bowed air columns. Bowed air columns might have had higher blowing resemblance ratings because the air column might have influenced participants to perceive these sounds as blown. This demonstrates that our mental models are quite specific for blown sounds: they are more frequently applied to air columns.

Striking resemblance ratings (Fig. 3.20) were highest for struck plates, followed by struck air columns, and then struck strings. Struck sounds had higher striking ratings because they are impulsive and easily distinguished from continuous sounds. It is possible that the temporal centroids—distinguishing struck from bowed and blown sounds (Fig. 2.29)—of the stimuli informed listeners' judgements. The striking ratings were highest for the struck plates, probably because these sounds represent a typical excitation-resonator interaction. In the case of the struck string, which is a more common excitation-resonator interaction than a struck air column, we speculated that it depended on whether listeners heard the struck string as a plucked string. Although the actions of plucking and striking are very different, when they act on a string, what is heard is the displacement of the string. Consequently, if listeners did not perceive striking and plucking as belonging in the same category (in the context of a string), they may have assigned lower striking ratings to struck strings. The next highest striking rating was assigned to bowed plates. There are two potential reasons. The first is that the plate biased listeners to perceive bowed plates as struck rather than bowed. The second is due to the artifact produced during the beginning of the sound when the bow makes contact with the plate; this might have influenced some participants to perceive the bowed plates as struck.

It is also of interest to investigate how well the stimuli were perceived to resemble the excitation that we synthesized. In other words, even though we used Modalys to synthesize the stimuli, did listeners perceive the actual action that produced them? Figure 3.3 displays the mean bowing, blowing, and striking ratings for bowed, blown, and struck stimuli. Resemblance ratings were highest when participants rated the action that actually produced the sound. That is, bowing ratings were highest for bowed stimuli, blowing ratings were highest for blown stimuli, and striking ratings were highest for struck stimuli. This implies that participants were generally able to perceive the excitations that Modalys generated, across all resonators they were applied to. These results highlight Modalys as a useful tool in synthesizing perceptually convincing excitation methods.

We also graphed the mean bow, blow, and striking ratings for string, air column, and plate stimuli (Fig. 3.4). Bowing resemblance ratings were highest for string stimuli; blowing ratings were highest for air column stimuli, and striking ratings were highest for plate stimuli. Generally, the action-resemblance ratings were highest when the object that produced the sound typically interacts with the target action in acoustic musical instruments. This supports our hypothesis that our mental models for musical instruments are specific and informed by previous experience. The object that an action typically interacts with in acoustic musical instruments biased the perception of that action. Although performance was likely influenced by the physical nature and our mental models of excitation methods, resonance structures, and their interactions, the current findings demonstrate sensitivity to the excitation methods of sound sources.

# Chapter 4

# Experiment 2: Perceived Resemblance of Resonance Structures—String, Air Column, and Plate

The current experiment consisted of a task similar to that of Experiment 1; however, instead of rating the resemblance of the excitations that produced the stimuli, participants rated the resemblance of the resonators. The resonators were a string, air column, and plate. We hypothesized that different excitations and resonators would influence listeners' recognition of objects that produced the sound. Comparing performance between musicians and nonmusicians allowed us to see if the degree of formal musical training was associated with differences in how the two groups rated the resemblance of the sounds to different resonators. Additionally, the results allowed us to examine the perceptual precision of Modalys at modelling resonators that interact with excitations, especially when the interaction is atypical of acoustic musical instruments. We might find that some sounds have high resemblance ratings to more than one resonator or none of the given resonators. Moreover, consistent with Lemaitre and Heller's (2012) finding, we predicted that listeners are less sensitive to perceiving the resonators of sounds compared to the excitations.

## 4.1. Method

Similar to Experiment 1, the current experiment was designed as a factorial mixed-measures paradigm with three dependent variables. The same 27 stimuli were rated according to how well they resembled three resonators: string, air column, and plate. The between-groups factor of musicianship and the two repeated-measures factors of excitations and resonators that produced the sounds were the same as Experiment 1.

### 4.1.1 Participants

We recruited 43 participants (25 female, 18 male) for this experiment from McGill University's mailing list and web-based advertisements. Participants had normal hearing, which was confirmed by a pure-tone audiometric test with octave-spaced frequencies from 125 Hz to 8000 Hz at a hearing threshold of 20 dB HL (ISO 398-8 2004; Martin & Champlin, 2000). Participants were awarded $5 for their participation. The current study was certified for ethical compliance by McGill's Research Ethics Board II.

Data from two participants were discarded. One participant reported not performing the task correctly and the other participant did not meet our hearing thresholds. Of the remaining participants, there were 21 musicians (8 female, 13 male) and 20 nonmusicians (16 female, 4 male). Musicians and nonmusicians are classified in the same manner as in Experiment 1. Musician participants had an average age of 26.05 years ($SD$=4.50), and nonmusicians had an average age of 22.30 years ($SD$=3.08). Musicians were slightly older than nonmusicians, $t(39)=3.10$, $p=.004$, because musicians had the restriction of being at least in their third year of an undergraduate program, whereas nonmusicians did not have this restriction. Musicians had an average of 16.00 years of formal musical training ($SD$=5.23) on at least one instrument. Nonmusician participants had an average of 1.58 years of formal musical training ($SD$=2.02). Musicians had significantly more years of formal music training than nonmusicians, $t(39)=11.53$, $p<.001$.

Experiments 1 and 2 were run simultaneously with two independent groups. We ran more participants for each experiment two months later. Some participants who participated in either Experiment 1 or 2 returned to participate in the one they had not done. Eight participants (3 musicians, 5 nonmusicians) who participated in the current experiment already participated in Experiment 1. These participants indicated that they remembered very little of the sounds and experimental procedure. The shortest time between participation in the two experiments was 69 days and the longest time was 91 days.

### 4.1.2 Stimuli and apparatus

The stimuli and apparatus were identical to Experiment 1 (refer to Sections 3.1.2 and 3.1.3).

### 4.1.3 Procedure

The terminology was defined for participants similarly to Experiment 1; instead of defining the three actions, we defined the three objects (i.e., resonators) as: string—an object that is a thin wire fixed at its endpoints; air column—an object that is a tube, sealed at one end, and open at the other end; and plate—an object that is a thin, rigid, and flat rectangle. When we discuss the experimental procedure and results, we will refer to the excitations as actions and resonators as objects. We did not discuss the definitions of the actions with the participants so they could focus only on the object properties when they made their judgements.

The procedure of the current experiment was similar to that of Experiment 1, except that the ratings were concerned with objects rather than actions. That is, participants were asked to rate each sound based on how well they resembled a string, air column, or plate. Like Experiment 1, there were three blocks in the experiment, each with 27 trials (Fig. 4.1). Each block concerned ratings for only one type of object and the order of blocks was randomized across participants. The format of the experiment and practice trials, ordering of trials per block, and opportunity to hear the full range of stimuli prior to the experiment were identical to Experiment 1.

Similar to Lemaitre and Heller's (2012) paradigm, the ratings for the resemblance of an object matched or mismatched the actual material (wood, plastic, metal, glass) that produced the sound. In Lemaitre and Heller's (2012) experiment, the distribution of resemblance ratings that either matched or mismatch the material that produced the sound was half and half. In the current experiment, one-third of the resemblance ratings matched the actual object that produced the stimuli, and two-thirds of the resemblance ratings did not match the actual object. We investigated whether the stimuli conveyed the true object that produced them, a different object, or none of the objects.

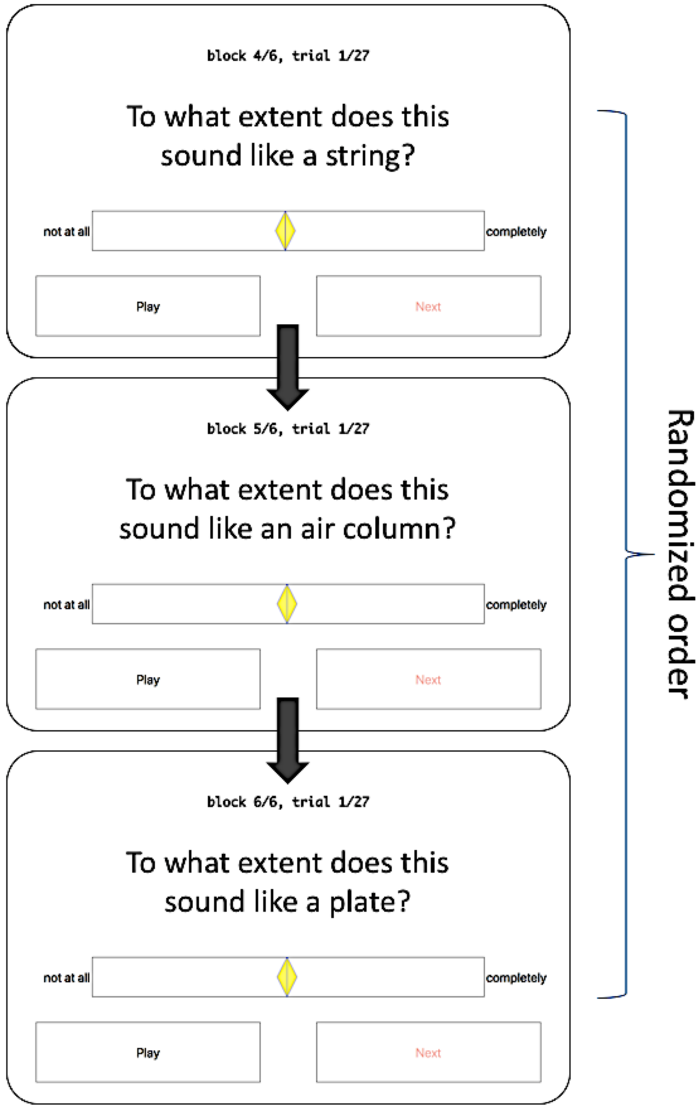


**Figure 4.1.** Summary of the experimental design and procedure of Experiment 2.

## 4.2 Results

Participants performed ratings for how well each stimulus resembled string, air column, and plate objects. The object-resemblance ratings for the three exemplars of each action-object interaction were averaged for each participant. Similar to Experiment 1, a 2×3×3 MANOVA was conducted. The between-groups factor was musicianship with two levels: musicians and nonmusicians. One repeated-measures factor was the object that produced the stimuli and has three levels: string, air column, and plate. The other repeated-measures factor was the action that produced the stimuli and has three levels: bowing, blowing, and striking.

#### 4.2.1 Multivariate analyses

SPSS did not compute Box's Test of Equality of Covariance Matrices for our data. The warning message was the same as that for Experiment 1. This was likely due to the fact that more than two levels of the independent variables shared at least 90% of the ratings on the dependent variables. Pillai's Trace, *V*, was used as the multivariate statistic as it is more robust to violations of assumptions (Olson, 1974).

MANOVA revealed that the between-groups effect of musicianship on the string, air column, and plate resemblance ratings was not statistically significant, $V$=.02, $F$<1. The musicians' mean object-resemblance ratings were very similar to those of the nonmusicians. The within-groups effect of objects that produce the stimuli was statistically significant, $V$=1.58, $F(6,154)$=95.13, $p$<.001, $\eta_p^2$=.79. This effect was characterized by the pattern that string ratings were highest for string stimuli, air column ratings were highest for air column stimuli, and plate ratings were highest for plate stimuli (Fig. 4.2). Plate and air column stimuli were perceived as more ambiguous than string stimuli. The within-groups effect of action properties of the stimuli was statistically significant, $V$=1.05, $F(6,154)$=28.41, $p$<.001, $\eta_p^2$=.53. Bowed stimuli were perceived as strings or air columns. Blown stimuli were mostly perceived as air columns, and struck stimuli were perceived as strings or plates (Fig. 4.3).

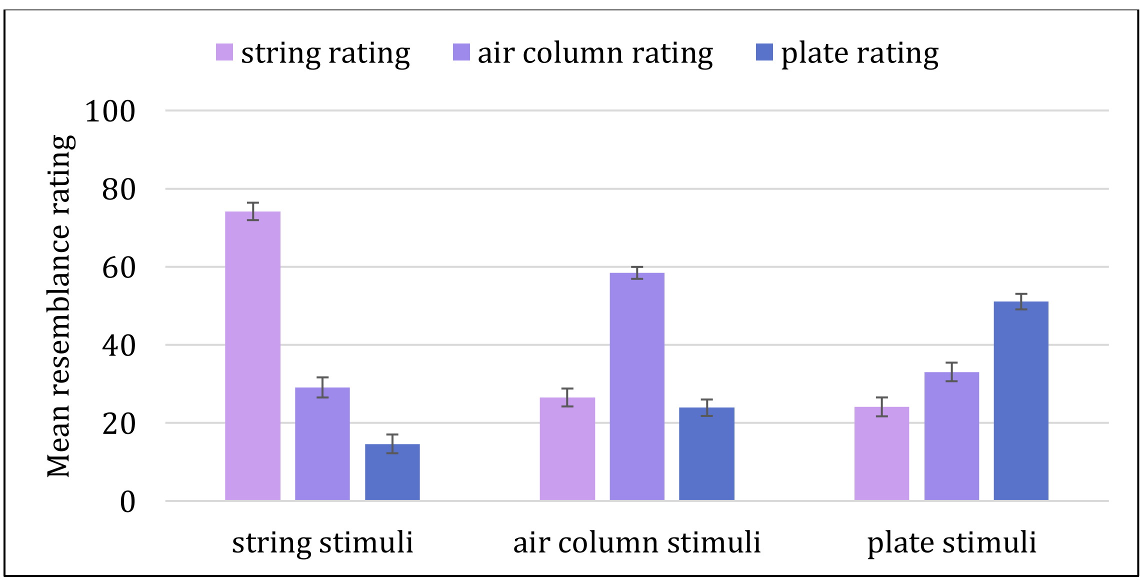


**Figure 4.2.** Mean string, air column, and plate resemblance ratings based on objects. Error bars indicate standard error of the mean.

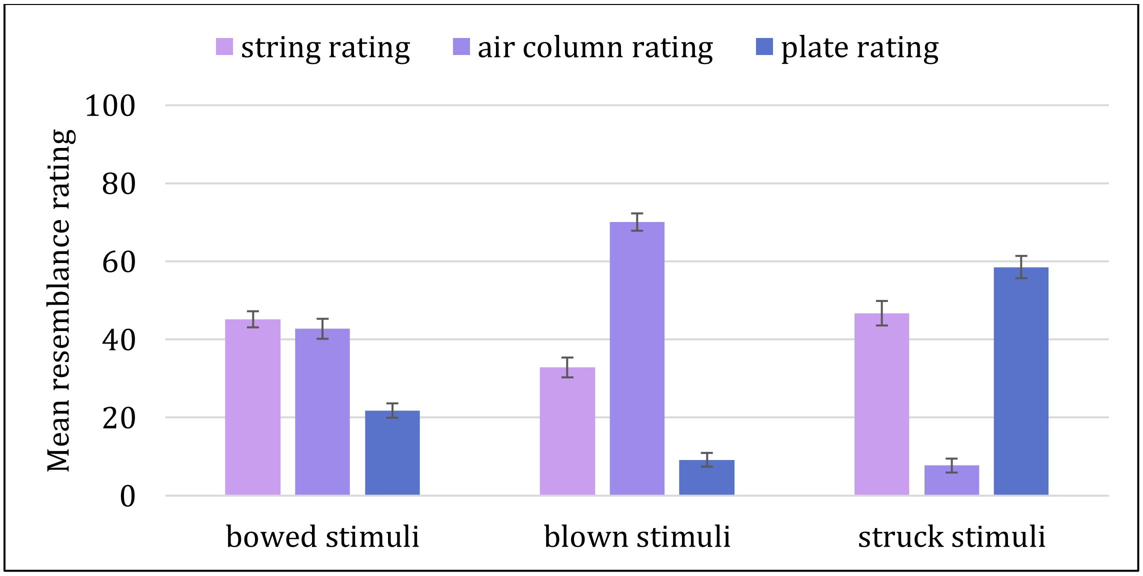


**Figure 4.3.** Mean string, air column, and plate resemblance ratings based on actions. Error bars indicate standard error of the mean.

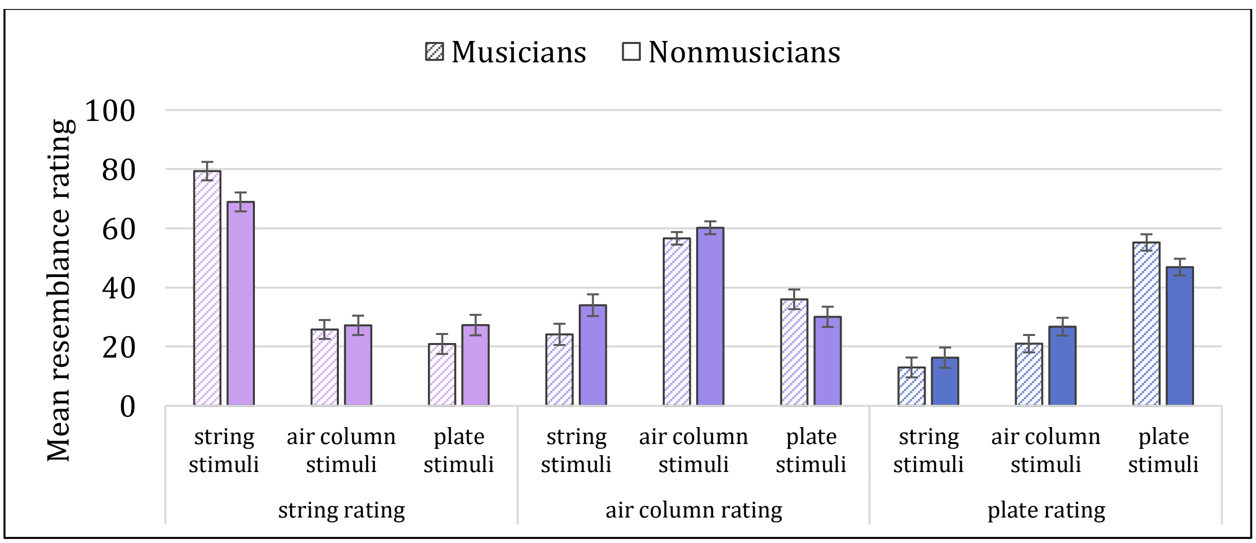


**Figure 4.4.** Musicians' and nonmusicians' mean string, air column, and plate resemblance ratings for each object property. Different colours represent different object-resemblance ratings. Solid bars represent nonmusicians and patterned bars represent musicians. Error bars represent standard error of the mean.

The interaction effect of musicianship and object properties was statistically significant, $V$=.23, $F$(6,154)=3.34, $p$=.004, $\eta_p^2$=.12 (Fig. 4.4). Musicians and nonmusicians rated the stimuli's resemblances to strings, air columns, and plates differently depending on the objects that produced them. Regardless of musicianship, listeners were more likely to perceive the object that actually produced the given stimulus. However, there were not any notable differences between musicians' and nonmusicians' ratings for particular object properties. The interaction effect of musicianship and action properties was not statistically significant, $F<1$. Musicians and nonmusicians did not assign resemblance ratings differently for different action properties. The interaction effect of action and object properties on the object-resemblance ratings was statistically significant, $V=1.08$, $F(12,468)=21.93$, $p<.001$, $\eta_p^2=.36$ (Fig. 4.5). Resemblance ratings were influenced by the combinations of actions and objects that produced the stimuli. Blown strings were equally perceived as air columns and strings. Blown plates were perceived as air columns. Struck air columns were equally perceived as strings and plates. The interaction effect of musicianship, action properties, and object properties on the three object-resemblance ratings was not statistically significant, $V=.09$, $F(12,468)=1.26$, $p=.242$, $\eta_p^2=.03$. Musicians' and nonmusicians' object resemblance ratings were not very different for different action-object combinations that produced the stimuli. However, one notable difference was that musicians perceived blown strings to resemble strings more than air columns, but nonmusicians perceived them to resemble air columns more than strings. Musicians also had higher plate ratings for bowed plates, but nonmusicians did not perceive these stimuli to resemble a particular object over the others (Fig. 4.6).

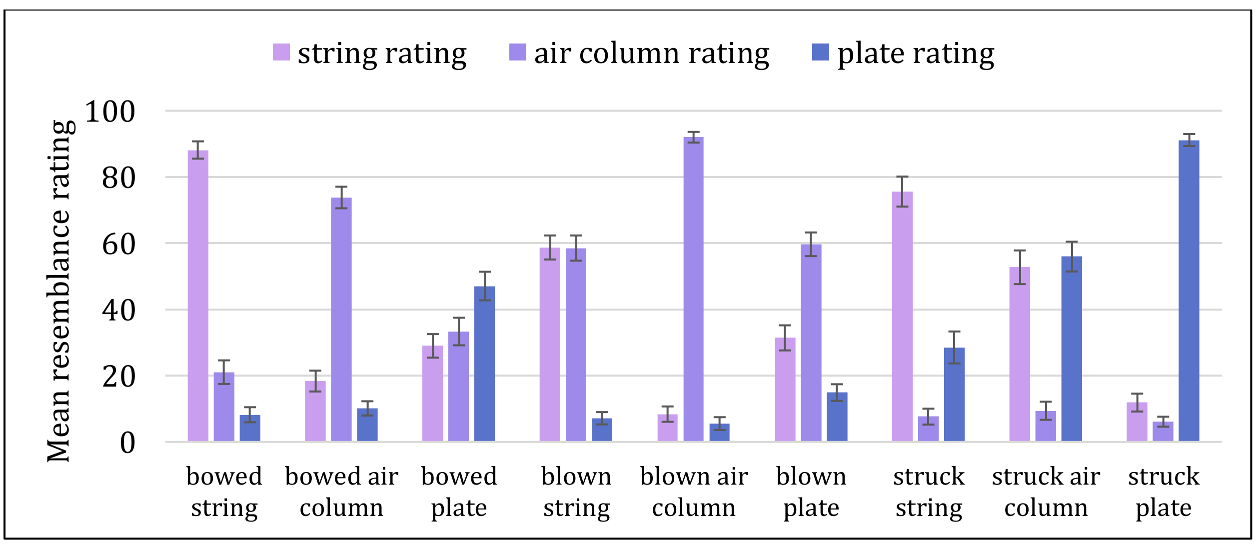


**Figure 4.5.** Mean string, air column, and plate resemblance ratings for each action-object interaction. Error bars represent standard error of the mean.

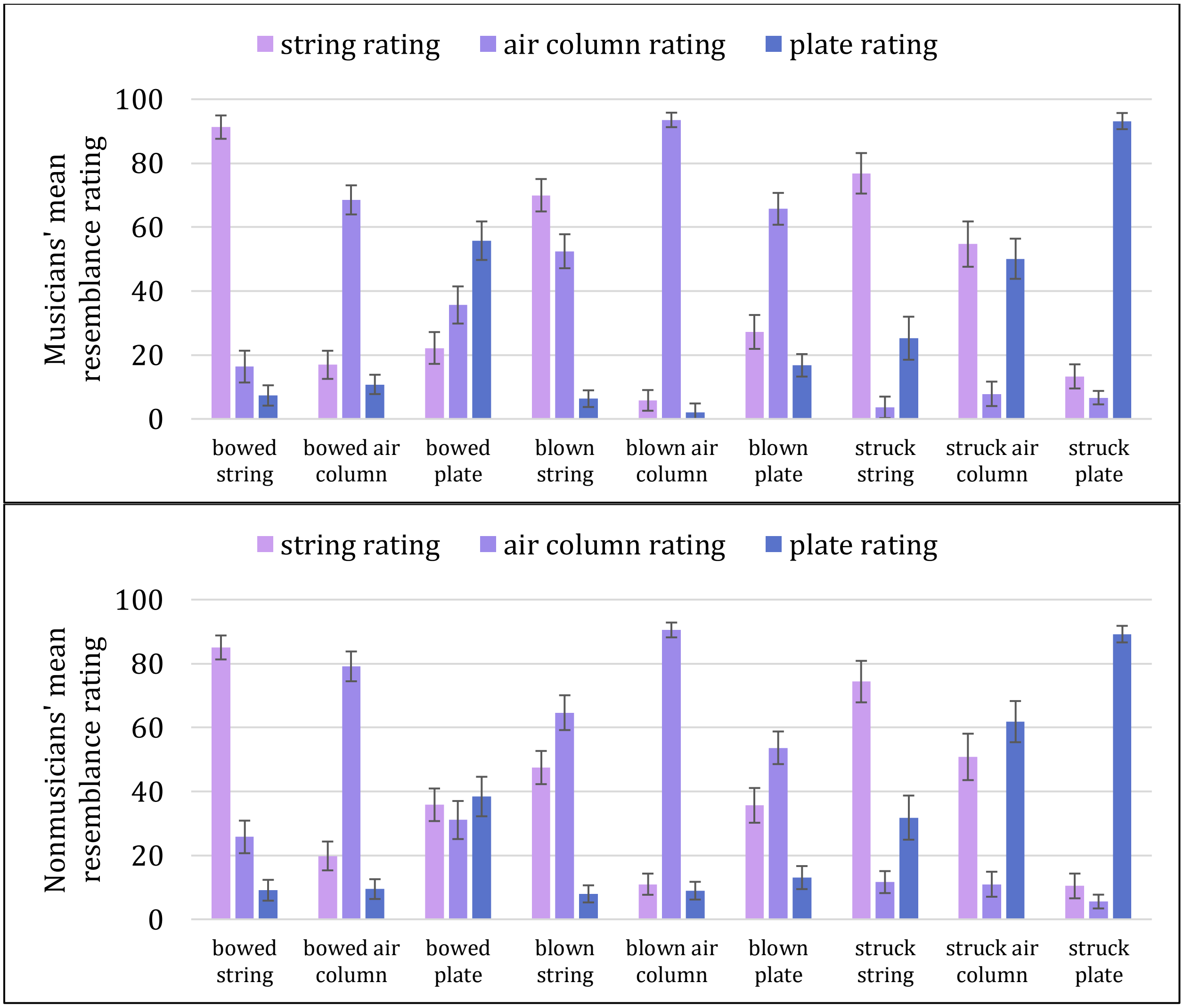


**Figure 4.6.** Mean object-resemblance ratings of musicians (top) and nonmusicians (bottom) for each action-object interaction. Error bars represent standard error of the mean.

### 4.2.2 Univariate analyses: String resemblance ratings

Univariate analyses were computed for string resemblance ratings. Mauchly's Test revealed that the sphericity assumption was violated for the repeated-measures effects of action properties, $p<.001$; combined action and object properties, $p=.002$; but not object properties, $p=.220$. We applied the Greenhouse-Geisser estimate of epsilon where appropriate to control for inflation of the $F$ statistic.

#### *4.2.2.1 Main effects*

Univariate analyses revealed that the main effect of musicianship on string ratings was not statistically significant, *F*<1. Musicians and musicians did not differ in their mean string resemblance ratings. The mean string ratings for musicians and nonmusicians were 42.02 (*SD*=36.42) and 41.15 (*SD*=34.56), respectively.

We found a significant main effect of object properties on string resemblance ratings, $F(2,78)=212.25$, $p<.001$, $\eta_p^2=.85$ (Fig. 4.7). Mean string resemblance ratings are 74.15 (*SD*=27.33) when stimuli are produced by a string, 26.50 (*SD*=30.13) when produced by an air column, and 24.10 (*SD*=23.41) when produced by a plate. Bonferroni-corrected pairwise comparisons revealed that mean string ratings were significantly different when stimuli were produced by a string and air column, $Z=47.65$, $p<.001$, and when stimuli were produced by a string and plate, $Z=50.06$, $p<.001$. However, the difference in mean string ratings was not significant when stimuli were produced by an air column and plate, $Z=2.40$, $p=.929$. *Z* represents the mean difference in resemblance ratings in absolute value. String ratings were highest for string stimuli. Moreover, string objects were not often confused with air column or plate objects.

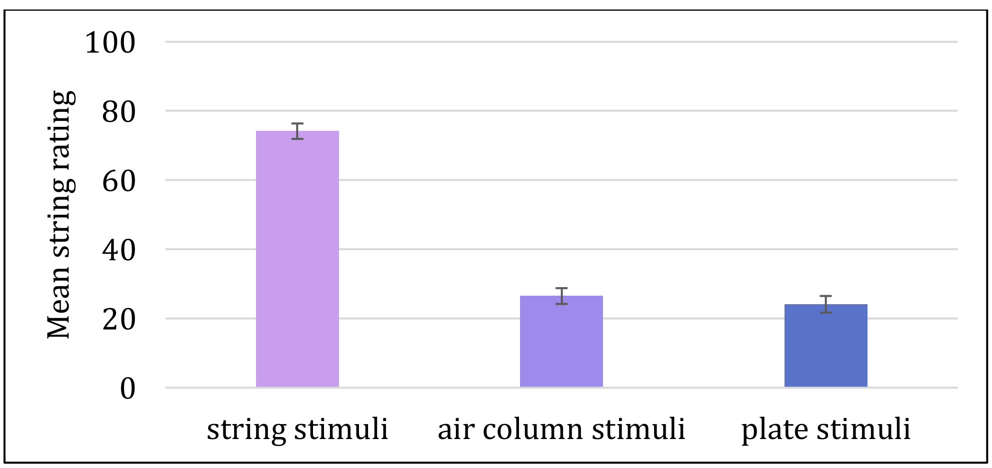


**Figure 4.7.** Main effect of object properties on string resemblance ratings. Error bars indicate standard error of the mean.

There was also a significant main effect of action properties on string resemblance ratings, $F(1.45,56.41)=9.70$, $p=.001$, $\varepsilon=.72$, $\eta_p^2=.20$ (Fig. 4.8). The mean string resemblance ratings were 45.18 (*SD*=36.88) for bowed stimuli, 32.84 (*SD*=30.24) for blown stimuli, and 46.73 (*SD*=37.49) for struck stimuli. Post-hoc analyses revealed that mean string resemblance ratings were significantly different for bowed and blown stimuli, *Z*=12.33, *p*<.001, and for blown and struck stimuli, *Z*=13.89, *p*=.006, but not for bowed and struck stimuli, *Z*=2.18, *p*=1.000. Although string ratings were higher for bowed and struck stimuli than for blown stimuli, the ratings indicated that perception of a string object was ambiguous across the three actions.

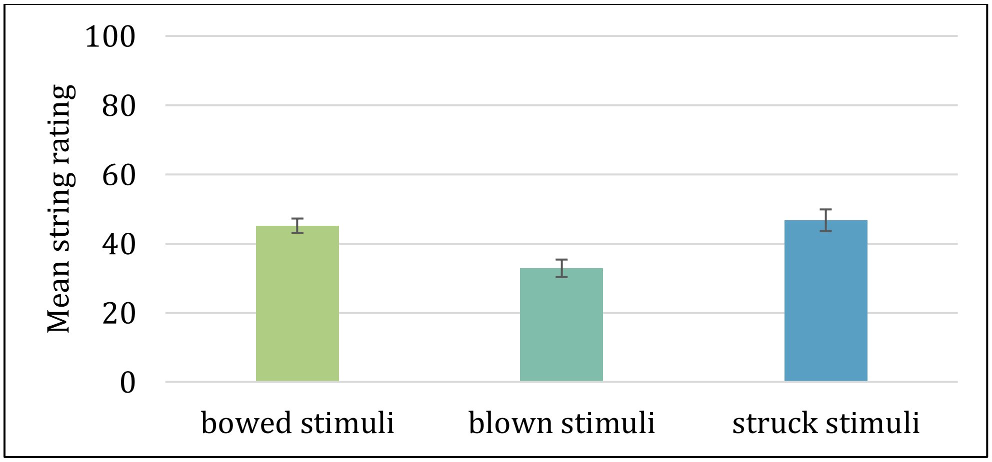


**Figure 4.8.** Main effect of action properties on string resemblance ratings. Error bars indicate standard error of the mean.

### *4.2.2.2 Interaction effects*

We found that the interaction effect between musicianship and object properties of the stimuli was statistically significant, $F(2,78)=4.96$, $p=.009$, $\eta_p^2=.11$ (Fig. 4.9). We performed a simple effects analysis to see how each musicianship group rated string resemblance differently for different objects. This simple effect was significant for both musicians, $F(2,78)=143.52$, $p<.001$, and nonmusicians, $F(2,78)=75.35$, $p<.001$. Both groups had the highest string ratings for string stimuli. Furthermore, we analyzed another simple effect to see if both groups assigned string ratings differently for each type of object property. The simple effect was significant for string stimuli, $F(1,39)=5.42$, $p=.025$, but not for air columns, $F<1$, and plates, $F(1,39)=1.75$, $p=.194$. Thus, musicians assigned higher string ratings for string stimuli than do nonmusicians. The interaction between musicianship and action properties was not significant, $F<1$. String ratings did not differ between musicians and nonmusicians when different actions produced the stimuli.

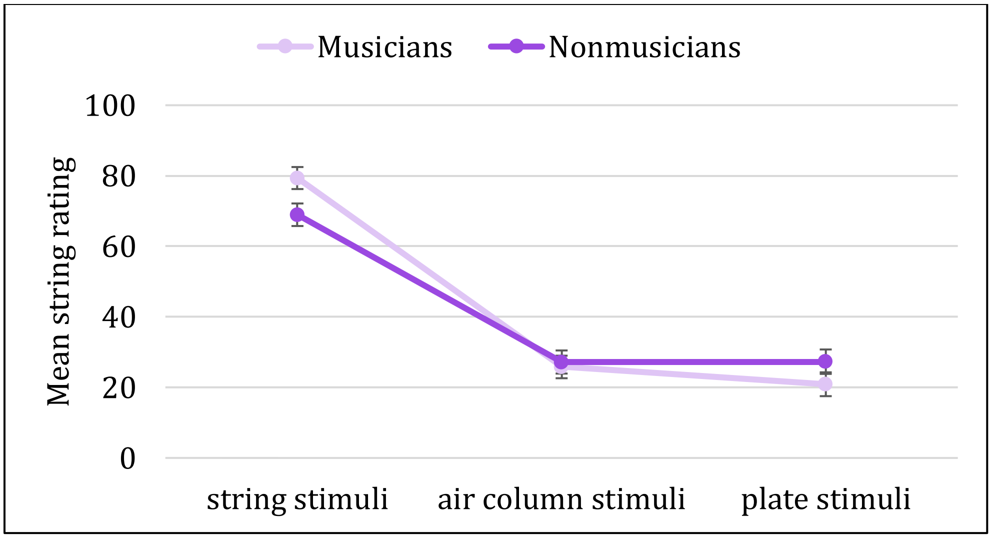


**Figure 4.9.** Interaction effect of musicianship and object properties on mean string resemblance ratings. Error bars represent standard error of the mean.

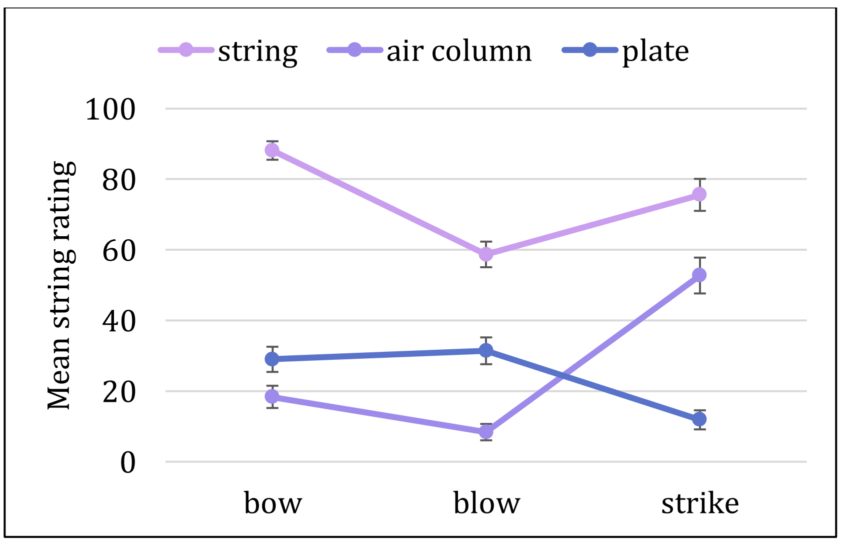


**Figure 4.10.** Interaction effect of combined actions and objects on string resemblance ratings. Each coloured line represents a different object and the actions are on the horizontal axis. Error bars indicate standard error of the mean.

The interaction effect between action and object properties on string ratings was statistically significant, $F(2.93,116.05)=43.81$, $p<.001$, $\varepsilon=.74$, $\eta_p^2=.53$ (Fig. 4.10). String ratings were highest for bowed, blown, and struck strings. Plates and air columns were less likely to be perceived as strings, with the exception of the struck air column. A non-parametric simple effects analysis will be examined in Section 4.2.5 to uncover more specific differences in mean string ratings across object properties at each level of action properties and vice versa. The three-way interaction effect of musicianship, actions, and objects was not statistically significant, $F(2.93,116.05)=2.59$, $p=.057$, $\varepsilon=.74$, $\eta_p^2=.06$. Musicians and nonmusicians' mean string ratings did not differ when they heard stimuli that were produced by different combinations of actions and objects.

### 4.2.3 Univariate analyses: Air column resemblance ratings

The univariate analyses were conducted for the air column resemblance ratings. Mauchly's test of sphericity revealed that none of the within-groups effects of object properties, action properties, or interactions between action and object properties violated the sphericity assumption, $p=.559$, $p=.393$, $p=.867$, respectively. Thus, we did not apply adjustments to the $F$ statistic.

#### *4.2.3.1 Main effects*

Univariate analyses revealed that the main effect of musicianship on air column ratings was not statistically significant, $F<1$. Musicians' and nonmusicians' mean air column ratings were very similar. The mean air column resemblance ratings were 38.92 ($SD=35.53$) for musicians and 41.43 ($SD=36.51$) for nonmusicians.

The main effect of object properties on air column ratings was statistically significant, $F(2,78)=72.26$, $p<.001$, $\eta_p^2=.65$ (Fig. 4.11). The mean air column ratings were 29.08 ($SD=30.39$) for string stimuli, 58.41 ($SD=39.36$) for air column stimuli, and 33.04 ($SD=30.36$) for plate stimuli. Post-hoc comparisons revealed that the difference in mean ratings was statistically significant for strings and air columns, $Z=29.33$, $p<.001$, and for air columns and plates, $Z=25.37$, $p<.001$, but not for strings and plates, $Z=3.97$, $p=.408$. Air column stimuli were most perceived as air columns in comparison to string and plate stimuli.

The main effect of action properties on air column ratings was statistically significant, $F(2,78)=276.94$, $p<.001$, $\eta_p^2=.88$ (Fig. 4.12). Mean ratings were 42.74 ($SD=32.57$) for bowed stimuli, 70.08 ($SD=25.68$) for blown stimuli, and 7.71 ($SD=14.54$) for struck stimuli. Pairwise comparisons revealed a significant difference in mean ratings for: bowed and blown stimuli, $Z=27.34$, $p<.001$; bowed and struck stimuli, $Z=35.03$, $p<.001$; and blown and struck stimuli, $Z=62.37$, $p<.001$. Although blown stimuli were most often perceived as air columns, listeners also occasionally perceived bowed stimuli as air columns.

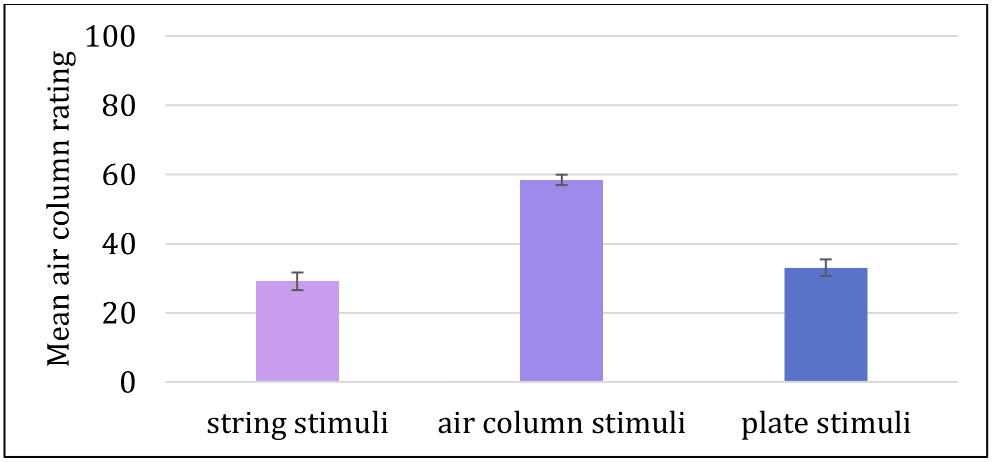


**Figure 4.11.** Main effect of object properties on air column resemblance ratings. Error bars indicate standard error of the mean.

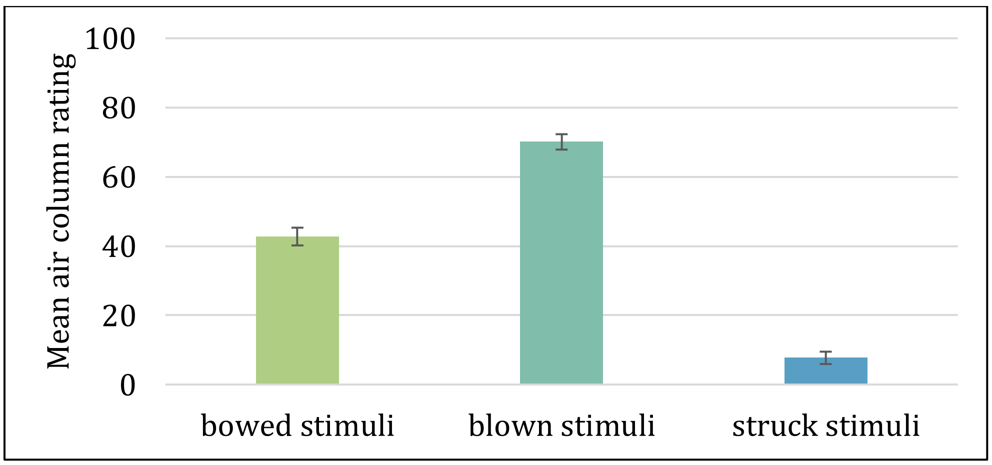


**Figure 4.12.** Main effect of action properties on air column resemblance ratings. Error bars indicate standard error of the mean.

#### *4.2.3.2 Interaction effects*

The interaction effect between musicianship and object properties on air column ratings was statistically significant, $F(2,78)=4.50$, $p=.014$, $\eta_p^2=.10$ (Fig. 4.13). A simple effects analysis was performed to test for differences in how each musicianship group assigned air column ratings to different objects. For musicians, the simple effect of object properties was statistically significant, $F(2,78)=39.50$, $p<.001$. Their air column ratings were highest for air column stimuli. The simple effect of object properties is also significant for nonmusicians, $F(2,78)=37.31$, $p<.001$. Nonmusicians were also more likely to perceive air column stimuli as air columns. Moreover, we tested the simple effect of musicianship at each level of the object properties. These simple effects analyses revealed that musicians and nonmusicians do not rate air column resemblance differently from each other regardless of the object that produces the stimuli: $F(1,39)=3.68$, $p=.053$ for string stimuli, $F(1,39)=1.37$, $p=.250$ for air column stimuli; and $F(1,39)=1.54$, $p=.222$ for plate stimuli. The interaction effect between musicianship and action properties of the stimuli on air column resemblance ratings was not statistically significant, $F<1$. Thus, musicians and nonmusicians did not rate air column resemblance differently when different actions produce the stimuli.

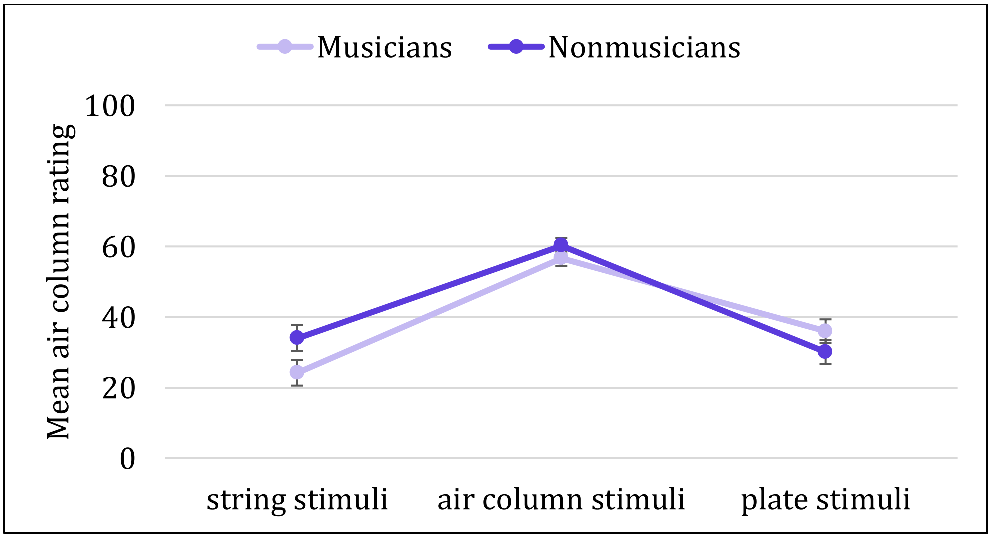


**Figure 4.13.** Interaction effect of musicianship and object properties on mean air column resemblance ratings. Error bars represent standard error of the mean.

The interaction effect between action and object properties on air column ratings was statistically significant, $F(4,156)=35.05$, $p<.001$, $\eta_p^2=.47$ (Fig. 4.14). Air column ratings were highest for blown stimuli and air column stimuli, with the exception of the struck air column. In fact, all struck stimuli were least likely to be perceived as air columns. This is probably because struck sounds are impulsive, but air columns are usually associated with continuous excitations. A simple effects analyses will be presented in Section 4.2.5 to investigate this interaction further. Lastly, the three-way interaction between musicianship, action properties, and object properties was not significant, $F(4,156)=1.27$, $p=.285$, $\eta_p^2=.03$. Musicians and nonmusicians did not reliably rate air column resemblance differently for different combinations of action and object properties.

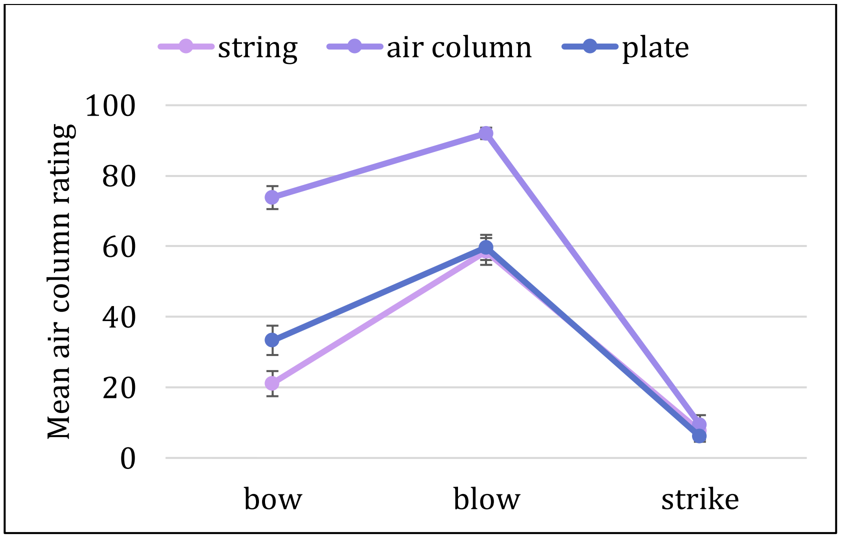


**Figure 4.14.** Interaction effect of combined actions and objects on air column resemblance ratings. Each coloured line represents a different object and the actions are on the horizontal axis. Error bars indicate standard error of the mean.

### 4.2.4 Univariate analyses: Plate resemblance ratings

Univariate analyses were conducted for the plate resemblance ratings. Mauchly's test revealed violations of the sphericity assumption for three within-groups effects of object properties, action properties, and the interaction between action and object properties, $p<.001$ for all. Consequently, we will use Greenhouse-Geisser's conservative adjustment of the $F$ ratio.

#### *4.2.4.1 Main effects*

There was no main effect of musicianship on plate ratings, $F<1$, meaning that musicians and nonmusicians did not differ in their mean plate ratings. The mean ratings were 29.73 ($SD$=34.05) for musicians and 29.97 ($SD$=34.60) for nonmusicians.

The main effect of object properties on plate ratings was statistically significant, $F(1.44,56.13)=108.48$, $p<.001$, $\varepsilon=.72$, $\eta_p^2=.74$ (Fig. 4.15). Plate ratings were influenced by the type of object that produced the sound. The mean plate ratings were 14.62 ($SD$=22.79), 23.88 ($SD$=30.19), and 51.05 ($SD$=37.19) for strings, air columns, and plates, respectively. Post-hoc analyses revealed that the mean difference in plate ratings was statistically significant for: strings and air columns, $Z=9.25$, $p<.001$; strings and plates, $Z=36.43$, $p<.001$; and air columns and plates, $Z=27.18$, $p<.001$. Plate stimuli were most perceived as plates and air columns were more likely perceived as plates than strings.

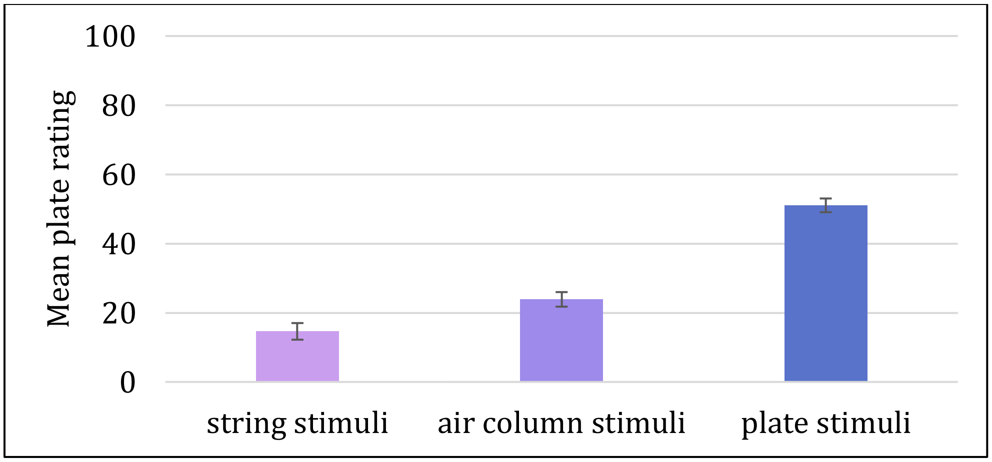


**Figure 4.15.** Main effect of object properties on plate resemblance ratings. Error bars indicate standard error of the mean.

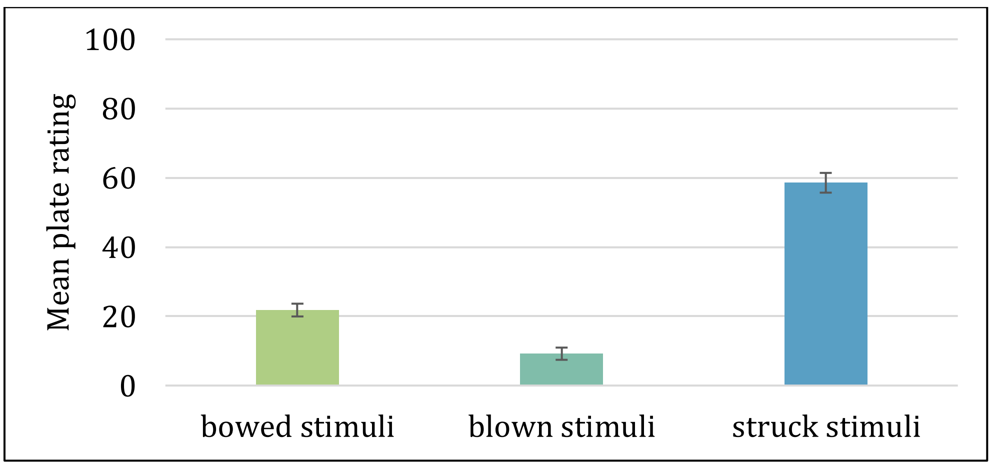


**Figure 4.16.** Main effect of action properties on plate resemblance ratings. Error bars indicate standard error of the mean.

Furthermore, univariate analyses revealed a significant main effect of action properties on plate resemblance ratings, $F(1.46,56.90)=185.67$, $p<.001$, $\varepsilon=.73$, $\eta_p^2=.83$ (Fig. 4.16). Mean plate ratings differed depending on the type of actions that produced the stimuli. The mean ratings for bowed, blown, and struck stimuli were 21.80 ($SD=26.93$), 9.20 ($SD=14.16$) and 58.55 ($SD=35.99$), respectively. Post-hoc pairwise comparisons revealed significant differences in mean ratings for: bowed and blown stimuli, $Z=12.61$, $p<.001$; bowed and struck stimuli, $Z=36.75$, $p<.001$; and blown and struck stimuli, $Z=49.36$, $p<.001$. Struck stimuli were most often perceived as plates. Bowed stimuli were more likely perceived as plates than blown stimuli.

#### *4.2.4.2 Interaction effects*

A significant interaction between musicianship and object properties on plate ratings was revealed, $F(1.44,56.13)=4.28$, $p=.030$, $\varepsilon=.72$, $\eta_p^2=.10$ (Fig. 4.17). We analyzed differences between musicians' and nonmusicians' plate ratings when stimuli were produced by different objects. Mean plate resemblance ratings were significantly different depending on the objects that produce the stimuli for both musicians, $F(2,78)=78.06$, $p<.001$, and nonmusicians, $F(2,78)=35.73$, $p<.001$. For both groups, ratings were highest for plates. We also analyzed differences in ratings between musicians and nonmusicians for each type of object. There was no significant difference between musicians' and nonmusicians' mean ratings for string stimuli, $F<1$, and air column stimuli, $F(1,39)=1.85$, $p=.181$. However, musicians' and nonmusicians' mean plate ratings did differ for plate stimuli, $F(1,39)=4.40$, $p=.042$: musicians were more sensitive to plate stimuli as resembling plates than were nonmusicians. The interaction between musicianship and action properties on plate ratings was not significant, $F(1.46,56.90)=2.01$, $p=.154$, $\varepsilon=.73$, $\eta_p^2=.05$. Musicians and nonmusicians' mean ratings did not differ depending on the actions that produced the stimuli.

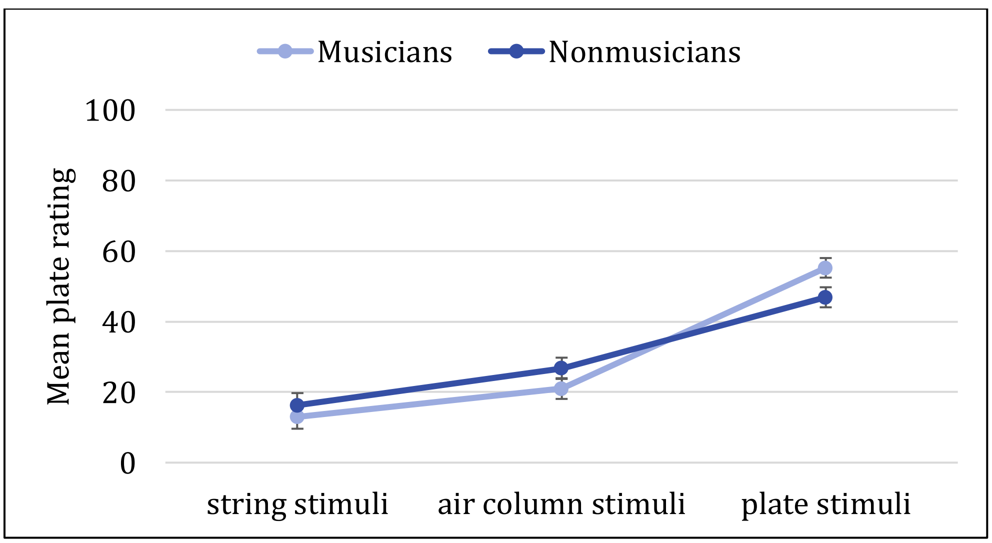


**Figure 4.17.** Interaction effect between musicianship and object properties on mean plate resemblance ratings. Error bars represent standard error of the mean.

A significant interaction effect was found between combinations of action and object properties, $F(3.24,126.51)=31.15$, $p<.001$, $\varepsilon=.81$, $\eta_p^2=.46$ (Fig. 4.18). Struck plates and air columns were more often perceived as plates, whereas struck strings were perceived as plates to a lesser degree. Struck strings represent a typical action-object interaction, so it was less likely to be heard as a plate. All blown stimuli had very low plate ratings. In the case of the blown plate, since this represents an atypical action-object interaction, these stimuli may have biased participants into thinking that they were not produced by plates. Other than the blown plate, the other plate stimuli were often perceived as plates, with a bias towards struck plates. A non-parametric simple effects analysis was conducted to further explain differences in plate ratings for this interaction (refer to Section 4.2.5). The three-way interaction between musicianship, action properties, and object properties was not significant, $F<1$, meaning that musicians and nonmusicians did not assign different plate ratings to different combinations of actions and objects.

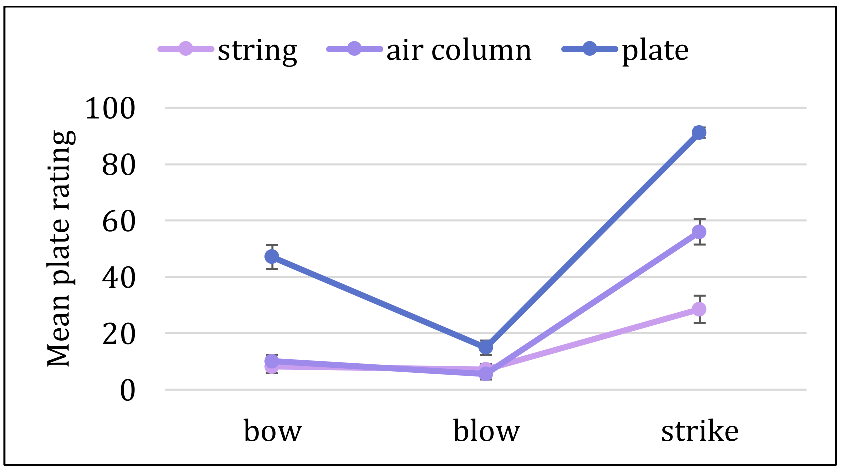


**Figure 4.18.** Interaction effect of combined actions and objects on plate resemblance ratings. Each coloured line represents a different object and the actions are on the horizontal axis. Error bars indicate standard error of the mean.

### 4.2.5 Simple effects analyses: Interaction between actions and objects

The univariate tests revealed significance of Levene's test of homogeneity of variance for several string, air column, and plate resemblance ratings. Thus, the homogeneity of variance assumption was violated. We then assessed the normality of the residuals from the MANOVA test. The normality assumption was violated for the residuals, as revealed by a significant Shapiro-Wilk test, $p<.001$. We used Friedman's test to analyze the simple effects for the interaction between action and object properties reported in the univariate analyses. We investigated the simple effects of both action properties for each type of object and object properties for each type of action. The method of running these simple effects analyses is identical to that explained in Experiment 1.

#### *4.2.5.1 String resemblance ratings*

We conducted simple effects analyses to explain the effect of the interaction between objects and actions on string resemblance ratings (Fig. 4.10). Friedman's test revealed a significant simple effect of action properties for string stimuli, $\chi^2(2, N=41)=30.98$, $p<.001$. Overall, bowed, blown, and struck strings were likely to be perceived as strings. String ratings were highest for bowed strings, followed by struck strings, then blown strings. These results reflect the bias of our mental models, which are stronger for bowed and struck strings, since they resembled typical action-object combinations. The simple effect of action properties for air columns was significant, $\chi^2(2, N=41)=34.91$, $p<.001$. Struck air columns were more likely perceived as strings than bowed and blown air columns. This is interesting, as the struck air column represents an atypical excitation-resonator interaction; thus, mental models for this interaction are weak, which might have biased participants into perceiving a string. Friedman's test revealed a significant simple effect of action properties for plates, $\chi^2(2, N=41)=26.97$, $p<.001$. String ratings were low overall for bowed, blown, and struck plates, which implies that strings were easily distinguished from plates.

Friedman's test revealed that the simple effect of object properties for bowed stimuli was significant, $\chi^2(2, N=41)=55.90$, $p<.001$. Bowed strings were most often perceived as strings; bowed air columns and plates were not often perceived as strings. These results indicate that mental models for bowed strings guided listeners to perceiving strings. The simple effect of object properties for blown stimuli was significant, $\chi^2(2, N=41)=61.83$, $p<.001$. Blown strings were more often perceived as strings than blown air columns and plates. The string ratings for the blown plate indicate that the plate was occasionally mistaken for a string when it was blown. A significant simple effect of object properties for struck stimuli was revealed, $\chi^2(2, N=41)=63.40$, $p<.001$. Struck plates were rarely perceived as strings; it might have been obvious that a string did not produce these sounds. Struck strings were often perceived as strings, and participants occasionally mistook the struck air column stimuli as produced by strings.

#### *4.2.5.2 Air column resemblance ratings*

Simple effects analyse were conducted to examine the effect of the interaction between objects and actions on air column ratings (Fig. 4.14). Friedman's test revealed a significant simple effect of action properties on air column ratings when a string produced the stimuli, $\chi^2(2, N=41)=62.92$, $p<.001$. Blown strings were more likely heard as air columns than bowed and struck strings. This demonstrates the bias of blowing, which typically interacts with an air column in acoustic musical instruments. Moreover, the blown strings we synthesized may resemble a saxophone, which explains why they may have been mistaken for air columns. When stimuli are produced by an air column, the type of action had a significant simple effect on the air column ratings, $\chi^2(2, N=41)=70.08, p<.001$. Blown and bowed air columns were often heard as air columns, with a bias towards blown air columns. However, struck air columns were rarely perceived as air columns, demonstrating that listeners did not associate air columns with striking. The simple effect of action properties on air column resemblance ratings was significant for plates, $\chi^2(2, N=41)=61.98, p<.001$. Bowed plates were sometimes perceived as air columns, but blown plates were more often perceived as air columns. The blown plate may sound like saxophone multiphonics, so it is not surprising that they were often mistaken as air columns. This speaks mostly to Modalys' ability to convey plates, since they have occasionally been mistaken for air columns.

A significant simple effect of object properties for bowed stimuli was revealed, $\chi^2(2, N=41)=51.16, p<.001$. Bowed air columns were most often heard as air columns and bowed plates were sometimes confused with air columns. Since plates are not typically associated with continuous excitations, the continuous nature of a bowed excitation may have confused listeners. The simple effect of object properties on air column resemblance ratings was significant for blown actions, $\chi^2(2, N=41)=54.64, p<.001$. Blown stimuli were often perceived as air columns, especially for blown air columns. Blown strings and plates may have been mistaken as air columns since blowing is typically associated with air columns. Friedman's test revealed no significant main effect of object properties on struck sounds, $\chi^2(2, N=41)=.24, p=.899$. None of the struck stimuli were perceived as air columns—not even struck air columns. Either Modalys was not entirely effective at conveying an air column through its interaction with striking, or participants were biased by the fact that struck air columns did not match their mental models for excitation-resonator interactions.

#### *4.2.5.3 Plate resemblance ratings*

Simple effects analyses were analyzed to further investigate the effect of the interaction between objects and actions on plate ratings (Fig. 4.18). Friedman's test revealed a significant simple effect of action properties for string stimuli, $\chi^2(2, N=41)=22.57, p<.001$. Overall, plate ratings were low for strings, but they were highest for struck strings. Generally, listeners were able to distinguish plates from strings. The simple effect of action properties on plate ratings for air columns was significant, $\chi^2(2, N=41)=56.38$, $p<.001$. Struck air columns were more likely to be perceived as plates than bowed and blown air columns. This might imply that Modalys was not very effective in conveying an air column when it was struck. The simple effect of action properties on plate ratings was significant for stimuli produced by a plate, $\chi^2(2, N=41)=72.63$, $p<.001$. Plate ratings were highest for struck plates, as expected. However, the bowed plate was perceived as a plate less often and the blown plate was hardly perceived as a plate. Modalys may not have conveyed a plate effectively or listeners mental models might be hindering their ability to perceive a plate.

Friedman's test revealed a significant simple effect of object properties for bowed stimuli, $\chi^2(2, N=41)=55.33$, $p<.001$. Bowed plates were more often perceived as plates than their air column and string counterparts. Participants were able to distinguish plates from air columns and strings, but the action of bowing hindered perception occasionally. We found a significant simple effect of object properties on plate ratings for blown stimuli, $\chi^2(2, N=41)=24.33$, $p<.001$. None of the blown stimuli were perceived as plates, although the highest plate rating was assigned to the blown plate. However, the rating was still low, which indicates the influence of listeners' mental models for blown sounds and how they cannot be applied to plates in the physical world. The simple effect of object properties on plate resemblance ratings was significant for struck stimuli, $\chi^2(2, N=41)=51.28$ $p<.001$. Struck plates were mostly perceived as plates and struck strings were not commonly perceived as plates. The ratings for struck air columns, however, fell somewhere in the middle. This demonstrates that listeners were biased to thinking an air column cannot be struck.

## 4.3 Discussion

We initially hypothesized that different excitation-resonator interactions would influence how listeners perceived the objects of the resulting sounds. Participants in our experiment rated 27 stimuli made up of different excitation-resonator interactions based on the extent to which they resembled three objects: string, air column, and plate. We found minimal differences between musician and nonmusician participants.

For string, air column, and plate resemblance ratings, there was consistently a main effect of object properties: different objects that produced the stimuli influenced how participants assigned the resemblance ratings. Ratings were highest when the object that produced the stimuli matched the object participants were rating (Figs. 4.7 for strings, 4.11 for air columns, and 4.15

for plates). This is consistent with previous findings that listeners can identify object properties (e.g., materials) of sounds quite well (Lemaitre & Heller, 2012; Hjortkjær & McAdams, 2016). There was also a main effect of action properties on the object ratings. String ratings were highest for bowed and struck stimuli (Fig. 4.8); air column ratings were highest for blown stimuli (Fig. 4.12); and plate ratings were highest for struck stimuli (Fig. 4.16). This supports our hypothesis that excitations that typically stimulate certain resonators in acoustic instruments biased the object-resemblance ratings. For example, string ratings were highest for the two actions that typically interact with strings in acoustic musical instruments. When participants heard a bowed or struck stimulus, they were primed to think that those actions were interacting with a string even when they were not. The main effect of action properties on object-resemblance ratings highlighted the specificity of our mental models for excitation-resonator interactions of musical instruments.

Additionally, we continuously found an interaction between action and object properties on each object-resemblance rating. By further dissecting the interaction through examination of simple effects, we can more clearly see what types of stimuli confused participants' resemblance ratings. String ratings (Fig. 4.10) were highest for bowed strings. This was expected, because those stimuli represented an interaction between an excitation and a resonator that is typical of acoustic musical instruments. The next highest string ratings were for the struck strings, which also represented a more typical excitation-resonator combination. The blown strings had the next highest string ratings, but the blowing excitation may have primed participants to hear the string as an air column. The string ratings for the struck air column indicated that the striking excitation biased participants to hear the air column as a string. Moreover, struck air columns would be harmonic sounds and striking an air column is not a normal case. The resulting sounds correspond more to popping the hand over the end of a pipe, rather than hitting the hair column, which produces resonances of the pipe shells instead of the air modes. The string ratings were lowest for blown air columns and struck plates: these stimuli represented typical excitation-resonator interactions, so it was obvious that they were not produced by strings.

In the case of air column resemblance ratings (Fig. 4.14), participants assigned the highest ratings to the blown air columns. As these stimuli represented a typical excitation-resonator interaction, it was obvious that an air column did indeed produce these sounds. Air column ratings were the lowest for all struck sounds. The impulsive nature of struck sounds made it obvious to listeners that they were not produced by an air column for the struck string and struck plate stimuli.

Additionally, the impulsive nature of struck sounds also biased participants not to hear the struck air column as produced by an air column. Bowed air column stimuli were also rated quite high for air column resemblance, which shows that listeners had some sensitivity to perceiving sounds produced by an air column. Moreover, the next highest air column ratings were given to blown string and blown plate stimuli. This demonstrates that blown actions biased participants to perceiving an air column, since listeners' mental models were much stronger for blown air column sounds in comparison to blown string or blown plate sounds.

Plate resemblance ratings (Fig. 4.18) were highest for struck plates. This is what we expected, given that these stimuli represented a typical excitation-resonator interaction. The next highest plate resemblance ratings were given to the struck air columns. The striking action biased participants to be more likely to perceive a struck air column as produced by a plate. In fact, plate resemblance ratings fell on the lower end for bowed and blown plates. This might be due to the continuous nature of the sounds, which are not common for sounds produced by plates. Plate resemblance ratings were lowest for bowed and blown strings and air columns. The actions of bowing and blowing biased participants into thinking that the stimuli could not be produced by a plate, as plates are often associated with impulsive impacts. Overall, it seems the plate resemblance ratings were quite biased by the actions that produce the stimuli rather than the objects.

We are curious whether participants perceived the resonators that produced the sounds. Figure 4.2 shows the mean string, air column, and plate resemblance ratings for stimuli that were produced by these objects. We can see that the highest resemblance ratings were assigned to the actions that actually produced the stimuli. In general, listeners were quite sensitive to the resonators that were produced by Modalys, regardless of the excitations that were applied to them. However, this might be attributed to Modalys being better at simulating resonators than excitations.

In Figure 4.3, we plotted the mean string, air column, and plate resemblance ratings for bowed, blown, and struck stimuli. Listeners assigned the highest string ratings to bowed and struck stimuli. They also gave the highest air column ratings to blown stimuli and the highest plate ratings to struck stimuli. These findings support our hypotheses regarding the biases in our mental models for how musical instruments are played. Listeners were more likely to assign higher resemblance ratings to an object when the action that produced the sound is one that typically interacts with it in acoustic musical instruments. Again, the limitations of our mental models for musical instruments were attributed to the limitations of how excitation methods and resonance structures can behave in the physical world. However, even with these limitations, listeners still demonstrated some sensitivity to the resonance structures of sound sources.

# Chapter 5

# General Discussion

Nine types of excitation-resonator interactions were synthesized with Modalys, which employed physically inspired modeling approaches. Three exemplars for each interaction were chosen as the stimuli for our experiments. Listeners rated how well the stimuli resembled three excitations in Experiment 1 and how well they resembled three resonators in Experiment 2. In our analyses, there was consistently a main effect of the action properties, a main effect of the object properties, and an interaction effect of the action and object properties on the resemblance ratings of excitations (bowing, blowing, striking) and resonators (string, air column, plate).

## 5.1 Summary of results across both experiments

### 5.1.1 Stronger mental models for typical excitation-resonator interactions

For the most part, participants were able to perceive the actual excitations and resonators that produced the stimuli when their interactions were typical of acoustic musical instruments. That is, the highest resemblance ratings were assigned to actual excitations and resonators that produced bowed string, blown air column, struck plate, and struck string stimuli. Moreover, resemblance ratings were much lower for the excitations and resonators that did not produce these particular stimuli. This supports our hypothesis that mental models are stronger for typical excitation-resonator interactions than for their atypical counterparts. Excitation-resonator interactions represent the musical instrument families that listeners are familiar with: bowed and struck string stimuli represent string instruments; blown air column stimuli represent wind instruments; struck plate stimuli represent percussive instruments. Since listeners have exposure to these instrument families, mental models for them influence how they perceive novel stimuli.

### 5.1.2 Differences between musicians and nonmusicians

Overall, there were minimal differences in how musicians and nonmusicians rated the stimuli. Some specific differences will be explained in this section. One difference in bowing ratings is worth mentioning: musicians' bowing resemblance ratings were lower than those of nonmusicians for struck string and struck air column stimuli. Musicians were probably more sensitive to differences between continuous (bowed and blown) and impulsive (struck) sounds. As soon as they perceived the struck string and struck air column as impulsive, they knew that bowing did not produce them. Nonmusicians, on the other hand, might have perceived struck air column stimuli as produced by a string. This hypothesis was derived from the interaction effect between musicianship, action properties, and object properties. Nonmusicians might have perceived a string from the struck string and struck air column stimuli, which may have biased them to hear these sounds as bowed, since bowing commonly interacts with strings in acoustic musical instruments. This demonstrates that nonmusicians' mental models were limited for bowing excitations.

We found an interaction effect between musicianship, action properties, and object properties on blowing ratings. Nonmusicians rated blown string stimuli higher on blowing than musicians, meaning that they perceived the blown string as produced by blowing more so than the musicians did. This may be attributed to the different ways in which musicians and nonmusicians processed the stimuli. Musicians seem to rely on their mental models: once they heard the string, they were biased to perceive bowing instead of blowing. Nonmusicians might have processed the blown string from a bottom-up mechanism. Their blowing ratings for the blown string did not seem to be limited by mental models which made it easier for them to separate out excitations from resonators to make their judgements. Another interesting difference was the difference in blowing ratings for bowed plate stimuli between musicians and nonmusicians. Musicians' blowing ratings were lower, meaning that nonmusicians perceived the bowed plate to be produced by blowing to a greater degree than did musicians. We speculate that this could be attributed to the notion that blown and bowed sounds are continuous. When nonmusicians heard continuous stimuli, they might have been biased to perceive them as blown or bowed.

Musicians and nonmusicians hardly differed in their striking ratings. The impulsive nature of struck sounds perhaps made it easier for listeners to distinguish them from the continuous sounds in the other two excitation categories.

Not many differences were seen between musicians' and nonmusicians' string, air column, and plate ratings, with two exceptions. First, musicians gave higher string ratings to string stimuli than did their nonmusician counterparts. Second, musicians' plate ratings for plate stimuli were also higher than those of nonmusicians. Other than that, however, there were minimal differences between musicians' and nonmusicians' string ratings. These results demonstrate that musicians were more likely to perceive strings and plates from stimuli produced by strings and plates, respectively.

### 5.1.3 Were ratings based on Modalys' efficacy or listeners' mental models?

Ratings for actions and objects were highest for the excitations and resonators that actually produced the stimuli, respectively. These results indicate that Modalys can generally convey different excitations and resonators effectively, depending on the type of interaction.

For typical interactions—bowed string, blown air column, struck plate, and struck string—listeners' ratings were highest for the excitations and resonators that actually produced the sounds. Bowing and string ratings were highest for bowed strings (Fig. 5.1a). Listeners rated blown air columns high for blowing and air column (Fig. 5.1e). Ratings for struck plates were highest for striking and plate (Fig. 5.1i). Lastly, listeners assigned high ratings to striking and string for struck strings (Fig. 5.1g). This is not surprising, since our mental models are strongest for these types of excitation-resonator interactions, as they are typical of acoustic musical instruments. Therefore, Modalys' ability to convey different excitation methods and resonance structures for typical interactions is confirmed by our perceptual data.

The atypical interactions included bowed air columns, bowed plates, blown strings, blown plates, and struck air columns. Bowed air columns were rated highest for air column; however, instead of having high ratings for bowing, they had high ratings for blowing (Fig. 5.1b). This result is interesting as it is contrary to previous studies that have demonstrated a greater sensitivity to actions than materials or objects (Lemaitre & Heller, 2012; Hjortkjær & McAdams, 2016). However, these studies were conducted on impacted materials, and the current study was conducted on musical stimuli employing both continuous and impulsive excitations. One possible explanation for the higher blowing ratings of bowed air column stimuli could be attributed to our mental models. Since air column ratings were high for bowed air columns, participants may have been biased to perceive blowing, which is typically applied to air columns in the physical world. A second possible explanation might be that Modalys was not entirely effective at conveying a bowing excitation for the bowed air column.

Ratings for bowed plates were highest for the plate and striking (Fig. 5.1c). Again, this result deviates from what previous studies predicted because listeners identified the plate but not bowing. Judgements may have been influenced by listeners' strong mental models for struck plates, which biased them toward hearing the bowed plate as struck rather than bowed. On the other hand, a more likely explanation is that Modalys did not convey bowing very well with bowed plate stimuli. The reason why this explanation is more likely is because there is an artifact at the beginning of the sound that is quite impulsive and can be perceived as a strike. If we removed this part of the sound, listeners might have rated these sounds differently.

Blown strings had higher ratings for bowing than blowing and ratings for string and air column were very similar (Fig. 5.1d). As bowing is typically applied to strings and blowing is typically applied to air columns, participants were confused by these excitations and resonators. Confusing bowing for blowing may have to do with the fact that both bowed and blown sounds result from continuous excitations. Confusing strings and air columns might be explained by mental models for bowed string and blown air columns; but Modalys may have also been less effective at physically modeling strings when blowing is applied to them.

Blown plates were rated highest for resembling blowing and an air column (Fig. 5.1f). This is an interesting case, as it was one of the interactions that was most difficult to synthesize. In fact, the majority of the time, blown plate stimuli were perceived as blown air columns to the experimenters. With spectrogram analysis, we found that the frequency spectrum shared a large resemblance to that of the struck plate. However, listeners' ratings confirmed that an air column was more likely to be perceived than a plate. This is most likely due to Modalys' estimation of the modes that represent a plate.

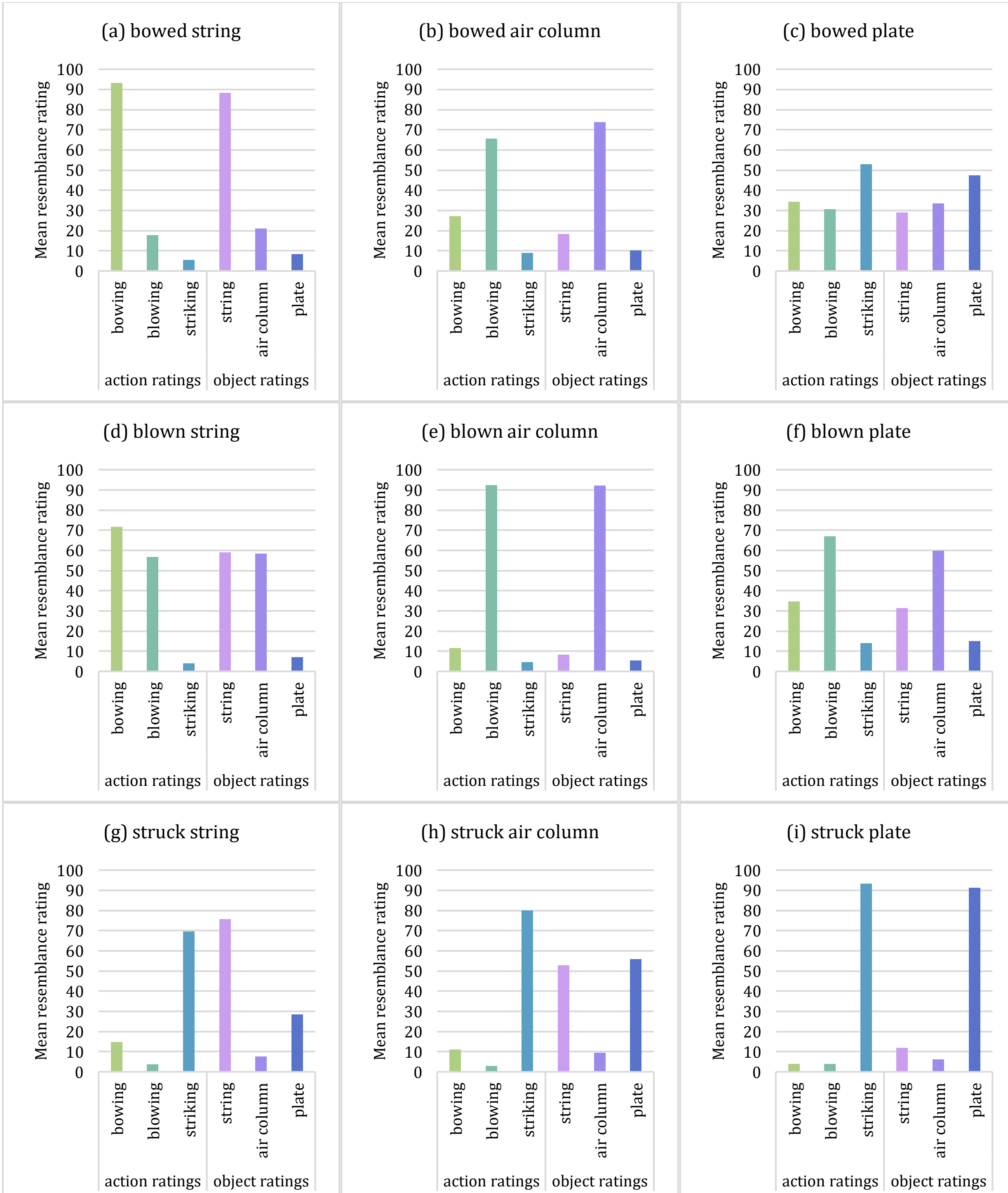


**Figure 5.1.** Mean bowing, blowing, striking, string, air column, and plate ratings for each type of action-object interaction: (a) bowed string, (b) bowed air column, (c) bowed plate, (d) blown string, (e) blown air column, (f) blown plate, (g) struck string, (h) struck air column, and (i) struck plate.

Ratings of struck air columns were highest for striking and very similar for string and plate (Fig 5.1h). This is consistent with previous literature describing greater sensitivity to actions than materials (Lemaitre & Heller, 2012; Hjortkjær & McAdams, 2016). Although these results might indicate that Modalys was not entirely effective at conveying an air column, a more probable explanation might have to do with listeners' mental models for struck sounds. Since striking would most typically apply to plates and strings in the physical world, this might bias listeners to perceive a string or plate even though the stimuli were produced by an air column.

## 5.2 Using perceptual data to inform physically inspired modeling

Our stimuli were synthesized with Modalys, which employs physical modeling approaches. The advantage of using Modalys is that it can apply different excitations to different resonators, even if these interactions are not physical by nature. On the other hand, modeling sounds that would be impossible to generate in the physical world is a complicated task. With our stimuli, we attempted to apply as much physical control of the parameters—choosing to manipulate only two parameters for each type of excitation-resonator interaction. In most cases, we synthesized stimuli that resembled typical excitation-resonator interactions, estimated the limits of the parameter manipulations, and then applied the same manipulations to different interactions. This proved to be the most effective technique as it was simpler to rely on our perceptual judgements when we synthesized the typical interactions. For the atypical interactions, we also relied on our perceptual judgements, but with the assumption that our parameter manipulations for typical interactions were correct.

Not surprisingly, the resemblance ratings were highest for excitations and resonators that actually produced stimuli that resembled typical interactions between them. Listeners were more confused, however, when distinguishing between excitations or resonators that produced the atypical interactions. Whether these confusions can be attributed to the specificity of mental models for how musical instruments are typically played or limitations of physically inspired modeling is a question worth exploring. Given that our mental models for excitation-resonator interactions are very specific for acoustic musical instruments, it may be difficult for listeners to perceive these mechanical properties outside of their typical contexts.

We generated the sounds using as much physical control of the parameters as possible. Spectrograms were analyzed and highlighted two general patterns: (1) stimuli produced by the same excitation methods have similar temporal envelopes; (2) stimuli produced by the same resonance structures have similar frequency spectra. The physical manipulations and spectrogram analysis we performed confirmed that Modalys can synthesize the different excitations and resonators. However, the results of our experiments demonstrated that listeners do not always perceive the intended excitation method or resonance structure. Thus, if physically inspired modeling manipulations are applied to a sound, to what extent are they informative if listeners cannot perceive those manipulations, or if they perceive something else? We could argue that more perceptual studies should be conducted to test the efficacy and limitations of physical models. By doing so, parameter manipulations could then be based on what listeners actually perceive, rather than just the physics. For example, with our stimuli, we might consider other exploratory approaches of manipulating different parameters, since the ones we have chosen for the atypical interactions did not seem to convey the intended excitations and resonators to listeners. Therefore, we could attempt to manipulate a variety of parameters to see which manipulations convey the intended mechanical properties the best.

## 5.3 Potential role of audio descriptors

Audio descriptors play a communicative role between the mechanical properties of sound sources and timbre perception. Sound sources are heard as having a timbre and can be described by audio descriptors, some of which seem to be related to certain aspects of timbre. We reported three audio descriptors of our stimuli in Chapter 2. The first was temporal centroid, which distinguished impulsive (struck) from continuous (bowed and blown) sounds. Impulsive sounds had lower temporal centroids and continuous sounds had higher temporal centroids. The second audio descriptor we analyzed was the spectral centroid, which did not seem to separate different resonators from one another. Lastly, the third audio descriptor was inharmonicity, which also did not distinguish the resonators. Different resonators may be distinguished by other descriptors, such as the spectral shape and its evolution over time or the odd-to-even ratio.

Struck stimuli were commonly rated as strongly resembling striking. Bowed and blown stimuli were often recognized as not produced by striking, but listeners sometimes mistook bowing and blowing for one another. We plan to conduct a correlational analysis between the variability of some audio descriptors and the perceptual data. We expect to find that bowing, blowing, and striking ratings can be predicted by differences in temporal centroids and logarithm of attack time (i.e., the onset of a sound that also distinguishes impulsive from continuous sounds, Peeters et al., 2011). On the other hand, string, air column, and plate ratings might be influenced by variability in spectral flux or other measures of spectral shape, which is the amount of variation of a signal's spectrum over time (Peeters et al., 2011).

## 5.4 Limitations

When we looked at listeners' ratings for the resemblance of the three different excitations, they were highest for the actions that actually produced the sounds, regardless of what object they were applied to (Experiment 1). Similarly, ratings for the resemblance of three different resonators were highest for the objects that actually produced the sound across the actions that were applied to them (Experiment 2). Ratings for specific excitation-resonator interactions indicated that listeners had highest ratings for actions and objects that produced typical interactions, but they were often confused for atypical interactions. Although these results align with our hypotheses, there were a few limitations to our study.

Although participants rated how well the stimuli resembled different excitations and resonators, the ratings themselves may not be as informative as we intended. The ratings do not entirely tell us whether listeners were influenced more by their mental models, Modalys' ability to convey excitations and resonators, or some combination of the two. Moreover, it is difficult to say whether listeners can actually identify the excitation or resonator that produced a stimulus. Even if ratings are "above chance", how much above should they be for us to conclude that listeners can perceive the excitations and resonators producing the stimuli? In this case, a categorization task might be useful; it could complement the current results with a forced-choice paradigm. We would then correlate the current ratings with discrete categorization to see whether listeners can actually distinguish between the mechanical properties.

Listeners only rated the resemblance of actions and objects for three versions of each of the nine excitation-resonator interactions. Moreover, each interaction was manipulated with very controlled approaches that were applied to only two parameters. The specific control we applied to each of the interactions might not be generalizable to how acoustic musical instruments are played in the physical world. A performer does not manipulate only two parameters at a time when they play different notes on their instrument; a variety of parameters are manipulated that change the timbre of the resulting sound. However, even with the two parameters we manipulated for each interaction, the timbre of the resulting sound changed substantially. Moreover, previous studies have shown that changing just one parameter (e.g., pitch) can alter how the timbre is perceived (Krumhansl & Iverson, 1992; Handel & Erickson, 2001; Marozeau et al., 2003). Also, the parameters we manipulated in our stimuli are among those that are commonly manipulated in the physical modeling of acoustic musical instruments: bow speed and pressure (Halmrast et al., 2010); breath pressure and squeeze parameter (Dalmont et al., 2005); and a variety of options for percussive sounds (Halmrast et al., 2010). Using different types of manipulations would be of interest for another exploratory approach in sound synthesis of typical and atypical excitation-resonator interactions.

Another limitation can be attributed to the recruitment of participants. It was easiest to find musician participants, as there is a mailing list for music students at McGill University. Nonmusicians were primarily recruited from McGill's certified Facebook page and from our lab's participant pool. However, nonmusicians showed less interest in participating in our experiments than musicians, which is why we asked some of them to return as participants for the experiment they had not taken part in. Moreover, some of the nonmusicians misrepresented the number of years of formal musical training they had or the number of instruments they played. We contacted the nonmusician participants following their participation to confirm their musical background. Even after including their confirmed responses, t-tests revealed a significant difference in the number of years of formal musical training between musician and nonmusician participants for both experiments. It might be interesting to examine the effect of musicianship as a covariate factor, to see if patterns of ratings change with increasing years of formal musical training.

## 5.5 Future directions

As mentioned, we would be interested in conducting a categorization task with our stimuli, to see if those results complement those of the current study. Moreover, a categorization task would indicate if listeners actually perceive the excitations and resonance structures we synthesized. For the resemblance ratings of the current study, it is difficult to discern how high a rating should be before we can infer that participants perceive that a specific excitation or resonator produced a stimulus. By correlating the results of a categorization task with those of the current study, we can examine whether the values of the ratings predict whether or not particular excitations or resonators are perceived.

This categorization task would follow a similar paradigm to that of Hjortkjær and McAdams (2016). These authors presented their stimuli to participants one at a time and asked them to identify either the action or material that caused the sounds. Their procedure included a forced-choice judgement with three options that corresponded to either three actions or three materials. The categorization task we would conduct during my doctoral studies would follow the same procedure.

Another procedure worth examining is that of the dissimilarity ratings. Similar to previous studies (McAdams et al., 1995; Caclin et al., 2005), listeners would be presented with pairs of stimuli and rate how dissimilar they perceive them to be. These results would complement those of the current study and categorization tasks: stimuli produced by the same category of excitation methods or resonance structure should be rated as more similar. If this holds true for the atypical excitation-resonator interactions, we can infer that participants can disentangle the mechanical properties implicitly: dissimilarity ratings reflect judgements that are based on the mechanical properties that are available to compare (i.e., just the excitation methods, just the resonance structure, both, or neither). We can then examine which audio descriptors contribute to the dissimilarity judgements with MDS models.

## 5.6 Conclusion

In the current study, we intended to examine whether our mental models for acoustic musical instruments influence how we perceive different excitation methods, resonance structures, and their interactions. We used physically inspired modeling approaches to synthesize three excitation methods (bowing, blowing, and striking) and three resonance structures (string, air column, and plate). We combined each excitation with each resonator to create nine types of interactions. Listeners rated how well each sound resembled the three excitations in the first experiment and how well they resembled the three resonators in the second experiment.

We found that overall, listeners rated stimuli highest for excitations and resonators that actually produced the stimuli. When examining resemblance ratings for each type of excitation-resonator interaction, ratings were only highest for excitations and resonators that produced the stimuli when the interaction was typical of acoustic musical instruments (bowed string, blown air column, struck plate, and struck string). Ratings for atypical interactions were higher for either excitation methods or resonance structures that did not produce the sounds. These findings confirm that our mental models are stronger for typical excitation-resonator interactions, as we have more exposure to these interactions than their atypical counterparts. Moreover, these findings speak to the abilities of physically inspired modeling techniques to convey different excitation methods and resonance structures. Our findings indicate that it was either difficult for participants to perceive the excitations or resonators that interact atypically, that physically inspired modeling approaches cannot entirely convey what listeners simply do not have the mental models for, or that such interactions have no physical basis and cannot be either properly modeled or produced. Consequently, the results of our perceptual studies can inform physical modeling approaches. Taken together, our studies reveal how timbre perception provides listeners with information about sound source mechanics to form mental models of musical instruments. We can then infer how novel sounds become integrated into mental models of sound sources in daily life.

## Appendix 1: Example 3—"tube, reed and hole connection" by Modalys

```
(new)

; RESONATOR: AIR COLUMN
(defvar my-tube)
(setq my-tube
      (set-pitch (make-object 'closed-open-tube
                              (modes 40)
                              (radius0 .01)
                              (radius1 .01))
                 'length 80.0))

; EXCITER: REED
(defvar my-reed)
(setq my-reed (make-object 'bi-two-mass
                           (small-mass   .000002)
                           (large-mass   .000002)
                           (stiffness0      200)
                           (stiffness1      200)
                           (const-loss0      60)
                           (const-loss1      60)
                           (freq-loss0      260)
                           (freq-loss1      260)))

; BREATH ENVELOPE
(defvar breath-env2)
(setq breath-env2 (make-controller 'envelope 1
                                   (list (list 0.00   70)
                                         (list 0.10   70)
                                         (list 0.30   70)
                                         (list 0.60   75)
                                         (list 0.80   75)
                                         (list 1.00  100)
                                         (list 1.30  100)
                                         (list 1.60  120)
                                         (list 2.00  120)
                                         (list 2.10  100)
                                         (list 2.50  100)
                                         (list 2.60   75)
                                         (list 3.00   75)
                                         (list 3.10   70)
                                         (list 3.50   70)
                                         (list 4.25    0))))

; ACCESSES
(defvar my-tube-top)
```

```
(setq my-tube-top (make-access my-tube (const 0) 'long))
(defvar my-reed-tip)
(setq my-reed-tip (make-access my-reed (const 1) 'trans0))

(make-connection 'reed my-tube-top my-reed-tip 0.001
                 (make-controller 'arithmetic 1 '* (list (const 6) breath-
env2))
                 (const 0.000276))

; EXCITATION: BLOWING
(defvar my-reed-bse)
(setq my-reed-bse (make-access my-reed (const 0) 'trans0))
(make-connection 'position my-reed-bse (const 0))

; MAKE THE HOLES
(defvar hole-radius1)
(setq hole-radius1 (make-controller 'dynamic 1 -1 (list .0)))
(defvar hole-radius2)
(setq hole-radius2 (make-controller 'dynamic 1 -1 (list .0)))
(defvar hole-radius3)
(setq hole-radius3 (make-controller 'dynamic 1 -1 (list .0)))


(defvar my-tube-hle1)
(setq my-tube-hle1 (make-access my-tube (const 0.7937) 'long))
(make-connection 'hole my-tube-hle1 hole-radius1 (const 0.))

(defvar my-tube-hle2)
(setq my-tube-hle2 (make-access my-tube (const .6674) 'long))
(make-connection 'hole my-tube-hle2 hole-radius2 (const 0.))

(defvar my-tube-hle3)
(setq my-tube-hle3 (make-access my-tube (const .5) 'long))
(make-connection 'hole my-tube-hle3 hole-radius3 (const 0.))

; OUTPUT
(defvar my-tube-out)
(setq my-tube-out (make-access my-tube (const .25) 'long))
(make-point-output my-tube-out)

(run 0.5)

(set-breakpoint hole-radius1 (list 0.05 0.01)) ;; open first hole in 0.05 sec
(run 0.5)

(set-breakpoint hole-radius2 (list 0.05 0.01)) ;; open the second hole
(run 0.5)

(set-breakpoint hole-radius3 (list 0.05 0.01)) ;; open the third hole
(run .5)

(set-breakpoint hole-radius3 (list 0.05 0.0)) ;; close the third hole
(run .5)

(set-breakpoint hole-radius2 (list 0.05 0.0)) ;; close the second hole
(run .5)
```

```
(set-breakpoint hole-radius1 (list 0.05 0.0)) ;; close the first hole
(run 1.5)

(play)
```

## Appendix 2: Example 6—“bowed string” by Modalys

```
(new)
(load-component "connection")

; RESONATOR: STRING
(defvar my-string)
(setq my-string (make-object 'bi-string
                             (modes 80)
                             (length .5)
                             (tension 150)
                             (density 1000)
                             (radius .001)
                             (young .001)
                             (freq-loss 1)
                             (const-loss 1)))

; EXCITER: BOW
(defvar my-bow)
(setq my-bow (make-object 'bi-two-mass
                          (small-mass .05)
                          (large-mass .05)
                          (stiffness0 50000)
                          (stiffness1 50000)
                          (freq-loss0 100)
                          (freq-loss1 100)
                          (const-loss0 0)
                          (const-loss1 0)))

; ACCESSES
(defvar my-bow-h-bpt)
(setq my-bow-h-bpt (make-access my-bow 1 'trans0))

(defvar my-bow-v-bpt)
(setq my-bow-v-bpt (make-access my-bow 1 'trans1))

(defvar my-string-h-bpt)
(setq my-string-h-bpt (make-access my-string .1 'trans0))

(defvar my-string-v-bpt)
(setq my-string-v-bpt (make-access my-string .1 'trans1))

; EXCITATION: BOWING
(defvar bc)
(setq bc (make-connection 'bow
                          my-bow-v-bpt my-bow-h-bpt .01
                          my-string-v-bpt my-string-h-bpt
```

```
                           '(2 10 5 4)))

(defvar my-bow-h-mov)
(setq my-bow-h-mov (make-access my-bow (const 0) 'trans0))

(defvar my-bow-v-mov)
(setq my-bow-v-mov (make-access my-bow (const 0) 'trans1))

; BOW SPEED
(make-connection 'speed my-bow-h-mov
                 (make-controller 'envelope 1
                                  (list (list 0.00  1)
                                        (list 0.10  4)
                                        (list 9.99  1))))

; BOW PRESSURE
(make-connection 'position my-bow-v-mov
                 (make-controller 'envelope 1
                                  (list (list 0.0  0.010)
                                        (list 0.2 -0.001)
                                        (list 0.4 -0.001)
                                        (list 0.5  0.010)
                                        (list 0.6  0.010)
                                        (list 0.7 -0.001)
                                        (list 1.2 -0.001)
                                        (list 1.3  0.010)
                                        (list 9.9  0.010))))

; OUTPUT
(make-point-output (make-access my-string (const .6) 'trans0))
(run 3)
(play)
```

## Appendix 3: Example 13—"simple blow" by Modalys

```
(new)
(set-precision 'float)

; BREATH ENVELOPE
(setq breath-env (make-controller 'dynamic 1 -1 0 "breath-pressure"))

(setq filtered-breath-env
                       (make-controller
                        'variable-second-order-filter 1 0
                        breath-env
                        (make-controller 'dynamic 1 -1 0 "filter-f")
                        (make-controller 'dynamic 1 -1 5 "filter-w")))

(setq filtered-breath-env breath-env)

; RESONATOR: STRING
(setq resonator (make-object 'mono-string))

; ACCESSES OF RESONATOR
(setq resonator-in (make-access resonator (const 0.1) 'trans0))
(setq resonator-out (make-access resonator (const .5) 'trans0))
(setq rad (make-object 'radiator (angle 90)(radius 3e-2)))
(setq rad-1 (make-access rad (const 1) 'normal))
(make-connection 'adhere resonator-out rad-1)


; EXCITER: REED
(setq reed
      (make-object 'single-point
         (freqs (make-controller 'dynamic 1 -1 1400 "reed-freqs"))
         (bws (make-controller 'dynamic 1 -1 50000 "reed-bws"))
         (amps (const 1))))

; ACCESS OF REED
(setq reed-access (make-access reed (const 1) 'normal))

; EXCITATION: BLOWING
(setq reed-connection
      (make-connection 'normalised-valve reed-access 0.001 resonator-in
                 filtered-breath-env
                 (const 1.2)      ; air density
                 (make-controller 'dynamic 1 -1 3 "valve-zeta")
                 (const 180)
                 (const 1)
```

```
))


(setq pl (make-plot))
(plot-value pl "speed" (make-controller 'access-speed 1 rad-1))

; OUTPUT
(make-point-output rad-1)

(run .1)
(send-message "breath-pressure" .7 0) ;; BREATH PRESSURE
(run 1)
(send-message "breath-pressure" 0 .2) ;; BREATH PRESSURE
(run 2)
(play)

(get-info 'max-sample)
(plot pl "all")
```

# Appendix 4: Example 2b—“plate and hammer” by Modalys

```
(new)

; RESONATOR: PLATE
(defvar my-plate)
(setq my-plate (set-pitch (make-object 'rect-plate
                                        (thickness .01)
                                        (density 10000)
                                        (freq-loss 0.1)
                                        (const-loss 0.5))
                          'size 330))

; EXCITER: HAMMER
(defvar my-hammer)
(setq my-hammer (make-object 'bi-two-mass (small-mass 0.5)
                              (large-mass 0.5)
                              (stiffness0 100000)
                              (stiffness1 150000)))

; ACCESSES
(defvar my-plate-hit)
(setq my-plate-hit (make-access my-plate (const .6 .7) 'normal))
(defvar my-hammer-hit)
(setq my-hammer-hit (make-access my-hammer (const 1) 'trans0))

; EXCITATION: STRIKING
(make-connection 'strike my-plate-hit 0 my-hammer-hit .1)

(defvar my-hammer-mov)
(setq my-hammer-mov (make-access my-hammer (const 0) 'trans0))

; HAMMER PRESSURE
(make-connection 'position my-hammer-mov
             (make-controller 'envelope 1
                                  (list (list 0.00   .1)
                                        (list 0.05  -.0001)
                                        (list 0.10   .1)
                                        (list 0.50   .1)
                                        (list 0.55  -.0001)
                                        (list 0.60   .1))))

; OUTPUT
(defvar my-plate-out)
(setq my-plate-out (make-access my-plate (const .2 .1) 'normal))
(make-point-output my-plate-out)
```

```
(run 10)
(play)
```

## References

Abdi, H. (2010). Holm's sequential Bonferroni procedure. In N. Salkind (Ed.), *Encyclopedia of research design,* vol. 1 (pp. 1–8). Thousand Oaks, CA: Sage.

Agus, T., Suied, C., Thorpe, S., & Pressnitzer, D. (2012). Fast recognition of musical sounds based on timbre. *Journal of the Acoustical Society of America*, *131*(5), 4124–4133.

American National Standards Institute (1994). *American national standard acoustical terminology, ANSI S1.1-1994*. New York, NY: American National Standards Institute.

Aramaki, M., Besson, M., Kronland-Martinet, R., & Ystad, S. (2009). Timbre perception of sounds from impacted materials: Behavioral, electrophysiological, and acoustical approaches. In S. Ystad, R. Kronland-Martinet & K. Jensen (Eds.), *Computer music modeling and retrieval 2008*, LNCS 5493 (pp. 1–17). Berlin, Heidelberg: Springer.

Bensoam, J., & Roze, D. (2013). Solving interactions between nonlinear resonators. In *Sound and Music Computing Conference 2013* (pp. 576–583). Stockholm, Sweden.

Berger, K. W. (1964). Some factors in the recognition of timbre. *Journal of the Acoustical Society of America*, *36*(10), 1888–1891.

Böttcher, N., Gelineck, S., & Serafin, S. (2007, June). PHYSMISM: A control interface for creative exploration of physical models. In *Proceedings of the 2007 Conference on New Interfaces for Musical Expression* (pp. 31–36). New York, NY.

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.

Conan, S., Thoret, E., Aramaki, M., Derrien, O., Gondre, C., Ystad, S., & Kronland-Martinet, R. (2014). An intuitive synthesizer of continuous-interaction sounds: Rubbing, scratching, and rolling. *Computer Music Journal*, *38*(4), 24–37.

Coyle, W. L., Guillemain, P., Kergomard, J., & Dalmont, J. P. (2015). Predicting playing frequencies for clarinets: a comparison between numerical simulations and simplified analytical formulas. *Journal of the Acoustical Society of America*, *138*(5), 2770–2781.

Dalmont, J. P., Gilbert, J., Kergomard, J., & Ollivier, S. (2005). An analytical prediction of the oscillation and extinction thresholds of a clarinet. *Journal of the Acoustical Society of America*, *118*(5), 3294–3305.

Dudas, R. (2014). Modalys, Version 3.4.1 [computer software]. Paris: Institut de recherche et coordination acoustique/musique.

Eckel, G., Iovino, F., & Caussé, R. (1995, July). Sound synthesis by physical modeling with Modalys. In *Proceedings of the International Symposium on Musical Acoustics* (pp. 479–482). Dourdan, France.

Elliott, C. A. (1975). Attacks and releases as factors in instrument identification. *Journal of Research in Music Education*, *23*(1), 35–40.

Expressive E (2019). Arché [computer software]. Montreuil: Expressive.

Fabiani, M., & Friberg, A. (2011). Influence of pitch, loudness, and timbre on the perception of instrument dynamics. *Journal of the Acoustical Society of America*, *130*(4), EL193–EL199.

Giordano, B. L., & McAdams, S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *Journal of the Acoustical Society of America*, *119*(2), 1171–1181.

Giordano, B. L., & McAdams, S. (2009). The perception of musical timbre. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 72–80). Oxford, NY: Oxford University Press.

Giordano, B. L., & McAdams, S. (2010). Sound source mechanics and musical timbre perception: Evidence from previous studies. *Music Perception*, *28*(2), 155–168.

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, *61*(5), 1270–1277.

Grey, J. M., & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America*, *63*(5), 1493–1500.

Haack, P. A. (1975). The influence of loudness on the discrimination of musical sound factors. *Journal of Research in Music Education*, *23*(1), 67–77.

Halmrast, T., Guettler, K., Bader, R., & Godøy, R. I. (2010). Gesture and timbre. In R. I. Godøy

& M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 183–211). New York, NY: Taylor & Francis.

Handel, S., & Erickson, M.L. (2001). A rule of thumb: The bandwidth for timbral invariance is one octave. *Music Perception*, *19*(1), 121–126.

Handel, S., & Erickson, M. L. (2004). Sound source identification: The possible role of timbre transformations. *Music Perception*, *21*(4), 587–610.

Hjortkjær, J., & McAdams, S. (2016). Spectral and temporal cues for perception of material and action categories in impacted sound sources. *Journal of the Acoustical Society of America*, *140*(1), 409–420.

Houben, M., Kohlrausch, A., & Hermes, D. J. (2004). Perception of the size and speed of rolling ball by sound. *Speech Communication, 43*(4), 331–345.

IBM Corp. (2015). ICM SPSS Statistics for Macintosh, Version 23.0 [computer software]. Armonk, NY: IBM Corp.

ISO 389–8. (2004). Acoustics – Reference zero for the calibration of audiometric equipment – Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones (Tech. Rep.). Geneva, Switzerland: International Organization for Standardization.

Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America*, *94*(5), 2595–2603.

Klatzky, R. L., Pai, D. K., & Krotkov, E. P. (2000). Perception of material from contact sounds. *Presence: Teleoperations and Virtual Environments*, *9*(4), 399–410.

Kozak, M., & Piepho, H.-P. (2018). What's normal anyway? Residual plots are more telling than significance tests when checking ANOVA assumptions. *Journal of Agronomy and Crop Science*, *204*(1), 86–98.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 739–751.

Lakatos, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception and Psychophysics*, *62*(7), 1426–1439.

Lartillot O., Toiviainen P., Eerola T. (2008). A Matlab toolbox for music information retrieval.

In Preisach C., Burkhardt H., Schmidt-Thieme L., Decker R. (Eds.), *Data Analysis, Machine Learning and Applications* (pp. 261–268). Berlin, Heidelberg: Springer.

Lemaitre, G., & Heller, L. M. (2012). Auditory perception of material is fragile while action is strikingly robust. *Journal of the Acoustical Society of America*, *131*(2), 1337–1348.

Lutfi, R., & Oh, E. L. (1997). Auditory discrimination of material changes in a struck-clamped bar. *Journal of the Acoustical Society of America*, *102*(6), 3647–3656.

Mathworks. (2018). MATLAB, Version 9.4.0.813654 [computer software]. Natick: MathWorks Incorporated.

Martin, F. N., & Champlin, C. A. (2000). Reconsidering the limits of normal hearing. *Journal of the American Academy of Audiology, 11*(2), 64–66.

Marozeau, J., de Cheveigné, A., McAdams, S., & Winsberg, S. (2003). The dependency of timbre on fundamental frequency. *Journal of the Acoustical Society of America*, *114*(5), 2946–2957.

McAdams, S. (2013). Musical timbre perception. In D. Deutsch (Ed.), *The Psychology of Music* (3rd ed., pp. 35–67). San Diego, CA: Academic Press.

McAdams, S. (2019). The perceptual representation of timbre. In K. Siedenbeurg, C. Saitis, S. McAdams, R. N. Popper, & R. N. Fay (Eds.), *Timbre: Acoustics, Perception, and Cognition* (pp. 23-57). Berlin, Heidelberg: Springer Nature.

McAdams, S., & Goodchild, M. (2017). Musical structure: Sound and timbre. In R. Ashley & R. Timmers (Eds.), *The Routledge Companion to Music Cognition* (pp. 129–139). New York, NY: Taylor & Francis.

McAdams, S., Chaigne, A., & Roussarie, V. (2004). The psychomechanics of simulated sound sources: Material properties of impacted bars. *Journal of the Acoustical Society of America*, *115*(3), 1306–1320.

McAdams, S., Roussarie, V., Chaigne, A., & Giordano, B. L. (2010). The psychomechanics of simulated sound sources: Material properties of impacted thin plates. *Journal of the Acoustical Society of America*, *128*(3), 1401–1413.

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*(3), 177–192.

McMillan, A. J. (1997). A non-linear friction model for self-excited vibrations. *Journal of Sound* *and Vibration*, *205*(3), 323–335.

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

Modartt (2019). Pianoteq, Version 6.5.2 [computer software]. Tolouse: Modartt Software.

Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. *Journal of the American Statistical Association*, *69*(348), 894–908.

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams. (2011). The timbre toolbox: Extracting audio descriptors from musical signals. *Journal of the Acoustical Society of America*, *130*(5), 2902–2916.

Saldanha, E. L., & Corso, J. F. (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America*, *36*(11), 2021–2026.

Smith, B. K. (1995). PsiExp: An environment for psychoacoustic experimentation using the IRCAM musical workstation. In *Society for Music Perception and Cognition Conference '95*. Berkeley, CA: University of California, Berkeley.

Steele, K. M., & Williams, A. K. (2006). Is the bandwidth for timbre invariance only one octave? *Music Perception*, *23*(3), 215–220.

Srinivasan, A., Sullivan, D., & Fujinaga, I. (2002). Recognition of isolated instrument tones by conservatory students. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & E. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney* (pp. 720–723). Adelaide, Australia: Causal Productions.

van Dinther, R., & Patterson, R. D. (2006). Perception of acoustic scale and size in musical instrument sounds. *Journal of the Acoustical Society of America*, *120*(4), 2158–2176.

Vigué, P., Vergez, C., Karkar, S., & Cochelin, B. (2017). Regularized friction and continuation. *Journal of Sound and Vibration*, *389*, 350–263.

Warren, W. H., & Verbrugge, R. R. (1984). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 704–712.

Ystad, S., Aramaki, M., & Kronland-Martinet, R. (2019). Timbre from sound synthesis and high-level control perspectives. In K. Siedenbeurg, C. Saitis, S. McAdams, R. N. Popper, & R. N. Fay (Eds.), *Timbre: Acoustics, Perception, and Cognition* (pp. 361-389). Berlin, Heidelberg: Springer Nature.