# Effect of back wood choice on the perceived quality of steel-string acoustic guitars

Samuele Carcagno, Roger Bucknall, Jim Woodhouse, Claudia Fritz, Christopher Plack

## To cite this version:

Samuele Carcagno, Roger Bucknall, Jim Woodhouse, Claudia Fritz, Christopher Plack. Effect of back wood choice on the perceived quality of steel-string acoustic guitars. Journal of the Acoustical Society of America, 2018, 144 (6), pp.3533-3547. 10.1121/1.5084735 . hal-02025913

# Effect of back wood choice on the perceived quality of steel-string acoustic guitars

**DOI:**
10.1121/1.5084735

**Document Version**
Accepted author manuscript

**Citation for published version (APA):**
Carcagno, S., Bucknall, R., Woodhouse, J., Fritz, C., & Plack, C. (2018). Effect of back wood choice on the perceived quality of steel-string acoustic guitars. *The Journal of the Acoustical Society of America*, *144*(6), 3533-3547. https://doi.org/10.1121/1.5084735

**Published in:**
The Journal of the Acoustical Society of America

The Journal of the Acoustical Society of America

# Effect of back wood choice on the perceived quality of steel-string acoustic guitars

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | JASA-03561R1 |
| **Full Title:** | Effect of back wood choice on the perceived quality of steel-string acoustic guitars |
| **Short Title:** | Guitar back wood |
| **Article Type:** | Regular Article |
| **Corresponding Author:** | Samuele Carcagno, Ph.D.<br>Lancaster University<br>Lancaster, Lancashire UNITED KINGDOM |
| **First Author:** | Samuele Carcagno, Ph.D. |
| **Order of Authors:** | Samuele Carcagno, Ph.D. |
| | Roger Bucknall |
| | Jim Woodhouse |
| | Claudia Fritz |
| | Christopher J Plack |
| **Section/Category:** | Musical Acoustics |
| **Keywords:** | musical instruments;  endangered timbers;  subjective evaluation;  perception |
| **Abstract:** | Some of the most prized woods used for the backs and sides of acoustic guitars are expensive, rare, and from unsustainable sources. It is unclear to what extent back woods contribute to the sound and playability qualities of acoustic guitars. Six steel-string acoustic guitars were built for this study to the same design and material specifications except for the back/side plates which were made of woods varying widely in availability and price (Brazilian rosewood, Indian rosewood, mahogany, maple, sapele, and walnut). Bridge-admittance measurements revealed small differences between the modal properties of the guitars which could be largely attributed to residual manufacturing variability rather than to the back/side plates. Overall sound quality ratings, given by 52 guitarists in a dimly lit room while wearing welder's goggles to prevent visual identification, were very similar between the six guitars. The results of a blinded ABX discrimination test, performed by another subset of 31 guitarists, indicate that guitarists could not easily distinguish the guitars by their sound or feel. Overall, the results suggest that the species of wood used for the back and sides of a steel-string acoustic guitar has only a marginal impact on its body mode properties and perceived sound. |

# Effect of back wood choice on the perceived quality of steel-string acoustic guitars

Samuele Carcagno,[1] Roger Bucknall,[2] Jim Woodhouse,[3] Claudia Fritz,[4] and Christopher J. Plack[1, 5]

[1)] *Department of Psychology, Lancaster University, Lancaster, LA1 4YF, United Kingdom*[a]

[2] *Fylde Guitars, Penrith, CA11 9BD, UK*

[3] *Engineering Department, Cambridge University, Cambridge, CB2 1PZ, United Kingdom*

[4] *Sorbonne Université, CNRS, Institut Jean Le Rond d'Alembert, 75005, Paris, France*

[5] *Manchester Centre for Audiology and Deafness, University of Manchester, Manchester Academic Health Science Centre, M13 9PL, United Kingdom*

Some of the most prized woods used for the backs and sides of acoustic guitars are expensive, rare, and from unsustainable sources. It is unclear to what extent back woods contribute to the sound and playability qualities of acoustic guitars. Six steel-string acoustic guitars were built for this study to the same design and material specifications except for the back/side plates which were made of woods varying widely in availability and price (Brazilian rosewood, Indian rosewood, mahogany, maple, sapele, and walnut). Bridge-admittance measurements revealed small differences between the modal properties of the guitars which could be largely attributed to residual manufacturing variability rather than to the back/side plates. Overall sound quality ratings, given by 52 guitarists in a dimly lit room while wearing welder's goggles to prevent visual identification, were very similar between the six guitars. The results of a blinded ABX discrimination test, performed by another subset of 31 guitarists, indicate that guitarists could not easily distinguish the guitars by their sound or feel. Overall, the results suggest that the species of wood used for the back and sides of a steel-string acoustic guitar has only a marginal impact on its body mode properties and perceived sound.

---

[a] s.carcagno@lancaster.ac.uk

# I. INTRODUCTION

## A. Back wood in acoustic guitars

The acoustic guitar is one of the most popular musical instruments across the world. In the U.S. alone, more than one million acoustic guitars are sold each year (Music Trades, 2016a). The sound of an acoustic guitar depends in part on the woods used in its construction. It is universally agreed that the top plate is the most important component from this point of view (Richardson, 1994). However, the attention of guitar players is often focused on the choice of wood used for the back and sides of the soundbox (Johnston, 2006; 2013). For brevity, this will be referred to as the back wood.

Unfortunately, many traditionally prized back woods are not only expensive, but also rare and from unsustainable sources. For example, since 1992 Brazilian rosewood (*Dalbergia nigra*) has been in Appendix I of the CITES convention, which lists plant species threatened with extinction (Greenberg, 2016). In 2017 the whole Dalbergia genus was listed in Appendix II of the CITES convention (Bedell, 2017; Music Trades, 2016b). The scarcity and increasing cost of many traditional back woods has pushed guitar makers to explore alternative solutions. These have included the use of less familiar species of tropical woods that are more readily available (Ellis *et al.*, 2008), temperate woods such as maple and walnut, laminates (French and Handy, 2006), and synthetic materials such as carbon fiber or fiberglass composites (Forbes-Roberts, 2008; Pedgley *et al.*, 2009).

Opinions about the acoustical importance of the back wood are widely disparate. The famous maker Antonio de Torres Jurado, usually credited as the originator of the modern classical guitar, was so keen to demonstrate that the back is not very important that he made an instrument with a papier-mâché body (Romanillos, 1997). Other contemporary luthiers have also played down the importance of back wood choice arguing that within-species wood variability can be at least as big as between-species variation and that in any case the maker has a lot of freedom in constructional details that can be exploited to make excellent sounding instruments with any wood that is structurally suitable (Gore, 2011). Nonetheless, many guitar makers, dealers, and players have a strongly contrary view, which holds that the back wood makes a significant and immediately obvious difference to the sound of a guitar. Certain tropical hardwoods are highly prized: Brazilian rosewood in particular is often considered the "gold standard" for acoustic guitar backs. There are abundant anecdotal references to this belief in specialized guitar magazines and books (Bourgeois, 1994; Hunter, 2014; Johnston, 2011; Relph-Knight, 2011; Sandberg and Traum, 2000).

When commencing a new instrument, the maker chooses a wood species and then a particular sample of that species. The choice is not determined solely by acoustical considerations: visual appearance, ability to take the traditional high-gloss finish, stability and resistance to cracking, working qualities on the bench, and long-term resistance to deformation in response to the static stress from the strings are all important as well. The maker then makes some choices of detailed dimensions (shape, thickness, bracing pattern etc.), and the result is a soundbox with certain vibration modes and associated resonance frequencies and damping factors. These modal parameters give the acoustical 'fingerprint' of the instrument, and determine many aspects of the sound. But the sound is also affected by the player's choice of strings, and by details of the final set-up by the maker: for example the configuration of bridge saddle and nut, fretting and action height.

There has been a considerable amount of research on the acoustic properties of wood (Haines, 1979; Wegst, 2006; Yoshikawa, 2007). However, measurements of the mechanical properties of different wood species are of limited value unless they can be clearly related to the mechanical response of the guitar soundbox, and thence to the acoustical properties, and thence to listeners' perceptions and preferences (Fritz and Dubois, 2015; McIntyre and Woodhouse, 1978). There is some scientific literature relating to all these stages, reviewed briefly below, but to date there has been no clear synoptic effort to apply rigorous scientific methods to the overall question: 'does the back wood choice of an acoustic guitar, in itself, produce repeatable and recognisable differences of sound quality?' That is the main task of this study, to be tackled by a combination of physical measurements and blinded perceptual experiments.

### B. Guitar acoustics background

There is a considerable literature relating to the acoustics of the violin, and a somewhat smaller collection relating directly to the guitar. In both cases much emphasis has been placed on understanding a relatively small number of low-frequency modes of vibration: the only ones that an instrument maker can reasonably expect to control in detail (Woodhouse, 2002). For the violin, the picture is quite complicated: a clear physical understanding of these low frequency modes has only appeared quite recently (Gough, 2015), and it is also known that some aspects of behavior at much higher frequency are of great perceptual importance. Furthermore, the non-linear nature of a bowed string introduces another layer of complexity into understanding the "tone" of violins (Woodhouse, 2014).

The guitar is much simpler. Any plucked-string instrument is to a good first approximation a linear system, so that understanding the modes of the complete system should contain *all* relevant information. For the guitar specifically, it has been strongly argued that a few low-frequency modes are responsible for the dominant aspects of radiated sound and thus perception (Richardson, 1994, 2002). Furthermore, these low-frequency modes are rather easy to understand in physical terms. The rib structure of a guitar is much more solid than that of a violin, so there is very little deformation of the ribs at low frequency. The top plate behaves largely as if it were clamped along the rib line, as does the back plate. The coupling between the two is mainly through pressure changes in the air inside the cavity, plus some motion of the rib garland as a rigid mass.

The result is that the expected influence of the back plate is easy to predict. There is a triad of modes resulting from coupling of a 'Helmholtz-like' mode of the air and the lowest modes of the top plate and back plate separately. The other strong low-frequency modes are confined almost entirely to the top plate, with little or no coupling to the air and thus little or no influence of the back plate. So the influence of the back is expected to come mainly through the details of the triad of coupled modes, determined by a combination of the mass of the back and its stiffness in that lowest clamped mode. This mass and stiffness relate directly to things a maker is likely to be able to "feel in the hands" when working on the free back plate (Christensen and Vistisen, 1980; Fletcher and Rossing, 1998; Richardson, 1994, 2002). It follows that a skilled guitar maker might, to a very large extent, be able to compensate for the properties of different back woods by adjusting constructional details such as the thickness of the back plate, to achieve very similar-sounding instruments using different woods.

### C. Guitar psychoacoustics background

Some psychoacoustic experiments have been carried out on guitar sounds (Woodhouse *et al.*, 2012; Wright, 1996). Perceptual thresholds have been established when controlled changes are made to the modal properties of the instrument body (Woodhouse *et al.*, 2012). Most of the results were for nylon-strung guitars, but one would expect these thresholds to be little changed with steel-strung instruments. These studies were based on listening tests using synthesized sounds created from calibrated measurements on real instruments and strings. Related studies have been made for violin sounds (Fritz *et al.*, 2007). For the guitar, the most acute listeners were able to discriminate a shift of low-frequency body mode frequencies of the order of 1–2%. Listeners were less sensitive to body damping: the Q-factors of body modes needed to be changed by around 20% to be audible.

Listening tests based on modal synthesis can provide valuable information on the relative acoustic discriminability of changes in different physical parameters of musical instruments. However, the absolute discriminability measured with such listening tests may not tell the whole story in relation to playing tests performed in a musical context. There are many complicating factors: for example interactions between the player and the instrument that are only present in playing tests (Fritz and Dubois, 2015), the effect of variability in players' performance, and differences in the test conditions.

Characterizations and comparisons of different woods based on players' perceptions and preferences abound in guitar magazines and online forums. However, these subjective characterizations and comparisons are rarely (if ever) done under blinded conditions. As a result they may be severely biased by a variety of factors including prejudice for or against certain wood species, visual aesthetics, and price. The impact that such biases can have has been illustrated by a series of studies on players' preferences among modern and old violins (Fritz *et al.*, 2014, 2012). In these studies, conducted under blinded conditions, expert violinists showed a slight but distinct tendency to prefer certain high-quality contemporary violins over old Italian violins made by Stradivari and Guarneri, despite the fact that the old instruments are generally regarded by experts as being tonally superior and are typically worth many times more than the best contemporary violins. To the best of our knowledge no published studies have ever compared, using a blind test methodology, players' preferences or discrimination abilities between steel-string acoustic guitars built with different back woods. An unpublished study on classical guitars built with various tropical and non-tropical woods failed to find differences in guitarists' preferences for one type of wood over the other in blind testing sessions (Walraet and Garston, 2015).

### D. Overview of the current study

For the current study, a very experienced luthier (who is the second author of this paper) built six steel-string acoustic guitars to the same dimensional and material specifications of all their parts, except for the back and side plates, which used different woods ranging widely in mechanical properties, popularity among players, and price. Different back wood species are normally worked differently according to their mechanical properties, and this standard practice was followed in the current study. The typical working thickness for back plates and sides ranges from around 2 to 2.8 mm for different timbers. Sides made from denser timbers, like those in the rosewood family, become very difficult to bend at 2.8 mm, to the point of impossibility on sharp curves. Conversely, less dense woods such as mahogany and walnut are intrinsically weaker, and would be far too fragile if worked to 2 mm thickness over the width of a back, typically 200 mm, or the depth of the sides, typically 100 mm. It should be also pointed out that these thicknesses are starting points: after the initial thicknessing of the sides, then bending, most wood will distort from the heat and moisture, and must be levelled off over a width of around 100 mm, so it is not uncommon to have wood only about 1 mm thick in some places. This is not a concern if it occurs near an edge, but requires reinforcement if it occurs in the middle of a side. For all these reasons it would not be desirable or even possible to work different timbers to identical thickness.

For the reasons discussed above, the back and side plates of the six guitars differed not only for the wood species used but also in their detailed thickness distribution. It is important to emphasize that this is a strength, not a weakness, of the study. To ask "do guitars sound different if different back woods are used with identical thickness?" would not be an interesting research question, because no professional instrument maker would ever do this. The aim is to ask whether other woods can be used as a satisfactory substitute for Brazilian rosewood and other traditional choices, by experienced luthiers who make the best use of each alternative wood. For brevity in this paper we will refer to the combined differences in back wood species and in their dimensional specifications as “back wood” differences.

Experienced guitar players were asked to rate the six guitars on a series of perceptual attributes, including an “overall sound” quality rating. A subset of players was tested twice to assess the consistency of their ratings. The players’ ability to tell apart two different guitars was also assessed for three of the guitars using an objective ABX testing methodology. Both the rating and ABX tests were carried out in a dimly lit room with the players wearing welder’s goggles that allowed them to see the position of their fingers on the fingerboard, but prevented them from recognizing the back wood.

The guitars were also tested to establish some physical acoustical parameters. The bridge admittance was measured, which allows body mode properties to be extracted. The results can be combined with the earlier threshold data to predict the discriminability of any given pair of guitars. Those predictions were compared with the results of the playing tests described above. As a further step, short passages of music were synthesized using the measured modal properties. Those synthesized sounds were used in another series of ABX listening trials to test directly for the influence of modal frequencies, while controlling for other factors such as variability in player’s performance.

## II. GUITAR BUILDING PROCESS

Six steel-string acoustic guitars were hand-built at the workshop of Fylde Guitars (Penrith, UK), on commission for the experiment. The guitars were all based on the “Falstaff” model, which has a relatively large body (see Table S1 in the supplementary material[1] for dimensional specifications). A design with a two-piece back was used for the current study.

The top plates of all the guitars were made of Sitka Spruce (*Picea sitchensis*) taken from a single flitch of timber. Four of the plates were consecutive cuts from the same log. It was not possible to obtain six consecutive cuts, but all the chosen plates were manually matched for stiffness and weight, and for similar grain spacing and pattern. The necks were cut from one plank, as were the internal braces. The fingerboards and bridges were Cameroon ebony (*Diospyros crassiflora*). The component parts were worked exactly in parallel so that, for instance, the thickness of each top plate and each brace was determined using the same settings and at the same time. The brace shaping and neck shaping were done by hand, but with significant effort to keep them the same. Any gluing operation was performed as much as possible in parallel, using the same pressure, the same gluing time, and the same ambient temperature. A particular effort was put into matching the neck angle and string height.

The woods for the back and side plates were chosen to be representative of timbers commonly used in acoustic guitar making and to cover a wide range of price and availability (see Table S2 in supplementary material), they were: Brazilian rosewood, Indian rosewood (*Dalbergia latifolia*), South American mahogany (*Swietenia macrophylla*), sapele (*Entandrophragma cylindricum*), North American maple (*Acer macrophyllum*), and North American claro walnut (*Juglans hindsii*). The specific pieces of timber were chosen as being representative of that species, taking weight and stiffness as the most important parameters. Physical appearance was not a factor in the selection process. As already mentioned, the back and side plates were not built to identical dimensional specifications. Instead, each guitar was made exactly as if it were a standard order for a customer. The thickness of the back and side plates were individually adjusted according to the properties and "feel" of each one.

Different timbers also require different finishing schedules, because of variations in grain, but considerable effort was made to keep the finish balanced. No filler was used. Each guitar was sprayed with polyurethane lacquer and rubbed back until the grain was "just" full and level. Equal spray coats were then applied, levelled and polished so that the coating thickness was equal on each instrument. Each guitar was set up to similar measurements and monitored for about six weeks, then readjusted before the first tests. The guitars were fitted with Elixir Nanoweb Light Acoustic 80/20 Bronze strings (W. L. Gore & Associates, Inc., Newark, DE), which were changed twice over the course of the experiments on all guitars at the same time, including just before the bridge admittance measurements.

## III. PHYSICAL MEASUREMENTS

### A. Bridge admittance and modal parameters

The input admittance at the bridge of each guitar was measured by applying a controlled force via a miniature impulse hammer (PCB 086D80) in the direction normal to the plane of the top plate, and measuring the velocity response with a laser-Doppler vibrometer (Polytec OFV-056). The measurement point was on the bridge saddle between the $5^{th}$ and $6^{th}$ string. The strings were damped, and the instruments were supported in a vertical position with soft foam at the end button and a foam-lined clamp on the neck. Further details have been described by Woodhouse and Langley (2012).

The results for each guitar are shown in Fig. 1A, alongside those of a different steel-string acoustic guitar, a Yamaha FG-403MS (Yamaha Corporation, Hamamatsu, Japan) that was not part of the experimental set of guitars. It is immediately clear that all guitars have three strong peaks at low frequencies, after which the behavior settles into a relatively smooth and lower-level trace at higher frequencies. The three strong peaks were used in listening tests reported in Section VII, and their modal frequencies and damping factors are listed in Table I. The interpretation of these three modes is as follows: the first two are part of the coupled triad discussed earlier, while the third peak is a transverse bending mode of the top plate. The first two are significantly influenced by the back, while the third one should have only a very weak influence from the back: its frequency is determined mainly by the transverse stiffness of the top plate, including the effect of the wood, thickness, bracing and bridge (Christensen and Vistisen, 1980; Fletcher and Rossing, 1998; Richardson, 1994, 2002).

Figure 1B shows resynthesized bridge admittances corresponding to the subset of the guitars used in the tests to be described in Section VII. The Yamaha guitar has been used as a reference case, because a detailed modal fit up to 1200 Hz was available for this guitar. Modified versions were then computed in which the low-frequency modal properties were replaced with the measured values from the set of guitars, given in Table I. The resulting pattern is clear in the figure: in each case the low-frequency behavior is a close match to the measured response of the chosen guitar, while the response at higher frequencies is essentially identical in all cases (and includes no modes above 1200 Hz). It would, of course, be virtually impossible to achieve this kind of controlled variation of modal properties with physical instruments.

### B. Wood properties

It is useful to have some idea of the physical properties of the different woods used in this study. For logistical reasons it was not possible to test the particular specimens used to build the six guitars, but two examples of each wood type were selected for testing from the wood store of the same luthier. Back-plate blanks were thicknessed and shaped to rectangular form, then tested by the procedure described in McIntyre and Woodhouse (1988). Since the wood was intended for two-piece backs, the two halves from each set could be independently tested to give a direct measure of the consistency of results for samples of wood that are as similar as possible.

The measurements yield density, plus three stiffnesses called $D_1$, $D_3$ and $D_4$ by McIntyre and Woodhouse (1988). $D_1$ and $D_3$ quantify the long-grain stiffness and cross-grain stiffness, while $D_4$ quantifies the twisting stiffness. A fourth stiffness, $D_2$, relates to Poisson's ratio: it is more difficult to measure, but it has little influence on the vibration behavior of the plates and so it is ignored here. It is a deliberate choice to present results for the unfamiliar $D_j$ rather than the more familiar Young's modulus, shear modulus and so on. These more familiar moduli can only properly be regarded as properties of the solid wood, whereas the $D_j$ are properties of the actual plate as cut from the solid wood. The values are sensitively influenced by the way the wooden plates have been cut from the log as well as by the underlying wood properties: see McIntyre and Woodhouse (1988) for details.

The results are plotted in Figure 2. In all cases, the pairs of symbols indicating the two halves of a given set of wood fall reassuringly close together. The density results show that the four non-rosewood timbers have rather similar densities: the between-species variation is no bigger than the within-species variation shown by the pairs of tested sets. For the stiffnesses, Brazilian rosewood generally shows the highest values, while walnut shows the lowest. In most cases the results for the two tested sets of each wood fall close together. The conspicuous exceptions come from the values of $D_3$ for mahogany and sapele: in both cases one set shows much lower values of this cross-grain stiffness. The difference is mainly attributable to the fact that the low values are associated with samples that have been cut a long way off the quarter.

### C. Discussion

The measured bridge admittances and deduced modal properties show subtle differences between the guitars. However, careful examination of the results suggests that the differences seen were *not* caused primarily by the choice of back wood. Despite best efforts to vary only the back wood and keep other variables constant, the subtle variations observed between the guitars are probably due to small differences in the top plates (with bridge, bracing and lacquer) rather than to the deliberate differences in back woods. The evidence for this suggestion is that the frequency $F_3$ — which should essentially be independent of the back

— in Table I shows variations at least as big as those in the other two low modes, and furthermore the three frequencies show strong correlation.

The guitars made with Indian rosewood and sapele show the most extreme differences: frequencies that are systematically sharp and flat, respectively, by about half a semitone compared to the mean of the set, so that relative to one another they show about a semitone difference. But these are not the woods that stand out as extreme on the basis of the measured wood properties shown in Fig. 2: Brazilian rosewood is both the densest and the stiffest, while walnut is probably at the other extreme.

These observations have two important consequences. First, the similarity of the responses despite the very different properties of the back woods confirms a possibility mentioned earlier. The guitar maker, by treating each back plate in the way that his experience suggested was best, has to a very large extent compensated for any physical differences between the types of wood. The biggest residual difference in acoustical behavior is between guitars built with two woods that were not judged to be the most extremely different from the maker's perspective (sapele and Indian rosewood). The guitar made with Brazilian rosewood, traditionally the favoured back wood, does not stand out in these results. Its acoustical response and modal properties fall in the middle of the range on all measures, specifically appearing to be rather close to the sapele instrument.

The second consequence is a prediction about perceptual discrimination between the guitars in the set. The thresholds found in earlier work (Woodhouse *et al.*, 2012) suggest that the sapele and Indian rosewood guitars should be sufficiently different in their low modal frequencies that a skilled listener would be able to tell them apart moderately reliably. But these are the extremes of the range, and all other pairs of guitars would be predicted to be hard to discriminate.

However, one should be cautious about generalizing predictions derived from absolute thresholds obtained in listening tests with synthesized sounds to real-life playing tests. On the one hand musical perception is subtle and context-dependent. It is possible that experienced musicians playing real instruments might be more discriminating than is suggested by these predictions, based as they were on very artificial listening conditions with headphones and synthesized sound fragments. On the other hand, it is possible that the constraints of real-life playing tests, such as variability in player's performance and delays between listening of two guitars due to the need to physically swap instruments, may make discrimination more challenging in these tests. The remainder of this paper presents the results of a variety of psychoacoustic tests using the guitars of the test set, to explore these possibilities.

## IV. PSYCHOACOUSTICS EXPERIMENTS METHODOLOGY

### A. Participants

Fifty-three guitar players took part in one or more experiments in the study. The participants were not screened for hearing loss, but two participants reported a diagnosis of hearing loss. One of these two participants reported a moderate hearing loss for both ears and his data were excluded from all analyses. The other participant was profoundly deaf in one ear since childhood; because his self-reported hearing was good for the other ear his data were not excluded from the analyses. The remaining 52 participants (51 males) had a mean age of 42 years (SD=16). They were classified according to their employment status as professional (if they worked full time as guitar players, e.g. artist, guitar instructor, etc.), semi-pro (if they worked part-time as guitar players), or amateur. On average, the players reported having 26 years (SD=15) of experience playing the guitar (30 years for the professionals, 25 years for the semi-pros, and 25 years for the amateurs). Fifteen participants performed the blinded sound rating test at the Ullapool (Scotland, UK) guitar festival in a large room normally used as a dance studio. All the other participants were tested in a medium-sized furnished room at Lancaster University. Participants were compensated for their time at a rate of £10 per hour.

### B. Blinded tests procedures

Participants were tested while wearing welder's goggles in a dimly-lit room. These test conditions allowed them to see their fingers and the guitar neck while playing, but prevented them from visually identifying the wood of the guitar. To further limit the possibility that participants could recognize the guitars by non-acoustic cues, they were asked to close their eyes both while being handed each guitar by the experimenter, and when the experimenter picked up each guitar from them; they were also asked not to look at the body of the guitars during the test, and not to tap on the guitar body or inspect it in any other way. The lacquer on the guitars substantially eliminated possible odor cues that could make the different woods recognizable. Nonetheless, an air freshener placed on a table close to the player was used to mask any potential residual odor cues.

### C. Statistical analyses

Statistical analyses were performed using Bayesian general linear models (GLM) implemented with Markov Chain Monte Carlo (MCMC) methods (Kruschke, 2014). Details of the models are provided in the supplementary material. Because MCMC methods are stochastic their results can vary slightly on different runs. For this reason, as well as to avoid rejecting a given null hypothesis on the basis of a trivial effect size, we only rejected a null hypothesis for differences whose 95% credibility intervals (CIs) fell outside a region of practical equivalence (ROPE) around the null value (Kruschke, 2014). We will refer to such differences as *credible* differences. For all the measures used in this study the ROPE was set at ±0.2 because a magnitude of 0.2 represents a very small difference for the rating measures (5% of the scale range), for standardized measures (test-retest correlations and factor scores) and for $d'$ values. These ROPEs could be considered arbitrary, and sometimes the ROPE is not stated explicitly but left to the reader to decide. In practice, the vast majority of the conclusions in our analyses would not be affected by the choice of a different ROPE because most CIs crossed the null value, and therefore could not exceed the criterion for the rejection of the null hypothesis under any choice of ROPE.

## V. PSYCHOACOUSTICS EXPERIMENT 1: GUITAR RATINGS

### A. Method

Fifty-two guitar players (18 professional, 21 semi-pro, 13 amateur) performed the blinded rating experiment once. A subset of these players (eight professional, 17 semi-pro, nine amateur) performed it twice, so that the consistency of their ratings could be assessed. Guitarists played each guitar and were then asked to evaluate it on a 1-to-5 scale for "overall sound", playability, and 14 additional perceptual attributes listed in Table II. These additional perceptual attributes were chosen on the basis of a corpus analysis of online guitar reviews (described in the supplementary material), as well as consultation with the luthier. Some of the resulting rating attributes are likely to be semantically overlapping and thus redundant. However, it is not clear how players use these terms to describe the sound of a guitar. Instead of making a priori assumptions on how the terms relate to each other for guitarists, and arbitrarily select one among several possibly related terms to reduce the dimensionality of the dataset, a relatively large number of rating attributes was used at the data collection stage. The dimensionality of the dataset was reduced at the data analysis stage by means of an exploratory factor analysis (EFA), which is described in Section V D.

Players were informed that the rating attributes were selected on the basis of a linguistic analysis of guitar reviews and that some attributes may semantically overlap with each other. Although the number of attributes that players had to rate was somewhat higher than in previous studies of players evaluations of musical instruments (Fritz *et al.*, 2014; Saitis *et al.*, 2012), the fact that several attributes were probably semantically related should have considerably reduced the cognitive load needed to perform the task. The order of presentation of the guitars was randomly assigned by a computer algorithm for each player. The experimenter passed each guitar to the player by positioning it directly on the legs of the player. Players were allowed two minutes to play freely with each guitar any tune of their choice. They were allowed to use either a fingerpicking technique, or a pick, and to use a non-standard tuning if they wished. Once the two minutes of free play with a given guitar had elapsed the participants were asked to start the rating phase. They were allowed to keep the guitar during the rating phase and play it again as they wished in order to accurately rate the guitar on each attribute. There was no time limit for this rating phase; most participants completed it in about four to five minutes. For the ratings participants were presented with the 16 questions listed in Table II on the computer monitor, and were asked to answer each question with a 1-to-5 rating by dragging a slider through a mouse. The questions were presented in the same order shown in the table. Non-integer ratings up to one decimal place were allowed (e.g. 3.2). For participants who scored the guitars twice the whole procedure was repeated after a short (5–10 min) break. The total duration of the session for participants scoring each guitar twice was 90 minutes.

### B. Results: rating differences

The average "overall sound" ratings given to each guitar are shown in Fig. 3. To estimate the rating differences between guitars in the general population of guitar players, the ratings were modeled by a Bayesian GLM as a function of guitar back wood, player professional status, and individual player (see supplementary material for details). Figure 4 shows the estimates of the rating differences given by the model (filled symbols) alongside their 95% CIs, denoted by the horizontal segments. The estimated differences between guitars were all small ($< 0.2$). All of the 95% CIs crossed a difference of zero, indicating that none of the differences was credibly different from zero. Moreover, the CIs were relatively narrow; most of them did not exceed a difference of ±0.3 (7.5% of the scale range), suggesting that even if rating differences between guitars existed in the general population of guitar players, they would be at best small. A qualitatively similar pattern of results was observed for each player category (see Fig. S1A in supplementary material).

Because for practical reasons only one guitar of each back wood species could be tested in our experiment, specific comparisons between the individual guitars in our sample are of limited interest: The results strictly apply only to the exemplars tested and because of potential nuisance factors such as within-species variability in the quality of wood, or residual manufacturing variability, may not readily generalize to other guitars of the same species. However, the six guitars tested can be seen as a sample of a population of guitars (from the same maker) nominally differing only in their back woods. The variance in "overall sound" ratings between guitars of this "superpopulation" (Gelman and Hill, 2007) was estimated by our Bayesian model to be low. In standard deviation units its posterior mode was 0.06 with a 95% CI of 0.001 – 0.249. In other words, our results indicate that typical differences in "overall sound" ratings between guitars nominally differing only in their back woods would be in the range of 0.001 to 0.249 (between 0.02% and 6.2% of the scale range). If back woods were a major determinant of a guitar sound, it is very unlikely that in a sample of six guitars taken from that population, all six would be given very similar ratings.

Figure 3 shows the average playability ratings given to each guitar. Figure 4 shows the estimates of the rating differences given by the model (filled symbols), and their 95% CIs (horizontal segments). As for the "overall sound" ratings, the estimated differences in playability ratings were small. None of them was credibly different from zero, and the uncertainty of the estimates provided by the CIs suggests that even if rating differences between guitars existed in the general population of guitar players, they would be at best small. Qualitatively, a similar pattern of results was observed for each player category (see Fig. S1B in supplementary material). The superpopulation standard deviation of playability ratings between guitars was estimated to be 0.05 (95% CI: 0.001 – 0.195).

### C. Results: test-retest correlations

Figure S2 in the supplementary material shows the ratings given in each session by players who performed the blinded rating experiment twice. We used a Bayesian model to estimate the test-retest correlation for each individual player, as well as at the group level (see supplementary material for details). At the group level this correlation was 0.11, with a 95% CI ranging from -0.04 to 0.26. Hence, the group-level estimate of the test-retest correlation was not credibly different from zero, and taking into account the uncertainty of the estimate provided by the CI, it could be at most around 0.26, which would still indicate poor test-retest consistency of the "overall sound" ratings. The group-level test-retest correlation for the playability ratings estimated by the model was also low, 0.08, with a 95% CI ranging from -0.05 to 0.25.

### D. Results: exploratory factor analysis of rating attributes

The ratings given by guitarists on the additional 14 perceptual attributes were entered into an EFA. The purpose of this EFA was to reduce the dimensionality of the dataset, and to define some of the basic perceptual dimensions that are used by guitar players to judge the sound of a guitar. The EFA revealed three orthogonal factors that gravitate respectively around the dimensions of loudness, warmth, and clarity (see Fig. S3 and results section in supplementary material). In order to understand how the ratings on these perceptual dimensions relate to "overall sound" ratings, we extracted factor scores using the Anderson-Rubin algorithm (DiStefano *et al.*, 2009). A Bayesian multiple regression model (see supplementary material for details), predicting the "overall sound" ratings from the extracted factor scores, indicated that all three factors are credible predictors of "overall sound" ratings. The estimated slope coefficients, which represent the "overall sound" rating change for a unit change in factor score were 0.31 (CI: 0.26 – 0.37) for factor "loudness", 0.35 (CI: 0.29 – 0.41) for factor "warmth", and 0.42 (CI: 0.37 – 0.47) for factor "clarity". None of the slopes was credibly different from each other (see supplementary material for details).

The extracted factor scores were also used to test statistically whether the guitars differed from each other on these underlying dimensions. This analysis was performed using the same model used for analyzing "overall sound" and playability differences between guitars. The results did not show any credible differences between any pair of guitars for any of the factors (see Fig. S4 in supplementary material). However, there were trends for small differences between some guitar pairs in factor "loudness" and factor "warmth", and given the relatively large width of the CIs it is not possible to exclude the possibility that small or modest differences between guitars in these factors exist but could not be detected by our measurements.

# VI. PSYCHOACOUSTICS EXPERIMENT 2: ABX TEST

## A. Rationale

While overall sound quality ratings can provide information on the preferences of guitarists for different woods, they cannot determine whether guitarists can perceive acoustic differences between woods. It is possible that guitarists may be able to clearly hear or feel differences between back woods, but have no preference for one over the other. The ABX test is a simple discrimination test that is widely used in the audio engineering community to check whether there are audible differences between sounds (Boley and Lester, 2009). The simplicity of this test derives from the fact that it only requires the observer to respond on the basis of the perceived similarity between stimuli without the need to have a defined verbal/semantic characterization of the dimension(s) along which the stimuli differ (Hautus and Meng, 2002). In experiment 2 we used an ABX test to assess the ability of guitar players to discriminate between pairs of guitars by acoustic cues alone.

## B. Method

Thirty-one of the guitarists (seven professional, 14 semi-pro, 10 amateur) who performed the blinded acoustic rating test also took part in an ABX discrimination test. On each trial of this test guitarists played under blinded conditions first one guitar (guitar $A$) for one minute, then another guitar (guitar $B$) for another minute. They were then given again one of the two guitars (guitar $X$), and were asked to decide if it was guitar $A$, or guitar $B$. Because of time limitations only three guitars (Brazilian rosewood, sapele, and walnut) were used in this test. These three woods were chosen among the set of six woods to be representative of different levels of sustainability and price (Table S2), and based on the luthier's expectations of which woods would be more and less similar. Note from Fig. 2 that this selection does indeed include the woods with most disparate physical properties (Brazilian rosewood and walnut).

On each trial two of these three guitars were used. The participant was first given one guitar (guitar $A$) to play, and was then given the other guitar (guitar $B$). The participant could request to swap between guitar $A$ and $B$ as many times as s/he wished within a maximum time period of two minutes (most participants played guitar $A$ for one minute, and then guitar $B$ for another minute). At the end of the two-minutes period the participant was then given again one of the two guitars (guitar $X$) to play for a minute and was asked to report whether this guitar was guitar $A$ or guitar $B$. The three guitars were all tuned to the same nominal pitch prior to the beginning of a session. However, participants could potentially hear residual tuning differences and use them to discriminate between two guitars. In order to avoid this issue, guitar $X$ was quickly de-tuned and re-tuned to the same pitch with a digital tuner (Korg AW-2G Clip-On Tuner) by the experimenter before it was handed to the participant. This procedure lasted about 30 seconds. To minimize any possible distraction due to the re-tuning procedure, during this procedure the participant listened to pink noise played at a comfortable level through headphones.

The ABX test was completed in two 1-hour sessions that were run on different days. In each session, the players completed four trials for each guitar pair, one for each of the possible ABX stimulus sequences (correct response $A$: $<S_1 S_2 S_1>$, $<S_2 S_1 S_2>$; correct response $B$:

$<S_1S_2S_2>$, $<S_2S_1S_1>$). One of the players was unable to attend the second ABX session, therefore only four trials per guitar pair are available for this player. A roving design was adopted: on each trial both the guitar pair and the stimulus sequence were randomly sampled without replacement from the set of the 12 possible combinations of guitar pairs and stimulus sequences by a computer algorithm. No immediate feedback was given to the participants upon completion of a trial. However, participants were informed of the total number of correct responses at the end of each session. The $d'$ values were computed from hit rates (proportion of times guitar $X$ was correctly identified as guitar $A$) and false alarm rates (proportion of times guitar $X$ was incorrectly identified as guitar $A$) assuming that guitarists used a differencing strategy to perform the ABX task (Macmillan and Creelman, 2004). This assumption is motivated by the fact that the experiment had a roving design in which the guitar pair for each trial was randomly chosen among the three possible guitar pairs. To avoid undefined values of $d'$, hit and false alarm rates of 0 and 1 were converted to $1/(2N)$ and $1 - 1/(2N)$, respectively (Macmillan and Creelman, 2004).

### C. Results

Figure 5 shows the average performance in the task for each guitar pair as a function of player professional status. Although there was some variability between player categories, it is evident that $d'$ values were very low overall (a $d'$ value lower than 1 indicates poor discriminability). These $d'$ values were modeled as a function of guitar pair, player professional status, and individual player, using a Bayesian GLM (see supplementary material for details). The $d'$ CIs for each guitar pair as a function of player professional status are shown in

Fig. 6. When considering the scores across all guitar pairs and player categories the results show that, although performance was credibly above chance, it was poor, with the 95% CI for $d'$ falling entirely below 1 (CI: 0.21 – 0.8). Discriminability was not credibly different between the three guitar pairs tested (see Fig. S5 in supplementary material). There were trends for professional and amateur players to perform better than semi-pro players, but these differences were not credibly different from zero (see Fig. S6 in supplementary material). It should also be noted that these trends are not entirely in the expected direction, as one would not expect amateur players to perform better than semi-pro players. However, this odd trend is consistent with random sampling variability, rather than with a real performance advantage of amateur over semi-pro players. Interactions between player category and guitar pair were not credibly different from zero.

## VII. PSYCHOACOUSTICS EXPERIMENT 3: ABX TEST WITH SYNTHESIZED GUITARS

### A. Rationale

In the ABX test described in Section VI it was difficult to implement a control condition to check that players would perform it well when comparing guitars differing not only in their back woods but also in other component materials and dimensional specifications. A guitar of a different size, for example, could have been discriminated just by touch even without being played. For this reason, a subset of guitar players performed an additional ABX task in which the stimuli were sounds synthesized on the basis of the bridge admittance measurements of the guitars described in Section III. This allowed us to include a condition comparing one of the six guitars built for the study with a guitar of a different make and model that should have been easy to discriminate. Additionally, the results of this test could be more directly compared with previously published threshold data for changes to the modal properties of the instrument body obtained with synthesized guitar sounds (Woodhouse *et al.*, 2012). We used the same synthesis model of Woodhouse *et al.* (2012), which allows an accurate representation of the coupled mechanical vibration of the strings and guitar body.

The musical excerpt chosen as stimulus consisted of the first two measures of the song "Tears in Heaven". This excerpt was chosen as being typical of fingerpicking acoustic guitar style and for covering a relatively wide register of notes. The same three guitars that were tested in the previous ABX test were again tested. In addition, discrimination performance was assessed between the Indian rosewood and the sapele guitars, because these two guitars have, among the set of the six guitars built for the study, the most disparate modal frequencies. Based on the earlier threshold tests (Woodhouse *et al.*, 2012), this pair would be predicted to be the most readily discriminable. Finally, discrimination performance between one of the six guitars built for the study (Indian rosewood) and a guitar of a completely different make and model (Yamaha FG-403MS) was also assessed to check that when the guitars differed not only in their back woods, but also in other characteristics, the ABX test could be performed proficiently by the players.

### B. Method

Seven guitarists (three professional, four semi-pro), who had previously taken part in both experiment 1 and experiment 2 took part in this test. The stimuli for this test were synthesized from measurements of the bridge admittance functions of four of the Fylde guitars employed in the other tests (Brazilian rosewood, Indian rosewood, sapele, and walnut), and a guitar of a different make and model (Yamaha FG-403MS steel string guitar). Each stimulus had a duration of five seconds. Data collection occurred over the course of two sessions. During the first session participants completed first a practice block of two trials per guitar pair, and then a block of 36 trials per guitar pair. These blocks were randomly ordered. During the second session participants completed an additional block of 36 trials per guitar pair. The presentation intervals were marked by flashing lights on a computer screen, and were separated by 750-ms silent intervals. Participants indicated their responses via key presses mapped to the two response alternatives. Immediate feedback after each trial (a green or red flashing light for correct or incorrect responses respectively) was given during the practice blocks but not during the main blocks (a white flashing light simply acknowledged that a response had been given). The stimuli were generated using a 22,050 Hz sampling rate and 16-bit depth. They were sent to a digital-to-analog converter (E-MU 0204 USB), and played diotically through Sennheiser HD650 headphones. Participants were tested in a IAC double-walled sound-insulating booth. Because a non-roving, blocked design was used for this experiment, $d'$ values were calculated assuming that participants used an independent observations strategy to perform the ABX task (Macmillan and Creelman, 2004).

### C. Results

The average $d'$ values obtained by the seven guitarists who took part in this test are shown in Fig. 7. The $d'$ values were modeled as a function of guitar pair, and individual player, using a Bayesian GLM (see supplementary material for details). The $d'$ CIs for each guitar pair are shown in Fig. 8. Performance was very good for the discrimination of the Indian rosewood and Yamaha guitars (posterior $d'$ mode=3.24, CI: 2.74 – 3.78), indicating that players were able to do the test proficiently when the guitars were of a different make and model.

Performance in the discrimination of the three guitars that had been tested in the ABX test with real guitars was slightly better than in the previous test, but was nonetheless poor, with $d'$ values around 1 (Braz. rosewood vs sapele $d'$ CI: 0.22 – 1.23; Braz. rosewood vs walnut CI: 0.22 – 1.24; sapele vs walnut $d'$ CI: 0.44 – 1.47). Discrimination performance for the guitar pair with the most divergent modal frequencies was better, but still mediocre, with a $d'$ around 2 (Indian rosewood vs sapele $d'$ CI: 1.54 – 2.51). Overall these results show that although performance in this test with virtual guitars was better than in the playing ABX test, discrimination performance among most of the guitar pairs nominally differing only in their back woods was generally poor. It was still far from perfect even for the guitar pair with the most divergent modal frequencies, but the pattern of the results was generally consistent with the predictions based on earlier threshold tests (Woodhouse, 2002).

## VIII. GENERAL DISCUSSION

The results of this study indicate that the choice of back wood has a minimal effect on the quality of acoustic steel-string guitars, provided that the guitar maker uses each different wood in a suitable way. This conclusion is consistent with earlier expectations from acoustic and psychoacoustic studies of the guitar, and it is supported by converging evidence from our experiments:

1. differences in the body mode properties of the set of six guitars were small, and furthermore the residual differences did not seem to stem mainly from the back wood;
2. players's average ratings for "overall sound" and for playability were similar between the six guitars;
3. performance in a simple ABX discrimination test run under realistic playing conditions was very poor, only slightly better than chance level;
4. performance in an ABX discrimination test of guitar sounds synthesized from the bridge admittance functions of the six guitars was poor for most guitar pairs tested. It was mediocre for the guitar pair with the largest differences in bridge admittance functions; differences that likely reflect residual top plate variability.

### A. Rating tests

Experienced guitar players gave, on average, very similar "overall sound" and playability ratings to six guitars with backs and sides made of different woods. Guitarists who rated the guitars twice were also not consistent in their ratings from one session to another. There are several possible explanations for this pattern of results. One possibility is that guitarists did have preferences for different back woods but were somehow unable to express them consistently using the rating scale, perhaps because they were unable to keep track of the ratings they were giving to each guitar. This explanation seems implausible in light of the fact that guitarists did show clear preferences between some instruments in a rating test of the visual appearance of the guitars (described in the supplementary material), which had essentially the same format as the acoustic rating test. The only study that used a rating procedure similar to that employed in the current study for the evaluation of guitar sounds failed to find differences between guitar recordings of three guitars of varying commercial value (Inta, 2007). However, only five participants were tested in that study, and they were asked to listen to several recordings of each guitar, and each time rate it on 34 attributes, which may have substantially increased listener fatigue during the test. Given these issues it is difficult to draw conclusions on the validity of the rating method used in our study from the study of Inta (2007). It should be noted that rating methods similar to the one used in the current study have been often used in studies of consonance perception, in which, for example listeners may be asked to rate the pleasantness of various musical intervals on a scale from -3 to 3 (Bones *et al.*, 2014; Bones and Plack, 2015a,b; McDermott *et al.*, 2010). These studies of consonance perception reliably find differences in the pleasantness ratings given by both musicians and non-musicians to musical intervals according to their degree of consonance.

Another potential explanation for the pattern of results obtained in the blinded sound rating test is that back wood preferences are highly subjective: individual players might have clear favorites among woods, but the highly rated woods would be different from player to player. When the ratings are averaged across all players there would be, overall, no ratings differences between guitars. However, this explanation seems implausible in the light of the fact that guitarists that rated the guitars twice were not consistent in their ratings from one session to another, indicating that individual players did not have strong preferences for one or more woods. A parsimonious explanation for both the poor test-retest rating correlations, and the similarity of the average ratings across guitars, is that the perceived differences between the six guitars were so small that players could not rank them consistently either within or across sessions. This explanation is corroborated by the results of the ABX discrimination test, which show that players' ability to discriminate between three of the six guitars was poor, just above chance level.

Our finding of poor test-retest consistency of player ratings seems superficially at odds with previous studies on violin players showing that musicians can be self-consistent in their ratings of musical instruments (Saitis *et al.*, 2015). Unlike the guitars of our study, however, the instruments used in these previous studies were of different makes and models, and were intentionally selected to sound different from each other. Given that simply discriminating between the different guitars in the current study was very hard, it is not surprising that players failed to be consistent in their ratings. Additionally, a violinist has access to significantly more information about the instrument than a guitarist does, because of the sensitive way that the "playability" of a bowed string depends on details of the body vibration. Different bowing gestures elicit different transient behavior, whereas in a plucked instrument like the guitar the player has very little influence over the form and length of initial transients (Woodhouse, 2014).

In an additional experiment described in the supplementary material (unblinded sound ratings), guitarists did not give credibly different ratings for overall sound or playability to the guitars even when they could see the guitars in full light. This suggests that biases for guitars with prized tropical woods may not be as strong or widespread as surveying the specialized guitar press suggests they are. However, it should be kept in mind that guitarists performed the unblinded rating test only after the ABX test, on which they were given feedback at the end of each session. Because most guitarists were barely able to discriminate the guitars in the ABX test, and were aware of this fact, it is possible that they were more cautious in their ratings in the unblinded rating test, reducing any pre-existing bias for prized tropical woods.

### B. Dimensions used by players to evaluate guitars

The results of the EFA on the acoustic ratings of the guitars indicate that a number of perceptual attributes commonly used by guitarists to describe the sound of a guitar are closely related to each other. These attributes revolve around the dimensions of loudness, warmth, and clarity. Our results suggest that these are three of the basic underlying dimensions used by players to evaluate the sound of a steel-string acoustic guitar, and each of these dimensions is positively related to the "overall sound" rating. The number of factors that can be extracted using an EFA depends on the number and the diversity of descriptors that are used in the rating questionnaire, both of which were limited by time constraints in our experiment. Therefore, we do not claim that loudness, warmth, and clarity are the only dimensions on which the sound of a steel-string acoustic guitar is judged. It should also be kept in mind that a positive association between these dimensions and overall sound rating does not imply a causal relation between them.

### C. ABX discrimination

The results of the ABX tests indicate that discriminating between two guitars nominally differing only for their back wood is very difficult. Although performance was on average slightly above chance level, it was poor. In the test with synthesized guitar sounds performance was only slightly better. Despite the fact that listening conditions were probably more favorable in this test, because of the absence of any variability in the rendition of the excerpts due to variability in player's performance and the absence of long delays between excerpts due to the need to physically swap instruments, performance was still poor for most guitar pairs tested, and was only mediocre for the guitar pair with the most divergent body mode frequencies (Indian rosewood vs sapele). The results of the modal analysis strongly point to residual differences in the top plates as the main cause of differences in body mode frequencies between the guitars. Because performance was poor between guitars with minimal differences in body mode frequencies, and improved only for the guitar pair with the most divergent body mode frequencies, it seems reasonable to attribute this improvement to these differences in modal body frequencies, and hence to residual top-plate differences. Importantly, the ABX test with synthesized guitar sounds explicitly demonstrated excellent discrimination performance between the Indian rosewood guitar and a guitar of a different make and model. Listeners could easily and reliably perform the ABX task when the differences between guitars were sufficiently large, and not limited to their back wood.

A previous study using synthesized guitar sounds established thresholds for detecting a shift of low-frequency body modes. The average performance of the subjects in that test gave a threshold around 3%, while the most acute listeners achieved values around 1% (Woodhouse *et al.*, 2012). Thresholds in that study were measured using an adaptive procedure which tracked a $d'$ of 1.6. For comparison, the $d'$ measured in our study for the discrimination of the sapele vs walnut guitars, which have a 3% difference in their low frequency peaks, was around 1. This figure is thus somewhat lower than would be expected on the basis of the results of the previous study. Besides sampling error, this difference may be due to methodological differences between the two studies, such as the use of different stimuli, and the fact that the “natural” modal shifts between the guitars in our study do not precisely match any of the conditions tested by Woodhouse *et al.* (2012), which consisted of either coherent shifts of all low body mode frequencies, or of a shift of only the 200 Hz mode.

Some musicians are said to have “golden ears” because of their ability to discriminate subtle differences between sounds that other musicians cannot. Large interindividual differences in the ability to discriminate changes in basic acoustic attributes such as pitch are common in psychoacoustic studies, although interindividual variability tends to be smaller for trained musicians compared to non-musicians (e.g. Micheyl *et al.*, 2006). One might wonder whether, although on average discrimination performance in the ABX tasks between guitars nominally differing only for their back wood was rather poor, some players were consistently good at this task. For the playing ABX test, with only eight trials per condition, the variance of $d'$ is too large to make meaningful statements on individual players' performance (Hautus and Lee, 1998; Macmillan and Creelman, 2004). Keeping this caveat in mind, the best performing player, L7, had an average $d'$ across conditions of 2.1 (19/24 correct answers), a medium performance level. Individual results for the ABX test with virtual guitars, in which listeners completed 72 trials per condition, are shown in Fig. S7. One listener, L18, tended to perform better than the others across conditions, achieving a relatively high $d'$ of $\sim 3$ in the Indian rosewood vs sapele condition and in the sapele vs walnut condition. This listener could be said to be discriminating quite well some of the virtual guitar pairs nominally differing only for their back wood. However, given that the results of the modal analysis strongly suggest that modal differences did not stem mainly from the back wood, it is likely that his ability to discriminate between these guitars was based on modal cues caused by residual differences in the top plates, rather than by the different back plates.

### D. Generalizability of results

There are many factors that could conceivably limit the generalizability of our results, including the limited number of wood species that was tested in our study, the within-species variability of mechanical wood properties, the various possible choices in the guitar design including its shape, size, and bracing patterns, and the testing conditions. Each of these factors will be discussed below.

Only a limited number of wood species could be tested in our study. However, the woods selected included some of the most popular woods used for guitar backs, such as Indian rosewood and mahogany, as well as Brazilian rosewood, which is often considered the “gold standard” among back woods. These woods were compared to some less well regarded wood choices such as sapele and walnut. The lack of significant differences in “overall sound” ratings between the woods tested in our study may not extend to other untested wood species, but this does not invalidate the main conclusion of our study that, under blinded conditions, highly prized and expensive woods such as Brazilian rosewood are not necessarily preferred by guitar players to lesser-known, much less expensive woods. Our study investigated only solid woods, and therefore does not address the question of whether other materials like laminates or carbon fiber composites, already introduced in the market by several guitar companies, can substitute solid woods without loss of sound quality. This is an important question that we hope will be addressed by future studies.

Wood is a biological tissue, and as such it inevitably shows some variability in mechanical properties within the same species. Furthermore, as illustrated in Fig. 2, the properties of the final back plate can be sensitively influenced by how the board is cut from the log. However, the wood for each guitar used in our study was carefully selected by the luthier to be typical of each species according to his decades-long experience in guitar making. For this reason, we believe that the acoustic characteristics of the guitars were typical of guitars with backs made of those woods. Residual variability of the top plate woods (even though they were all cut from the same flitch), and variability due to small details of fabrication (French, 2008), could have affected the results of individual comparisons between our guitars to some extent. However, it is very unlikely that variability due to all of these factors could have conspired to make all six guitars have very similar ratings for overall sound quality and playability. This argument is supported by the fact that the playability and overall sound ratings variance estimates of the superpopulation of guitars nominally differing only for their back wood was low, indicating that back wood differences have only a minor impact on overall sound and playability ratings.

The design of a steel-string acoustic guitar involves many choices, including its shape, size, and bracing pattern. Only a single design was tested in our study, so it is not guaranteed that our results would generalize to guitars built with different designs. It should be pointed out, however, that the design chosen for our guitars is one of the most commonly used for steel-string acoustic guitars, so our results should generalize to a large number of guitars that are built with similar designs. Furthermore, the results shown here have confirmed predictions from earlier acoustical modelling: the influence of the back plate on the low-frequency modes is limited and well-understood (Christensen and Vistisen, 1980; Fletcher and Rossing, 1998; Richardson, 1994, 2002), and it should be possible for the guitar maker to compensate for material variations to a very large extent.

The time that each guitarist could spend playing and listening to each guitar in our experiment was limited, in the order of minutes. It is possible that if guitarists could have played the guitars for longer periods of time, in the order of hours, days, or perhaps years, their ability to discriminate between the guitars, as well as the consistency of their ratings across sessions, could have been higher. It is well known that the ability to discriminate sounds on the basis of simple perceptual attributes such as pitch or loudness (Carcagno and Plack,

2011) can improve dramatically with practice, a phenomenon known as auditory perceptual learning. This phenomenon, however, is much less pronounced for musicians (Micheyl *et al.*, 2006), who are already experienced listeners owing to years of musical practice. Despite this, we cannot exclude the possibility that longer practice with the instruments could have improved the results of the discrimination test, and possibly led to a different ranking of the guitars in terms of their overall sound and playability. However, the very idea that even experienced guitar players may need an extended period of perceptual learning to appreciate differences between back woods would indicate that such differences are very small. Large differences in the sound of two guitars should be immediately apparent without training, especially for experienced guitar players. The constraints on the time each guitarist could play the guitars were imposed mainly for practical reasons (longer or additional testing sessions would have been required otherwise). However, these time constraints are also relevant because they mimic to a certain extent what would typically happen in a guitar shop, where a player may spend minutes (rather than hours or days) playing and evaluating a small number of guitars before making a decision on which one to buy.

### E. Conclusions

The results of our study indicate that steel-string acoustic guitars with backs and sides built using traditionally prized, expensive, and rare woods are not rated substantially higher by guitarists than guitars with backs and sides built using cheaper and more readily available woods. The poor ability of guitarists to discriminate under blinded conditions between guitars with backs and sides made of different woods suggest that back wood has only a marginal impact on the sound of an acoustic guitar.

## ACKNOWLEDGMENTS

We would like to thank the Associate Editor and two anonymous reviewers for their valuable comments on a previous version of this manuscript, and Richard Lindsay for providing us a venue to test guitarists at the Ullapool guitar festival. SC was supported by a grant (BB/M007243/1) from the Biotechnology and Biological Sciences Research Council. Roger Bucknall is the owner of the "Fylde Guitars" guitar company. The underlying data in this paper is available from https://osf.io/f4pqa/?view_only=885abfa348324c5cbe8bf409a543601a [This is a temporary URL, we will provide a DOI once the manuscript is accepted].

---

[1] See supplementary material at [URL will be inserted by AIP] for additional methods details, additional figures, and additional results.

Bedell, T. (**2017**). "Rosewood, the blood diamond of music wood," Music Trades 26–30.

Boley, J., and Lester, M. (**2009**). "Statistical analysis of ABX results using signal detection theory," Audio Engineering Society.

Bones, O., Hopkins, K., Krishnan, A., and Plack, C. J. (**2014**). "Phase locked neural activity in the human brainstem predicts preference for musical consonance," Neuropsychologia **58**, 23–32, doi: `10.1016/j.neuropsychologia.2014.03.011`.

Bones, O., and Plack, C. J. (**2015**a). “Losing the music: aging affects the perception and subcortical neural representation of musical harmony,” J. Neurosci. **35**(9), 4071–4080, doi: 10.1523/JNEUROSCI.3214-14.2015.

Bones, O., and Plack, C. J. (**2015**b). “Subcortical representation of musical dyads: individual differences and neural generators,” Hear. Res. **323**, 9–21, doi: 10.1016/j.heares.2015.01.009.

Bourgeois, D. (**1994**). “Tapping tonewoods,” Acoustic Guitar **23**.

Carcagno, S., and Plack, C. J. (**2011**). “Pitch discrimination learning: specificity for pitch and harmonic resolvability, and electrophysiological correlates,” J. Assoc. Res. Otolaryngol. **12**(4), 503–517, doi: 10.1007/s10162-011-0266-3.

Christensen, O., and Vistisen, B. B. (**1980**). “Simple model for low-frequency guitar function,” J. Acoust. Soc. Am. **68**(3), 756 – 766, doi: 10.1121/1.384814.

DiStefano, C., Zhu, M., and Mindrila, D. (**2009**). “Understanding and using factor scores: Considerations for the applied researcher,” Pract. Assess., Res. Eval. **14**(20).

Ellis, A., Saufley, C., and Teja, G. (**2008**). “The future of tonewood,” Acoustic Guitar **18**(8), 80–86.

Fletcher, N. H., and Rossing, T. D. (**1998**). *The Physics of Musical Instruments*, Chap. 9, 239–271 (Springer New York, New York, NY), doi: 10.1007/978-0-387-21603-4_9.

Forbes-Roberts, R. (**2008**). “Carbon-fiber guitars,” Acoustic Guitar **19**(1), 62–67.

French, M. (**2008**). “Response variation in a group of acoustic guitars,” Sound and Vibration (1), 18–22.

French, M., and Handy, R. (**2006**). "Sustainability and life cycle management in guitar production," in *Proceedings of The 2006 IJME - INTERTECH Conference*, New Jersey, USA.

Fritz, C., Cross, I., Moore, B. C. J., and Woodhouse, J. (**2007**). "Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin," J. Acoust. Soc. Am. **122**(6), 3640–3650, doi: `10.1121/1.2799475`.

Fritz, C., Curtin, J., Poitevineau, J., Borsarello, H., Wollman, I., Tao, F.-C., and Ghasarossian, T. (**2014**). "Soloist evaluations of six Old Italian and six new violins," Proc. Natl. Acad. Sci. U.S.A. **111**(20), 7224–7229, doi: `10.1073/pnas.1323367111`.

Fritz, C., Curtin, J., Poitevineau, J., Morrel-Samuels, P., and Tao, F.-C. (**2012**). "Player preferences among new and old violins," Proc. Natl. Acad. Sci. U.S.A. **109**(3), 760–763, doi: `10.1073/pnas.1114999109`.

Fritz, C., and Dubois, D. (**2015**). "Perceptual evaluation of musical instruments: State of the art and methodology," Acta Acust. United Ac. **101**(2), 369–381, doi: `10.3813/AAA.918833`.

Gelman, A., and Hill, J. (**2007**). *Data analysis using regression and multilevel/hierarchical models* (Cambridge University Press, Cambridge).

Gore, T. (**2011**). "Wood for guitars.," J. Acoust. Soc. Am. **129**(4), 2519–2519, doi: `10.1121/1.3588331`.

Gough, C. E. (**2015**). "A violin shell model: Vibrational modes and acoustics," J. Acoust. Soc. Am. **137**(3), 1210–1225, doi: `10.1121/1.4913458`.

Greenberg, J. B. (**2016**). "Good vibrations, strings attached: The political ecology of the guitar," Sociology and Anthropology **4**(5), 431–438, doi: 10.13189/sa.2016.040514.

Haines, D. (**1979**). "On musical instrument wood," Catgut Acoust. Soc. Newsletter **31**, 23–32.

Hautus, M. J., and Lee, A. J. (**1998**). "The dispersions of estimates of sensitivity obtained from four psychophysical procedures: implications for experimental design," Percept Psychophys **60**(4), 638–649, doi: 10.3758/BF03206051.

Hautus, M. J., and Meng, X. (**2002**). "Decision strategies in the ABX (matching-to-sample) psychophysical task," Perception & Psychophysics **64**(1), 89–106, doi: 10.3758/BF03194559.

Hunter, D. (**2014**). "What's the big deal about Brazilian rosewood?," Guitar Player **48**(8), 130.

Inta, R. (**2007**). "The acoustics of the steel string guitar," Ph.D. thesis, University of New South Wales.

Johnston, R. (**2006**). "Forest of sound," Acoustic Guitar **16**(9), 80–82.

Johnston, R. (**2011**). "Brazilian rosewood," Acoustic Guitar **22**(2), 61–66.

Johnston, R. (**2013**). "Back and side woods," Acoustic Guitar **23**(7), 52–56.

Kruschke, J. K. (**2014**). *Doing Bayesian data analysis, a tutorial with R, JAGS, and Stan*, $2^{nd}$ ed. (Elsevier, London).

Macmillan, N. A., and Creelman, C. D. (**2004**). *Detection theory: A user's guide*, $2^{nd}$ ed. (Lawrence Erlbraum Associates, London).

McDermott, J. H., Lehr, A. J., and Oxenham, A. J. (**2010**). “Individual differences reveal the basis of consonance,” Curr. Biol. **20**(11), 1035–1041, doi: 10.1016/j.cub.2010.04.019.

McIntyre, M. E., and Woodhouse, J. (**1978**). “The acoustics of stringed musical instruments,” Interdiscipl. Sci. Rev. **3**(2), 157–173, doi: 10.1179/030801878791926128.

McIntyre, M. E., and Woodhouse, J. (**1988**). “On measuring the elastic and damping constants of orthotropic sheet materials,” Acta Metallurgica **36**(6), 1397–1416, doi: 10.1016/0001-6160(88)90209-X.

Micheyl, C., Delhommeau, K., Perrot, X., and Oxenham, A. J. (**2006**). “Influence of musical and psychoacoustical training on pitch discrimination,” Hear. Res. **219**(1-2), 36–47, doi: 10.1016/j.heares.2006.05.004.

Music Trades (**2016**a). “The music industry census,” Music Trades .

Music Trades (**2016**b). “New rosewood trade restrictions challenge guitar industry,” Music Trades 26.

Pedgley, O., Norman, E., and Armstrong, R. (**2009**). “Materials-inspired innovation for acoustic guitar design,” METU Journal of the Faculty of Architecture **26**(1), 157–175.

Relph-Knight, T. (**2011**). *Guitar Mechanics* (BookBaby).

Richardson, B. E. (**1994**). “The acoustical development of the guitar,” Catgut Acoust. Soc. J. **2**(5), 1–10.

Richardson, B. E. (**2002**). “Simple Models as a Basis for Guitar Design,” Catgut Acoust. Soc. J. **4**(5), 30–36.

Romanillos, J. L. (**1997**). *Antonio de Torres, guitar maker: His life and work* (Bold Strummer, Westport: CT).

Saitis, C., Giordano, B. L., Fritz, C., and Scavone, G. P. (**2012**). “Perceptual evaluation of violins: a quantitative analysis of preference judgments by experienced players,” J. Acoust. Soc. Am. **132**(6), 4002–4012, doi: 10.1121/1.4765081.

Saitis, C., Scavone, G. P., Fritz, C., and Giordano, B. L. (**2015**). “Effect of task constraints on the perceptual evaluation of violins,” Acta Acust. United Ac. **101**(2), 382–393, doi: 10.3813/AAA.918834.

Sandberg, L., and Traum, A. (**2000**). *The acoustic guitar guide: everything you need to know to buy and maintain a new or used guitar* (A Cappella).

Walraet, J., and Garston, B. (**2015**). “The Leonardo guitar research project (LGRP),” American Lutherie **124**, 23.

Wegst, U. G. K. (**2006**). “Wood for sound,” Am. J. Bot. **93**(10), 1439–1448, doi: 10.3732/ajb.93.10.1439.

Woodhouse, J. (**2002**). “Body vibration of the violin — what can a maker expect to control?,” Catgut Acoust. Soc. J. **4**(5), 43–49.

Woodhouse, J. (**2014**). “The acoustics of the violin: A review,” Rep. Prog. Phys. **77**(11), doi: 10.1088/0034-4885/77/11/115901.

Woodhouse, J., and Langley, R. S. (**2012**). “Interpreting the input admittance of violins and guitars,” Acta Acust. United Ac. **98**(4), 611–628, doi: 10.3813/AAA.918542.

Woodhouse, J., Manuel, E. K. Y., Smith, L. A., Wheble, A. J. C., and Fritz, C. (**2012**). “Perceptual thresholds for acoustical guitar models,” Acta Acust. United Ac. **98**(3), 475–486, doi: 10.3813/AAA.918531.

Wright, H. (**1996**). “The acoustics and psychoacoustics of the guitar,” Ph.D. thesis, University of Wales.

Yoshikawa, S. (**2007**). “Acoustical classification of woods for string instruments,” J. Acoust. Soc. Am. **122**(1), 568–573, doi: 10.1121/1.2743162.

## TABLES

| Back wood | $F_1$ (Hz) | $Q_1$ | $F_2$ (Hz) | $Q_2$ | $F_3$ (Hz) | $Q_3$ |
|---|---|---|---|---|---|---|
| Brazilian rosewood | 97 | 34 | 177 | 18 | 336 | 36 |
| Indian rosewood | 101 | 38 | 188 | 26 | 368 | 58 |
| Maple | 100 | 35 | 187 | 25 | 363 | 53 |
| Mahogany | 99 | 33 | 184 | 25 | 351 | 35 |
| Sapele | 96 | 39 | 175 | 17 | 335 | 44 |
| Walnut | 98 | 42 | 182 | 28 | 347 | 24 |

TABLE I. Measured modal frequencies and Q factors of low-frequency modes of the set of guitars.

TABLE II. List of questions used in the sound rating tests.

| | |
|---|---|
| Q1 | How do you like the overall sound of the guitar? |
| Q2 | How do you like the playability of the guitar? |
| Q3 | How bright is the sound of the guitar? |
| Q4 | How clear is the sound of the guitar? |
| Q5 | How warm is the sound of the guitar? |
| Q6 | How mellow is the sound of the guitar? |
| Q7 | How balanced is the sound of the guitar? |
| Q8 | How defined is the sound of the guitar? |
| Q9 | How strong is the tone separation of the guitar? |
| Q10 | How complex is the sound of the guitar? |
| Q11 | How rich is the sound of the guitar? |
| Q12 | How strong is the projection of the guitar? |
| Q13 | How strong is the sustain of the guitar? |
| Q14 | How wide is the headroom of the guitar? |
| Q15 | How loud is the guitar? |
| Q16 | How big is the sound of the guitar? |

## FIGURE LEGENDS

**FIG. 1.** (Color online) (A) Bridge admittance of the six experimental guitars. For comparison the bridge admittance of a different steel-string acoustic guitar (Yamaha FG-403MS) that was not part of the experimental set is also plotted. (B) Bridge admittances used for the synthesis-based listening tests described in Section VII, calculated from the measured admittance of a reference guitar, but modifying the low-frequency modal parameters as described in the text.

**FIG. 2.** (Color online) Measured (A) density and (B) stiffnesses of the tested wood samples. For each wood, two black symbols correspond to the two halves of one set, and two red symbols to the two halves of the second set. For the stiffnesses, stars denote $D_1$, circles $D_3$ and + symbols $D_4$.

**FIG. 3.** (Color online) Average "overall sound", and playability ratings obtained for each guitar in the blinded sound rating test. Error bars denote ±1 standard error of the mean (s.e.m.).

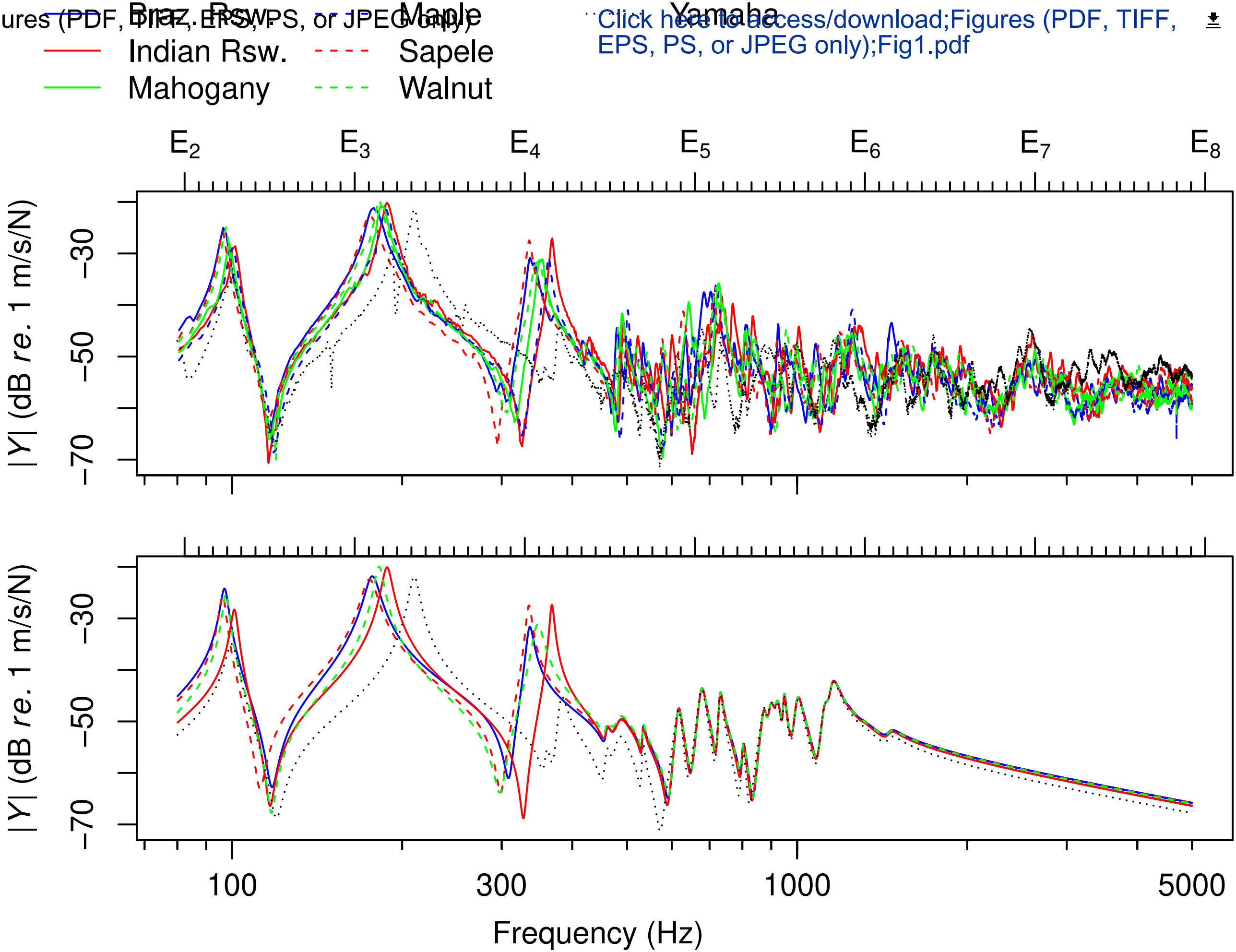
**FIG. 4.** (Color online) Bayesian posterior mode estimates (filled symbols) and 95% CIs (horizontal segments) for the rating differences in overall sound and playability between each guitar pair in the blinded sound rating test. The gray dotted lines mark a distance of ±0.2 points on the rating scale.

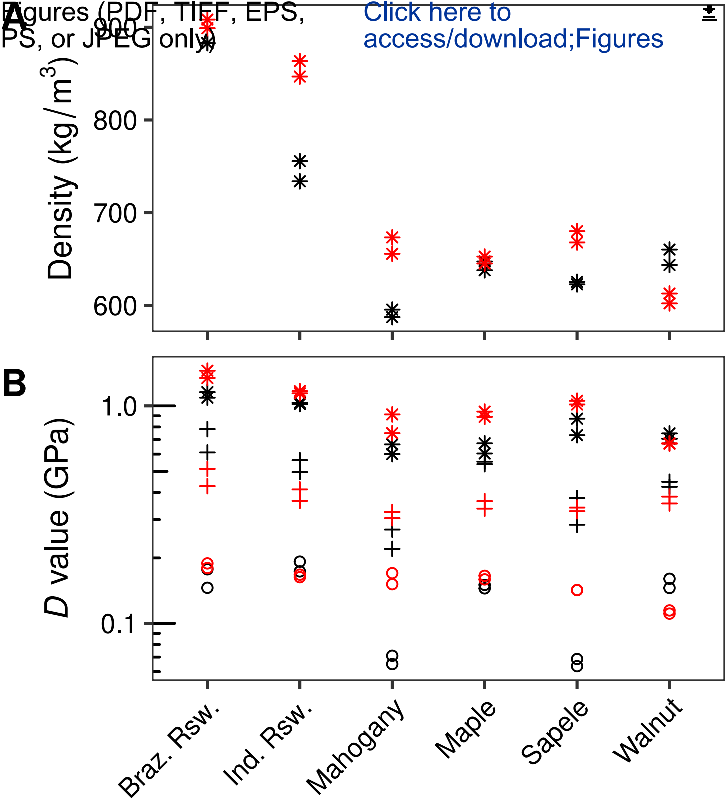
**FIG. 5.** (Color online) Average $d'$ values obtained for each guitar pair in the ABX test, as a function of player professional status. Error bars denote ±1 s.e.m.

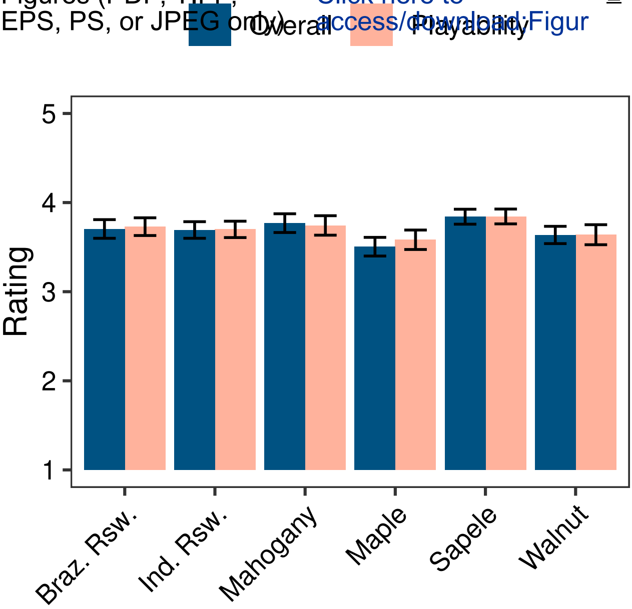
**FIG. 6.** (Color online) Bayesian posterior mode estimates (filled symbols) and 95% CIs (horizontal segments) of $d'$ for each guitar pair in the ABX test, as a function of player professional status.

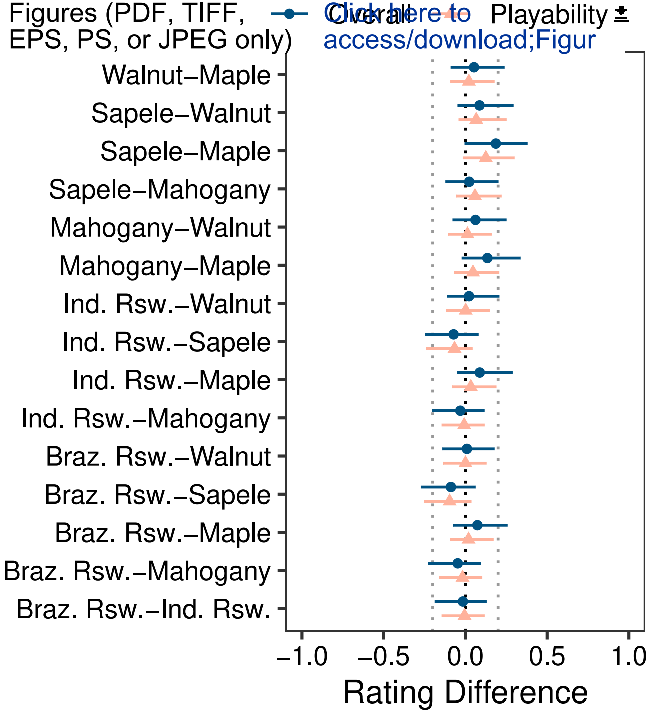
**FIG. 7.** (Color online) Average $d'$ values obtained for each guitar pair in the ABX test with virtual guitars. Error bars denote ±1 s.e.m.

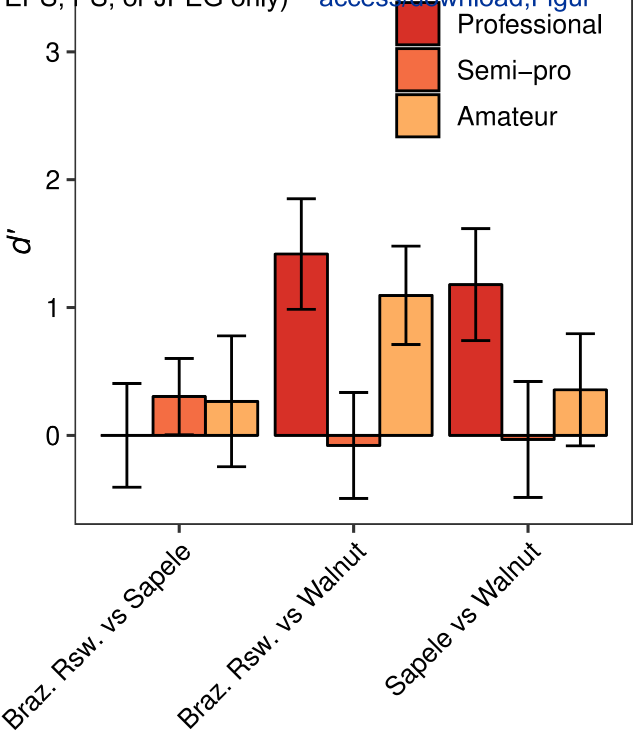
**FIG. 8.** Bayesian posterior mode estimates (filled symbols) and 95% CIs (horizontal segments) of $d'$ for each guitar pair in the ABX test with virtual guitars.

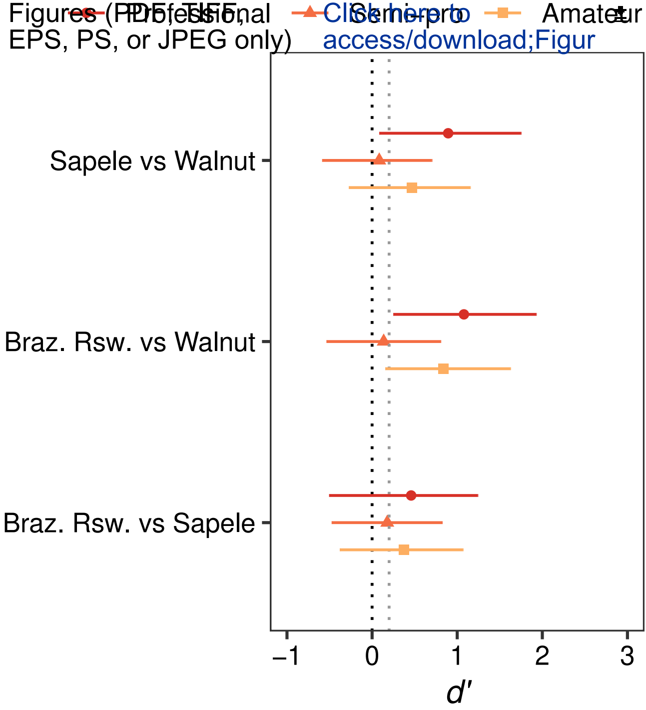

Braz. Rsw.
Maple
Yamaha
Indian Rsw.
Sapele
Mahogany
Walnut
E2
E3
E4
E5
E6
E7
E8
|Y| (dB re. 1 m/s/N)
−30
−50
−70
|Y| (dB re. 1 m/s/N)
−30
−50
−70
100
300
1000
5000
Frequency (Hz)

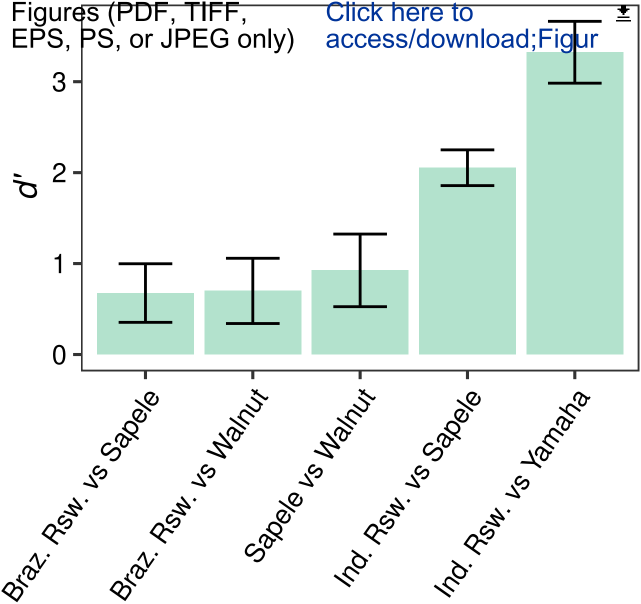

A
900
800
700
600
Density (kg/m$^3$)
B
1.0
0.1
$D$ value (GPa)
Braz. Rsw.
Ind. Rsw.
Mahogany
Maple
Sapele
Walnut

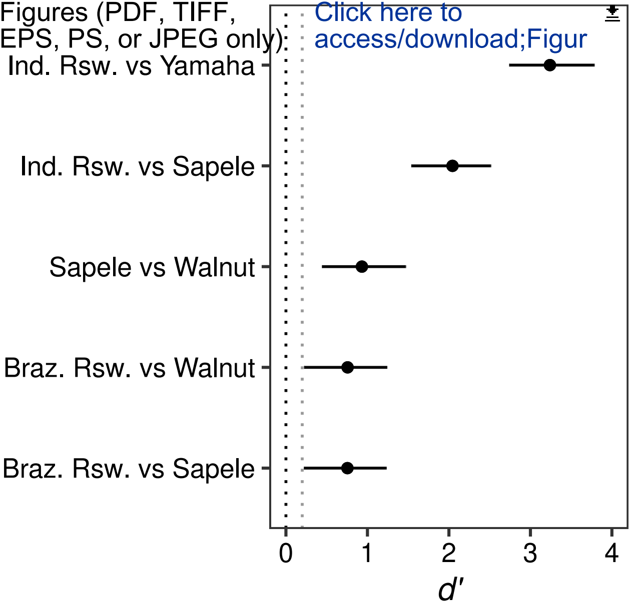

Overall
Playability
Rating
5
4
3
2
1
Braz. Rsw.
Ind. Rsw.
Mahogany
Maple
Sapele
Walnut


Overall
Playability
Walnut–Maple
Sapele–Walnut
Sapele–Maple
Sapele–Mahogany
Mahogany–Walnut
Mahogany–Maple
Ind. Rsw.–Walnut
Ind. Rsw.–Sapele
Ind. Rsw.–Maple
Ind. Rsw.–Mahogany
Braz. Rsw.–Walnut
Braz. Rsw.–Sapele
Braz. Rsw.–Maple
Braz. Rsw.–Mahogany
Braz. Rsw.–Ind. Rsw.
−1.0
−0.5
0.0
0.5
1.0
Rating Difference


Professional
Semi–pro
Amateur
d′
3
2
1
0
Braz. Rsw. vs Sapele
Braz. Rsw. vs Walnut
Sapele vs Walnut


Professional
Semi-pro
Amateur
Sapele vs Walnut
Braz. Rsw. vs Walnut
Braz. Rsw. vs Sapele
−1
0
1
2
3
d′


*d′*
3
2
1
0
Braz. Rsw. vs Sapele
Braz. Rsw. vs Walnut
Sapele vs Walnut
Ind. Rsw. vs Sapele
Ind. Rsw. vs Yamaha


Ind. Rsw. vs Yamaha
Ind. Rsw. vs Sapele
Sapele vs Walnut
Braz. Rsw. vs Walnut
Braz. Rsw. vs Sapele
0
1
2
3
4
$d'$

Supplemental Files for Publication

Rebuttal letter

Click here to access/download

**Helpful/Supporting Material for Reviewer**

Carcagno_et_al_rebuttal_letter.pdf

Manuscript with tracked changes

Supplemental material with tracked changes