Listener evaluations of new and Old Italian violins

Claudia Fritz, Joseph Curtin, Jacques Poitevineau, Fan-Chia Tao

Listener evaluations of new and Old Italian violins

Claudia Fritz[a], Joseph Curtin[b], Jacques Poitevineau[a], Fan-Chia Tao[c]

a Lutheries–Acoustique–Musique Team, Institut Jean Le Rond d'Alembert UMR 7190, Université Pierre et Marie Curie Univ Paris 06, CNRS, Sorbonne Universités, 75005 Paris, France;
b Joseph Curtin Studios, Ann Arbor, MI 48103;
c Research & Development, D'Addario & Company, Farmingdale, NY 11735

Published in PNAS, May 2017 (doi: 10.1073/pnas.1619443114)

Abstract

Old Italian violins are routinely credited with playing qualities supposedly unobtainable in new instruments. These qualities include the ability to project their sound more effectively in a concert hall – despite seeming relatively quiet under the ear of the player, compared with new violins. While researchers have long tried to explain the "mystery" of Stradivari's sound, it is only recently that studies have addressed the fundamental assumption of tonal superiority. Results from two studies show that under blind conditions experienced violinists tend to prefer playing new violins over Old Italians. Moreover, they are unable to tell new from old at better than chance levels. The current study explores the relative merits of Stradivari and new violins from the perspective of listeners in a hall. Projection and preference are taken as the two broadest criteria by which listeners might meaningfully compare violins: Which violins are heard better, and which are preferred? In two separate experiments, three new violins were compared with three by Stradivari. Projection was tested both with and without orchestral accompaniment. Projection and preference were judged simultaneously by dividing listeners into two groups. Results are unambiguous. The new violins projected better than the Stradivaris, whether tested with orchestra or without; the new violins were generally preferred by the listeners; and the listeners could not reliably distinguish new from old. The single best-projecting violin was considered the loudest under the ear by players, and on average, violins that were quieter under the ear were found to project less well.

Significance Statement

Old Italian violins are widely believed to have playing qualities unobtainable in new violins, including the ability to project their sound more effectively in a hall. Because Old Italian instruments are now priced beyond the reach of the vast majority of players, it seems important to test the fundamental assumption of their tonal superiority. A recent study found that under blind conditions violin soloists generally prefer new violins, and are unable to distinguish between new and old at better than chance levels. This paper extends the results to listeners in a hall. We find they generally prefer new violins over Stradivaris, consider them better-projecting, and are no better than players at telling new and old apart.

Introduction / context

Violinists generally agree that individual violins vary considerably in their ability to project – *projection* referring to how well an instrument can be heard at a distance, especially over a background of competing musical sound. The paradigmatic test for projection is the violin concerto, where a soloist is expected to carry over a full orchestra, often in a very large hall. Old Italian instruments are commonly believed to have the advantage over new ones in this regard. "What made the old instruments great was their power in a hall," wrote the distinguished violinist Earl Carlyss [1] in response to a 2010 blind study [2]. Somewhat paradoxically, Old Italian violins are also commonly described as being relatively quiet under the ear of the player, compared with new instruments. According to concertmaster Frank Almond (who plays a Stradivarius): "a peculiar (and sublime) aspect of great old Italian instruments is that the sound somehow expands and gains more complexity from a distance, especially in a concert hall." He contrasts this with many modern instruments, which seem to have a large sound under the ear, but may not "carry past the sixth row." [3]. Similarly, renowned cellist Steven Isserlis (who plays a Stradivarius cello) wrote, "A famous (and curious) feature of Stradivarius instruments is that their tone seems to increase with distance. As a rule, if my tone sounds small to me, it means that it is travelling out into the hall . . ." [4].

Most celebrated violinists since the early 1800s have played instruments by Stradivari or Guarneri *del Gesu*, and this is often taken as evidence that these violins possess some combination of playing qualities not found in newer instruments. Over the past two centuries, numerous informal playing and listening tests have challenged the notion [5,6,7]. More recently, a pair of well-controlled studies invited players to blind-test new violins against those by Stradivari and other Old Italian makers [2,8]. In both studies the most-preferred violin was new, and there was a general preference for new instruments over old. Moreover, players appeared unable to reliably distinguish between new and old.

These studies focused on judgments made by violinists while playing. Because projection must by definition be judged by listeners at a distance, many questions remained unanswered. Do Stradivari violins in fact out-project high-quality new instruments? Are better-projecting instruments typically quieter under the player's ear? Are the instruments that soloists prefer to play the ones that are found to project best in a hall? And can listeners tell whether a soloist is playing an Old Italian violin rather than a new one?

To answer these questions, we ran two experiments in which neither players nor listeners knew the identity of the test instruments. Projection and preference were taken as the broadest criteria by which listeners might meaningfully separate two violins in a hall: Which violin is heard better, and which is preferred? The first study was held in 2012 in a small concert hall near Paris, France, immediately following the above-cited preference test [8]. The second was held in 2013 in a larger hall in New York City. These studies, referred to here as the Paris and New York experiments, involved 55 and 82 listeners, respectively. In each case three new violins were tested against three by Stradivari. In Paris, the violins were tested with and without orchestral accompaniment. In New York, all tests were unaccompanied.

Materials and Methods

General methodology

In each experiment, a group of listeners was presented with a series of pair-wise comparisons. Each pair consisted of a new violin and a Stradivarius (the selection of the instruments is detailed below), and was played by a soloist behind an acoustically transparent screen. Players used their own bows, as they typically would when evaluating instruments in real-life. Modified welders' goggles together with much-reduced ambient lighting made it impossible for players to identify instruments by eye.

The Paris experiment was conducted in the 300 seat concert hall 'A coeur de ville' in Vincennes (a hall well-regarded for its acoustics), and the New York experiment in the 860 seat Great Hall at the Cooper Union building. Listeners were free to sit anywhere in the halls, though in Paris the first five rows were excluded.

Seven internationally renowned soloists (named Soloist 1-7 in SI) took part in the Paris experiment, six of whom participated in the player-preference experiment [8] conducted in the preceding two days. In New York, one soloist from the Paris experiment was joined by another distinguished violinist (SI Text). In both tests, the number of soloists who normally performed on an Old Italian instrument was about equal to the number normally performing on a new one (SI Text).

In Paris we invited experienced listeners, by which we mean listeners with expertize relevant to our subject; they included violin makers, players, musicians, audiophiles, music critics, composers, and acousticians. The New York experiment was a public event within Mondomusica (an international exhibition of handcrafted musical instruments); listeners were people who saw the advertisement and were interested in the topic. In both experiments, listeners who returned incomplete evaluation sheets were omitted from the analysis. The distribution of the remaining listeners (55 for Paris, 82 for New York) is provided in SI. At the beginning of the experiment, listeners were informed orally that participation was voluntary, and that all answers would be anonymized. Under these conditions, CNRS (the employer of the principal investigator) does not require any formal approval from a licensing committee prior to the experiment.

Detailed procedure for the Paris experiment

This experiment was divided into three parts: Parts 1 and 2 were concerned with the relative projection of old and new violins; Part 3 tested whether listeners could tell if the violins being played were new or old. The full score sheet given to the participants is available in SI.

1) Part 1

a) Testing Projection

Relative projection was tested by a series of pair-wise comparisons, structured as follows. A soloist played a short excerpt on each violin of a pair, then did the same thing again. This so-called ABAB format ensured listeners could hear each violin both before and after the other. The whole sequence was done twice for each pair, first using a solo excerpt, and then a concerto excerpt accompanied by the orchestra (details about the orchestra can be found in SI). It was made clear to listeners that the same pair was being played throughout. In order to test each of the Stradivari violins against each of the new, nine new/old pairs were formed. These were played in random order by Player 1, then after a short break by Player 2, but with the order of the violins within the pairs reversed, and the order of the pairs modified. Note that in choosing soloists for this test, we selected Player 1 as one who regularly plays a new violin (Soloist 1, see SI), and Player 2 as one who regularly plays an old (Soloist 2).

To facilitate testing, a researcher placed each pair of violins in turn on a table. A large sign (legible even in the relative darkness) placed in front of each instrument indicated to the soloist whether it was violin A or violin B. To help listeners keep track, a large video screen indicated which violin of which pair was being played (e.g., Pair 5, violin A).

Musical excerpts were selected from the standard repertoire, and were well-known to all soloists. The unaccompanied excerpts were chosen to cover a good deal of the instrument's range within about ten seconds. Soloist 1 played the opening solo (bars 23-27) from the first movement of the Tchaikovsky Violin Concerto op. 35. Soloist 2 played the opening solo (bars 90-94) from the first movement of the Brahms Violin Concerto in D major, op. 77. The accompanied excerpts were a bit longer (~20 s), and were passages in which the solo violin might easily be 'covered' by the orchestra: Soloist 1 played bars 187-202 from the third movement of the Brahms Concerto, and Soloist 2 played bars 41-43 from the 2nd movement of the Sibelius Violin Concerto, op. 47.

While projection is presumed to be an intrinsic quality of an instrument, it may well be affected by how good a "fit" a particular instrument is for a particular player. An experimental design allowing separation of these two factors would, however, require the audience to rate nine pairs of violins twice. The resulting 36 repetitions of the same excerpt would, we believe, have been an unreasonable challenge to listener concentration. To help avoid boredom and fatigue, the second soloist was therefore given a different excerpt. This trade-off between experimental design and the capacities of the listeners made it impossible to later decouple the intrinsic qualities of the violins from the influence of individual players.

For each pair of violins, listeners were required to answer the question: "Which violin projects best?" The meaning of *projection* was left undefined, but a questionnaire (SI Text) on the subject was sent to participants after the experiment ([9] provides a psycho-linguistic analysis of the answers). Listeners recorded their evaluations by placing a mark at the position of their choice along a continuous scale, where one end indicated that violin A was far superior to B in terms of projection. A mark at the opposite end indicated the converse, and a mark at the midpoint indicated the two projected equally well (SI Text). Though it was impossible for the soloists to evaluate violins in this way while playing them, they did indicate (to CF) which instrument of each pair they believed projected better, or that the two were equal.

b) Selection of test violins

In the Paris preference tests [8], ten soloists were presented with six new and six Old Italian violins (including five by Stradivari), which had themselves been preselected from a pool of nine Old Italian and 15 new violins made available by dealers, collectors, players, and makers (see [8] for details of the pre-selection process). Over the course of two 75-minute individual testing sessions, the soloists were asked to (1) eliminate any instruments they found unsuitable; (2)

choose from the remaining their four favorites, then rank these four in terms of overall preference, and (3) select the single violin that would best replace their own violin for an upcoming concert tour. Time limits restricted the number of violins that could be tested for projection to three new and three old. The following factors were considered in selecting these from the six old and six new violins which the soloists had evaluated:

1. The number of times an instrument was the top choice of a soloist. Five violins were chosen by at least one soloist. Three were new (N5, N10, & N9) and two were by Stradivari (O1 & O4).

2. Overall preference-score, based on soloists' choices of favorite violins. The three top-scoring new were, in descending order, N5, N10 and N11. The three top-rated old were O1, O8, and O4.

3. Measured acoustical output. Acoustical measurements were made of all instruments (SI Text). Preliminary results were used for a quick measure of the relative sound output of each instrument, per unit force at the bridge. The three new violins with the highest output were (in descending order) N5, N11, & N10. The three old were O6, O4, & O1.

Violin	Times chosen as favorite	Rank by overall preference score within old/new category	Rank by acoustical output within old/new category
N5	4	1	1
N10	1	2	3
N11	0	3	2
N9	1	3	4
O1	3	1	3
O4	1	3	2
O8	0	2	4
O6	0	4	1

Table 1: D*ata used in the selection of three new and three old violins for the listening experiment*

Based on the data in Table 1, we chose the following violins: N5, N10, N11, O1, O4 and O6. The inclusion of O6 needs justification, as it was not chosen by any soloist and was rated 9th of 12 for preference. It was selected mainly on the basis of its acoustical output, which was the highest of all the old violins. Moreover, unlike the well-liked O8 (also a strong contender) it was a Stradivarius, and so of most direct relevance to our hypothesis.

2) Part 2

Would the violins most frequently chosen by soloists in preference tests out-project those the soloists most often rejected? To test this, we had listeners compare the two most frequently chosen violins (N5 and O1) with two most often rejected (N2 and O12). The resulting four new/old pairs were tested as in Part 1, with players and listeners being asked the same questions. The solo excerpts were as above, but to increase musical variety, different orchestral excerpts were used: Player 1 (Soloist 3, see SI) played a 20s excerpt from the second movement of the

Brahms concerto (bars 93-98), and Player 2 (Soloist 4) played a 16 s excerpt from the third movement of the Sibelius concerto (bars 254-262).

3) Part 3

To test whether listeners could distinguish old violins from new, they were presented with a succession of seven concerto excerpts, each played (with orchestra) by a different soloist (Soloists 1-7, see SI). Unbeknown to listeners, the soloists played their own personal violins, thus forestalling concerns that they might be insufficiently acquainted with the test instruments to fully exploit their individual qualities. Each soloist had prepared a three-minute excerpt from either the Mendelssohn, Sibelius, or Beethoven concertos. After each performance, listeners were asked: "Do you think the violin is old or new? Why?"

Detailed procedure of the New York experiment

a) Testing projection and preference

The New York experiment tested how well the listeners' judgments of projection correlated with judgments of preference. It was structured in much the same way as Part 1 of the Paris experiment, with three Stradivari and three new violins yielding nine new-old pairs. Listeners were not asked to judge projection and preference at the same time: such a task would be cognitively challenging; moreover, the answer to the first question might influence the answer to the second, creating artificial correlations. Instead, listeners were randomly divided in two groups; one being asked first which violin of a pair they preferred, and why, and group 2 which violin projected better. After Player 1 (Soloist 8, see SI) had played the nine pairs, the questions were switched for Player 2 (Soloist 9), so that the first group now judged projection, and the second group, preference. Both projection and preference were reported on continuous scale, as in Paris. Unlike Paris, listeners were asked to define what projection meant to them by means of a multiple-choice questionnaire (SI Text). As in Paris, the soloists were asked to estimate projection on a simpler scale: A, B or equal. All tests were done without orchestral accompaniment.

b) Selection of test violins

We would have liked to use the same old and new test violins as in Paris, but just two of them were available: the new violin N5 and the Stradivari O6. Two additional new violins were chosen by a pre-selection process (SI Text) from 15 violins submitted by violin-makers. The two additional Stradivari were the only ones made available to us at the time.

Results

In this study, some Bayesian statistical procedures supplement traditional frequentist inferences, using LePAC software [10,11]. They assume an uninformative prior distribution (i.e. no information is used other than that contained in the data, and no particular hypothesis is favored *a priori*). Hereafter, Bayesian statements are notated *Pr**. They are naturally conditional on the data at hand, but conditional notation is omitted for the sake of brevity.

Paris experiment

a) Relative projection of three new violins and three Stradivarius (Part 1)

Figure 1 shows the relative projection of the nine new/old pairs tested in Paris. On the left are results for unaccompanied violin, and on the right for violin with orchestra. N10 was about equal to O1 and O4 in terms of projection, but in all other pairings, the new violins out-projected the old. The spaces between the blue and red error bars (which are especially wide for N11) suggest that projection depends somewhat on the player. While this may well be the case, no firm conclusions can be drawn, since the two players used different excerpts.

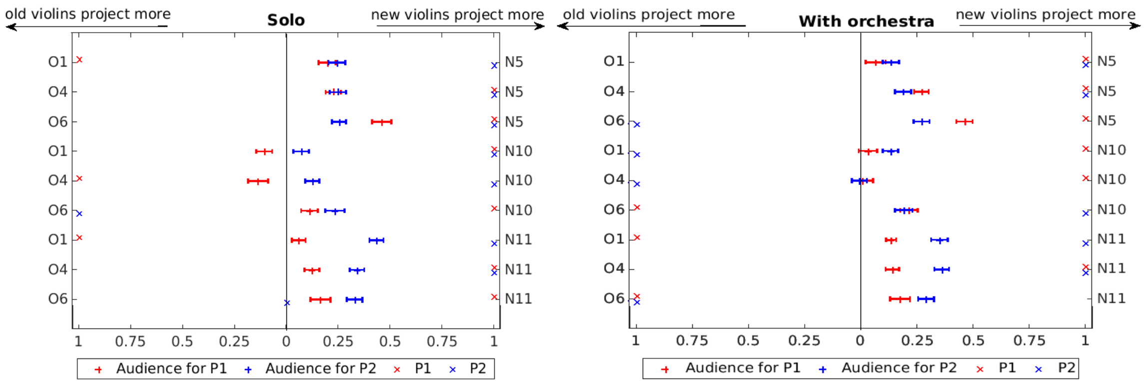

Figure 1: *Relative projection of the nine new/old pairs tested in Paris. On the left are results for unaccompanied violin, and on the right for violin with orchestra. Listeners scored projection on a continuous scale between 0 and 1. The charts show the average of their scores for each of two players – P1 in red, and P2 in blue. Error bars correspond to the standard error of the mean. The two players were asked to choose one of the following for each pair: (1) violin A projects best (indicated by a mark at the left edge of the graph; (2) violin B projects best (mark at right edge), or (3) the two violins project equally well (mark at center line). P1's judgements are shown in red and P2's in blue.*

Do violins that project well when played alone also project well over an orchestra? We find a striking consistency across the two conditions. Averaging the evaluations of all listeners and both players, the observed effect is -0.002 (solo - orchestra), $t(52)$ =-0.25, p = 0.40. Furthermore, we can state that $Pr^*(|\text{true effect}| < 0.02) = 0.95$ (on a 0-2 scale). This means, given the present data, a 95% probability that the true effect (i.e. the effect in the whole population of experienced listeners) is between -0.02 and 0.02. While non-null, the 0-1 scale means its value is very small, and a conclusion of negligibility can be drawn. In other words, very similar results are obtained whether testing projection with or without orchestra.

Do listeners from different professional backgrounds evaluate projection differently? By design, the entire audience consisted of experienced listeners, but they came from various professional backgrounds, and included musicians, violin-makers, and acousticians. All groups produced rather similar results. The observed root mean squared of the differences between groups is 0.04, $F(2,52) = 1.75$, p = 0.18 and $Pr^*(|\text{true effect}| < 0.08) = 0.95$ (on a 0-2 scale).

Given the consistencies described above, we felt justified in averaging the data from all listeners, with and without orchestra. Figure 2 shows the differences in projection between each new violin and all three old violins, and then between each old violin and all three new. Again we see a large new/old effect, with new violins judged to project better than old. The observed difference in projection is 0.19, $t(52) = 15.14$, $p < 10^{-4}$, 95% CI: [0.16; 0.21]. Pr^*(true effect > 0) > $1\text{-}10^{-4}$) and Pr^*(true effect > 0.17) = 0.95 (on a 0-1 scale). The apparent effect of the player is moderate (0.13) but clear as $t(52) = -5.82$, $p < 10^{-4}$, 95% CI: [0.09; 0.18] (in this case of a one-degree-of-freedom comparison with exactly the same assumptions as in the frequentist theory, the frequentist CI and the standard Bayesian CI coincide so that we can state Pr^*(0.09 < true effect < 0.18) = 0.95. Therefore, for similar cases in the following, we report only the CI limits.). Again, we cannot know whether this is due to the player or to the use of different excerpts.

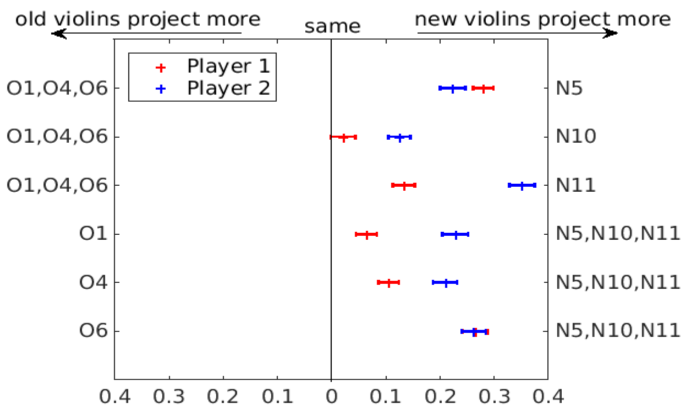

Figure 2: *Differences in projection between each new violin and all three old (top three rows) and between each old violin and all three new (bottom three rows) tested in Paris. Results are averaged for all listeners, with and without orchestra. Error bars correspond to the standard error of the mean.*

b) Relative projection of two preferred and two rejected violins (Part 2)

In Part 2 of the Paris experiment, the two most-preferred violins (N5 & O1) were compared with the two least-preferred (N2 & O12). Figure 3 shows that N5 clearly out-projected N2 & O12, and that O1 was about equal in projection to N2 & O12. There is good agreement between player and listener evaluations.

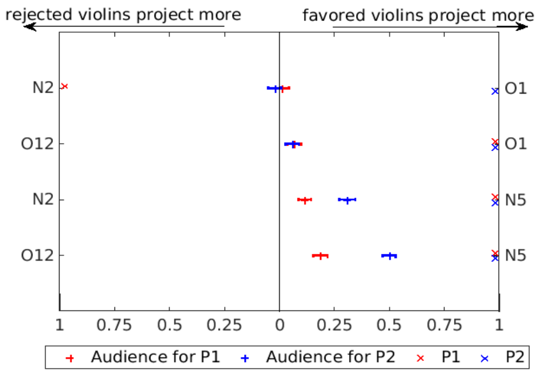

Figure 3: *Difference in projection between the two most preferred violins and the two most rejected. Error bars correspond to the standard error of the mean. The players' evaluations are single points at either 0 or 1 (left or right), with the same color code as for the audience. As the*

soloists gave the same answer with and without orchestra, there is no distinction between the two conditions, as in figure 1.

c) Comparison between soloist and listener evaluations of projection

Violinists generally believe that loudness under the ear is not directly related to projection, and that a violin's projection cannot be reliably estimated simply by playing it. During individual sessions that took place prior to the experiment [8], we asked each soloist to estimate projection and loudness-under-ear for three violins: their own, their favorite, and their next favorite of the opposite new/old category. This was done in a rehearsal room and again in the concert hall (see Table 2). Because of the selection method, only a few violins – N5, O1, and N10 – were rated by five or more soloists.

Violin	Projection	Loudness-under-ear
Soloists' own	8.2 (1.2)	7.9 (1.2)
Old	7.9 (1.3)	7.9 (1.2)
New	8.6 (1.3)	8.6 (1.2)

Table 2: *Mean (with standard deviation) of the loudness and projection ratings for soloists' own, old, and new violins, averaged across the two conditions – rehearsal room and concert hall.*

While the soloists gave their own violins (seven of which were old) a slightly higher rating for projection than for loudness (8.2 versus 7.9), when it came to the test instruments, the old got the same averaged rating (7.9) for both loudness and projection, as did the new (8.6). Since the new violins did in fact out-project the old, loudness-under-ear may be more closely related to actual projection than is generally believed.

Violin	Averaged estimated projection	Average loudness-under-ear	Number of estimates
N5	9,4	9,2	11
O1	8,8	7,6	6
N10	7,5	8,4	5

Table 3: *Averaged estimates for projection and loudness-under-ear for the three violins rated by at least five soloists during their private sessions in the Paris experiment*

Table 3 shows the averaged projection and loudness ratings for the three most-preferred violins, N5, O1, and N10. The projection-estimates wrongly suggest that O1 projects better than N10. The loudness ratings, however, at least get the ordering correct: N5, N10, O1 (in order of decreasing projection). The spread between projection and loudness seen in O1's ratings (8.8 v's 7.6) can be attributed to the four players who chose it as their favorite. Their ratings were (for projection and loudness respectively) 10 & 6; 9 & 7; 9 & 8, and 9.5 & 9. By contrast, the two soloists for whom O1 was not the favorite violin rated it very similarly for the two criteria (8 & 8 and 7.5 & 7). Given the small sample size, not too much can be made of this, though it does suggest that the widely-held belief that a Stradivari can project well despite sounding quiet under the ear may lead some players to over-estimate projection for instruments they like to play but find relatively quiet under the ear. There is indeed considerable anecdotal evidence that players and makers consider projection and loudness as two well separated concepts. This is illustrated in the analysis of the questionnaires filled in by participants in the New York experiment (73% of whom were makers and musicians). They were asked to select which of seven possible definitions best captured what projection meant to them (SI Text). Table 4 summarizes the results. The question referring to loudness was selected far less often than all but one of the others. More research is needed to investigate whether listeners do indeed evaluate projection and loudness differently, and if so, what the underlying acoustical parameters for each might be.

Possible definition of 'Which violin projects better?'	Number of times it was chosen
Which violin has the best carrying power?	29
Which sounds loudest?	14
Which sounds most full?	30
Which seems most "present" to your ears?	30
Which seems to fill the hall best?	30
Which comes through to your ears more clearly?	22
Which comes through to your ears more easily?	10

Table 4: *Number of times each possible definition of projection was chosen during the New York experiment.*

d) Distinguishing old from new (Part 3)

We asked listeners to guess whether each of seven violins was old or new. Thirty-nine of the 46 subjects made a guess about all seven violins. (Incomplete evaluation sheets were discarded from the analysis). The distribution of their correct guesses is shown in figure 4. It shows a unimodal distribution centered between three and four. In all, just 122 of 273 (or 44.7%) of the guesses were correct.

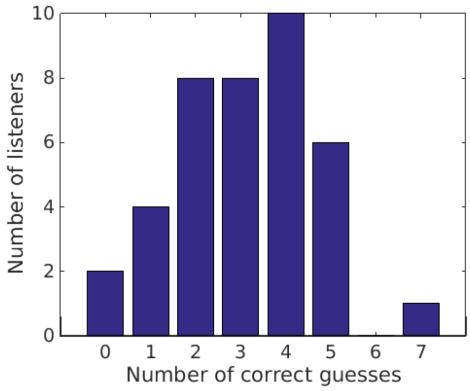

Figure 4: *Distribution of correct guesses about the age of seven violins by 39 subjects.*

It is interesting to see how the correct guesses were distributed among violins. For all but two instruments (one old, one new) the confidence intervals (SI Text) include 50%, so the chance hypothesis cannot be ruled out. For the other two violins, there is a systematic error: listeners guessed that the old violin was new, and the new one was old. A supplementary analysis built on the hypothesis that the guesses were entirely governed by chance is available in SI.

New York experiment

The Paris Experiment left at least one important question unanswered: How do listeners' evaluations of projection relate to their evaluations of preference? It is easy to imagine, for example, that violin A is found to project better than violin B, but violin B is preferred for its tone quality. Exploring the relationship between projection and preference was the principal objective of the New York Experiment.

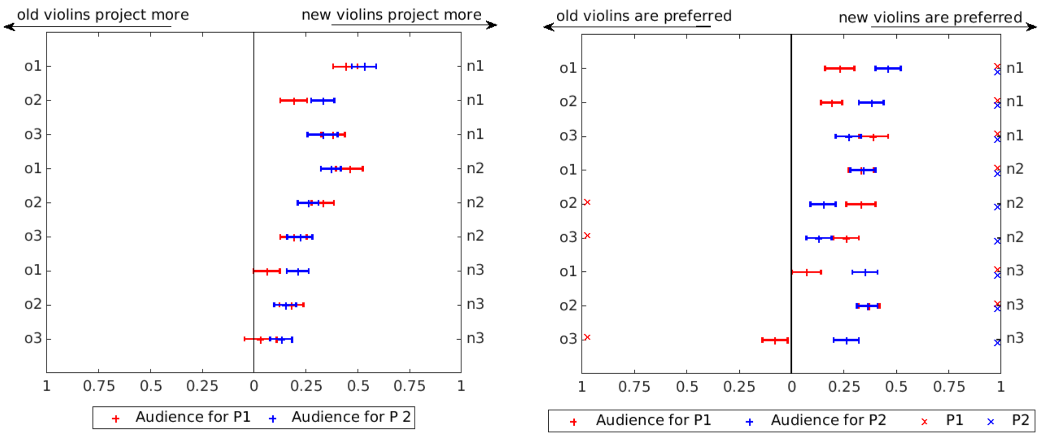

Figure 5: *Relative projection (left side) and preference (right) for each of the nine new/old pairs tested in New York. Listener scores are on a continuous 0 to 1 scale. The charts show the average of their scores for each of two players – P1 in red, and P2 in blue. Error bars correspond to the standard error of the mean. The two players were asked to choose one of the following for each pair: (1) I prefer violin A (*x *at the left edge of the graph; (2) I prefer violin B (*x *at right edge), or (3) I like them equally (*x *at center line). P1's judgements are shown in red and P2's in blue.*

Figure 5 shows relative projection (left side) and preference (right) for each of the nine new/old pairs. Looking first at projection, results are very similar to those of the Paris Experiment. New violins are found to project better than old, though in this case the difference happens to be even larger: observed difference = 0.27, $t(80)$ =14.55, $p < 10^{-6}$, 95% CI: [0.23; 0.31], Pr^*(true effect > 0) > 1- 10^{-6} and Pr^*(true effect > 0.24) = 0.95 (on a 0-1 scale). The better projection of the new compared to the old therefore seems robust across a change in halls, listeners, players, and violins. The observed effect of the player/order/excerpt is 0.03 (on a 0-2 scale), thus smaller than the effect of the player/excerpt in Paris (0.13 on a 0-2 scale). However, little can be inferred about the player effect as it cannot be separated from the effect of the excerpt (in Paris and New York) and the order (in New York only).

Looking at preference (Figure 5, right side), listeners clearly preferred new violins over old: observed difference = 0.27, $t(80)$ = 14.67, $p < 10^{-6}$, 95% CI: [0.23; 0.30], Pr^*(true effect > 0) > 1-10^{-6}) and Pr^*(true effect > 0.24) = 0.95 (0-1 scale). The figure shows slightly different results for the two players, but we cannot conclude there is a player effect, as it is inseparable from the possible effects of different excerpts and testing order (i.e., testing preference first rather than projection first). We can however conclude that if there is a player/order/excerpt effect on preference, it is rather limited: the observed value is 0.07 and $t(80)$ = 1.87, p = 0.07, Pr^*(|true effect| < 0.13) = 0.95 (on a 0-2 scale).

There is an almost 90% agreement between soloists and listeners about preference, at least when testing them together in a hall. Consider that player/listener agreement was just 70% when the soloists were asked to estimate projection under similar conditions in Paris.

Given that listeners generally preferred new violins over old *and* found them better-projecting, one may ask whether the new were preferred precisely because they projected better. As the listeners were asked to explain their preference choices, we look to the explanations for insight. Fifty-eight of 82 listeners used words related to projection in at least one case (SI Text), meaning that just 158 of the total of 548 collected statements refer to projection. (Listeners evaluated nine pairs, but many listeners didn't provide any explanations, or else provided them for just a few of the pairs). Among these 158 statements, some were of the kind 'I preferred violin A because of its tone quality, despite it being less loud/projecting than violin B' rather than 'I preferred violin A because it projects better'. This goes well with the extra definitions of projection given by two participants: “Doesn't always mean the best sound but one that cuts through noise, fills a room and allows the musician to be heard” and “full and loud, but not necessarily beautiful; sometimes harsh.” All of which to say, listener-explanations do not suggest any obvious link between preference and projection.

Conclusions and discussion

This study compares Stradivari and new violins principally from the listener’s perspective. Projection and preference are taken as the broadest criteria by which listeners might meaningfully separate violins in a hall. Which violins are heard better, and which are preferred? Also explored is the relationship between evaluations made by listeners and those made by players. Two separate experiments were conducted, one in the suburbs of Paris, the other in New York City. In Paris, violins were played with and without orchestra by blindfolded soloists behind an acoustically transparent screen. The same was done in New York, but without an orchestra.

The results are unambiguous: listeners found that new violins projected significantly better than those by Stradivari. Moreover, listeners preferred new violins over old by a significant margin. Though the listeners came from various professional backgrounds (they included musicians, violin-makers, and acousticians), very similar results were obtained from all backgrounds.

We find a strong correlation between projection with and without orchestra. This seems fortunate for both players and researchers, in that an orchestra is evidently not required to meaningfully test projection.

Projection estimates made by players during private sessions [8] did not agree very well with judgments made later by listeners in the hall. When playing for an audience in a hall, however, about 70% of player-estimates of projection agreed with listener-judgments. Put another way, players were wrong about 30% of the time, confirming the wisdom of bringing a trusted listener along when trying out instruments. On the other hand, when players in the New York experiment were asked, after playing each new/old pair, which violin they preferred, their preferences were in strong agreement with those of the audience.

During private sessions in Paris [8], four of ten soloists chose a Stradivari as the instrument "best suited to replace their own violin for an upcoming tour." We find here that these four soloists chose violins with significantly less projection than the two best-projecting new instruments. The most-preferred Stradivari barely out-projected the least-preferred Stradivari and the least-preferred new violin. The New York results suggest that the most-preferred Stradivari in Paris would also be less-preferred by listeners than the better-projecting new violins.

The most-preferred new violin in the Paris Experiment was considered louder under the ear than the most-preferred old – and indeed than every other violin that was rated for loudness. Soloists who chose a Stradivari may have preferred a violin that was less loud under the ear, at the same time believing it would project well, for the three players who chose the most-preferred Stradivari rated it notably higher for projection than for loudness.

Previous studies [2,8] indicate that under blind conditions violinists readily separate violins they like from those they do not, but are unable to reliably distinguish new from old. We show here that listeners (including violin-makers and musicians) are no better at telling whether a soloist's violin is new or old.

The test violins were necessarily few in number, and there is no way of knowing how representative they are of the larger population. While there is every reason to believe there are Stradivari violins with outstanding projection, we find little reason to suppose players would find them quiet under the ear. Violins (old or new) that are (1) relatively quiet under the ear, and (2) out-project apparently louder violins would be worthy subjects for future research. To date, we have found no evidence for their existence.

A player's requirements for projection are clearly dependent on musical context. Even a relatively quiet violin will be heard when played unaccompanied in a small hall, and a good conductor will do everything possible to ensure that a soloist with a relatively quiet violin can still be heard over an orchestra. In auditions and playing competitions, however, jurors listen to one player after another, making apparent the relative projection of their instruments. While it is common for gifted young players to borrow Old Italian violins for such occasions, our results suggest they would be better served by good new violins – at least so long as the identity of the instruments remains unknown to those judging.

Note that this study relied on soloists, whose primary need is to be heard over an ensemble. The vast majority of violinists, however, are required to blend into ensembles. Research is needed to clarify the relationship between projection and "blendability." Less-projecting violins could have the advantage in this regard.

A belief in the near-miraculous qualities of Old Italian violins has preoccupied the violin world for centuries. It may be that recent generations of violin-makers have closed the gap between old and new, or it may be that the gap was never so wide as commonly believed. Either way, the debate about old versus new can perhaps be laid aside now in favor of potentially more fruitful questions. What, for example, are the physical parameters determining the playing qualities of any violin, regardless of its age or country of origin?

Acknowledgements

We thank all dealers, makers, players, and collectors for their kindness and trust in making available these valuable instruments. Special thanks go to the soloists for their participation and enthusiasm, to the numerous audience members for their dedication and patience. Special thanks as well to the COSU players, their conductor Vincent Barthe and their administrators, Agnès Pussilieux and Lucie Paladino, for their kindness and professionalism! We acknowledge Philip De La Croix and Stéphane Agasse for the use of the Vincennes auditorium, and for their logistical help, and Kojiro Yamada and NHK for their amazing support in New York. We would also like to thank the LAM team (particularly Indiana Wollman and Laurent Quartier), the Borsarello family, David Griesinger, Stefan Avalos, and Suzanne Ortmeier for their kindness and help throughout the Paris experiment. And, finally, we are grateful to the Centre National de la Recherche Scientifique and Université Pierre et Marie Curie for funding this experiment and to the Violin Society of America for additional financial support.

References

1. Wade N (January 2 2012) In classic vs. modern violins, beauty is in the ear of the beholder. NY Times. Available at http://www.nytimes.com/2012/01/03/science/inplay-off-between-old-and-new-violins-stradivarius-lags.html. Accessed April 10, 2017.

2. Fritz C., Curtin J., Poitevineau J., Morrel-Samuels P. and Tao F.-C. (2012) Players preferences among new and old violins. Proceedings of the National Academy of Sciences of the USA. 109: 760-763.

3. Frank Almond (2012) “They blinded me with Science”. Personal blog. http://www.insidethearts.com/nondivisi/they-blinded-me-with-science/
4. Steven Isserlis (January3, 2012) “Stradivarius v modern violins: why this latest study strikes a discordant note”. *The Guardian* Available at https://www.theguardian.com/music/musicblog/2012/jan/03/stradivarius-v-modern-violins-study. Accessed April 10, 2017.
5. Moya, H. and Piper, T. (1916). Violin tone and violin makers (Chatto & Windus, London).
6. Windust E (2009) From the archive: 100 years ago, a Paris competition sets out to compare old and new violins. *The Strad* 120: 95 .
7. Coggins, A. (2007). “Blind faith”, The Strad 118: 52-55.
8. Fritz C., Curtin J., Poitevineau J., Borsarello H., Wollman I., Tao F.-C. et Ghasarossian T. (2014) Soloist evaluations of six Old Italian and six new violins. Proceedings of the National Academy of Sciences 111: 7224-7229.
9. Dubois D. and Fritz C. (2016) “La projection du violon : analyse sémantique” (Violin projection : semantic analysis), In *Le violon en France du XIX^{e} siècle à nos jours*, ed C. Fritz and S. Moraly (in press).
10. Lecoutre, B. and Poitevineau, J. (2005) Le logiciel “Le *PAC”* dédié aux procédures bayésiennes. *La Revue de Modulad* 33. Available at https://www.rocq.inria.fr/axis/modulad/logiciels.htm#lepac. Accessed April 10, 2017.
11. Lecoutre, B. and Poitevineau, J. (2014) *The Significance Test Controversy Revisited. The Fiducial Bayesian Alternative.* New York: Springer.

Supporting information

Information about the soloists and the violins they played during the experiment

Venue	Soloist	Name	Own violin	Favorite violin
Paris	1	Ilya Kaler	new	N5
	2	Yi-Jia Susanne Hou	old	N5
	3	Tatsuki Narita	old	O4
	4	Giora Schmidt	new	N10
	5	Solenne Paidessi	old	O1
	6	Elmar Oliveira	new	N9
	7	Marie-Annick Nicolas	new	
New York	8	Elizabeth Pitcairn	Strad	
	9	Giora Schmidt	new	

Table S1: *Information about the players who took part in each experiment*

Information about the listeners

Listeners who turned in incomplete data-sheets were excluded from the analysis. The distribution of the remaining (55 in Paris, 82 in New York) is described in Table S2:

	makers	violinists	other	N/A	Mean age
Paris	15	19	21	0	N/A
New-York / Group 1	18	11	13	3	49.2 years
New-York / group 2	11	10	13	3	46.8 years

Table S2: *Information about the listeners who participated in each experiment*

Information about the orchestra

The Sorbonne Universités student orchestra (called COSU) was supplemented by several professional players and conducted by Vincent Barthe. In all there were 10 violins, 5 violas, 3 cellos, 1 double bass, 2 flutes, 2 clarinets, 1 oboe, 1 bassoon, 2 horns, and timpani. To allow players to read their parts on the darkened stage, small directional lamps were attached to the music stands.

Acoustical measurements

The acoustical output (sound pressure per unit force at bridge) for each violin measured using an impulse hammer and microphone. The bridge was tapped both horizontally and vertically, and the sound pressure measured at 12 equatorial microphone positions 20 cm from the center of the violin in the plane of the bridge [Curtin J (2009) Measuring violin sound using an impact hammer. *J Violin Soc Am: VSAPapers* XXII:186–209]. This yielded 24 frequency response functions (FRFs) for each violin. The real average for all 24 FRFs was calculated and then reduced to a single average (in dB) for the frequency band spanning 200 – 6400 Hz.

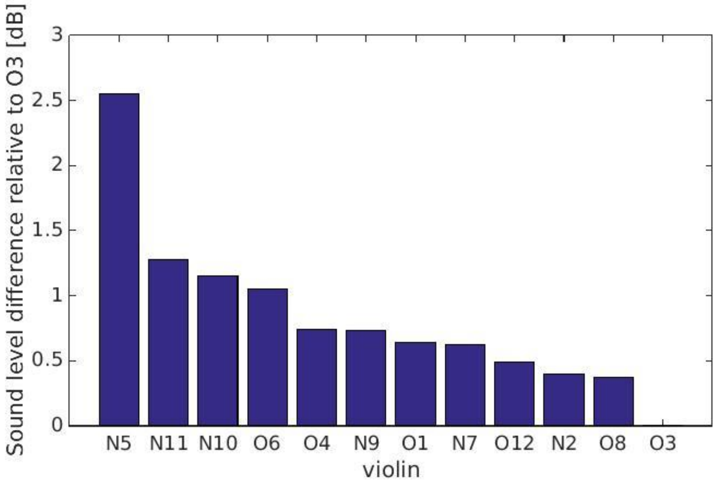

Figure S1: *Difference in averaged sound output level (per-unit-force at bridge) for each test violin relative to violin O3, for the 200 – 6400 Hz frequency band (represented in dB).*

The three new violins with highest levels are, in descending order, N5, N11, and N10, and the three old are O6, O4, and O1.

Selection of the violins for the New York experiment

New violins were chosen from a pool of 15 violins (including N5) submitted by their makers. A pre-selection took place in the same venue as the experiment. Author JC and the distinguished violin maker Sam Zygmuntovicz played the violins under blind conditions and selected what they considered the best eight. The selection of three new instruments was made by a blind test similar to the pre-selection in Paris [8]: Three blindfolded soloists (the two participating in the experiment, along with Karen Gyomo) each played the eight violins behind an acoustically transparent screen for a small audience (JC, F-C T, SZ and the two soloists who were not playing). Three violins (including N5) were chosen via an informal discussion of the player- and listener-preferences. There was good agreement among all parties.

Score sheet used in Paris

PART 1 **Name:**______________________________

Email:______________________________ **Seat Number in Hall**____________________

Expertise: (violin maker, professional etc.)____

In Parts 1 and 2, you will hear a series of comparisons between pairs of violins. The instruments in each pair will be referred to as Violin A and Violin B. A violinist will play a short solo excerpt on violin A, then violin B, then violin A again, then violin B again. Which violin projects better? Please see below for an example of how to mark your answer.

The player will next play an excerpt with orchestra on that same pair of violins in the same order. Which violin projects better now?

PLAYER 1

	A Superior	A=B	B Superior	
Example #1	|-----------------------------|---X------------------------|			B slightly superior to A
Example #2	|-----X----------------------|-----------------------------|			A very superior to B

	A Superior	A=B	B Superior
Q1a. Pair 1, Solo:	|-----------------------------|-----------------------------|		
Q1b. Pair 1, Orchestra	|-----------------------------|-----------------------------|		
[…]			
Q9a. Pair 9, Solo	|-----------------------------|-----------------------------|		
Q9b. Pair 9, Orchestra	|-----------------------------|-----------------------------|		
	A Superior	**A=B**	**B Superior**

Part 1 Continued:

PLAYER 2
Which violin projects better?

[Same score sheet as for Player 1]

PART 2

The instructions are the same as for part 1. Which violin projects better?

[Same score sheet as for Part 1, with 8 pairs played solo and then with orchestra]

PART 3

Each player will now play about 2:30 minutes from a concerto on just one violin. Do you think the violin is new or old? Why?

Player	New or old?	Why?
1	New Old	
2	New Old	

[...]		
7	New Old	

Questionnaire sent to the listeners after the Paris experiment

1. What is your definition of projection, i.e. the one that you used to evaluate the different violins?
2. Were there one or more musicians for whose the evaluation was easier/harder? Were there one or more excerpts for which the evaluation was easier/harder?

Score sheet used in New-York

Name:________________________________Age:_________ Email:_________________________

Expertise: (violin maker, professional etc.)______________ **Seat and row number in Hall: ___**

PART 1:

You will hear a series of comparisons between pairs of violins. The instruments in each pair will be referred to as Violin A and Violin B. A first violinist will play a short excerpt on violin A, then violin B, then violin A again, then violin B again. You then have to answer one question, before the violinist plays the next pair. After 10 pairs, the violinist will change, as well as the question (part 1B).

PART 1A: player 1

Question: Which violin do you prefer and why? Please see below for an example of how to mark your answer.

	A preferred	A=B	B preferred	Why?
Example #1 I much prefer A than B	|-----X----------------------|-----------------------------|			Because A has this and that while I don't like this and that in B.
Example #2	|------------------------X--|-----------------------------|			

I slightly prefer A than B

Example #3 |----- ------------------------X-----------------------------|

I like A as much as B

Example #4 |-----------------------------|-----X----------------------|

I prefer a bit more B than A

Example #5 |-----------------------------|-------------------------X--|

I considerably prefer B than A

	A preferred	A=B	B preferred	Why?
Pair 1 (practice)	|-----------------------------	|	-----------------------------|	
Pair 2	|-----------------------------	|	-----------------------------|	
[...]				
Pair 10	|-----------------------------	|	-----------------------------|	

PART 1B: player 2

Which violin projects better?

This question may mean slightly different things to different people, for example:

- Which violin has the best carrying power?
- Which sounds loudest?
- Which sounds most full?
- Which seems most "present" to your ears?
- Which seems to fill the hall best?
- Which comes through to your ears more clearly?
- Which comes through to your ears more easily?

Please check the one/s that best capture what projection means to you.

If you have some other definition, please write it here:

Please use the scale the same way as before by replacing “I prefer A more” by “A projects more” as illustrated in the example below:

A projects best **A=B** **B projects best**

Example #1 |-----X----------------------|-----------------------------|

A projects much more than B

Pair 1 (practice) |-----------------------------|-----------------------------|

Pair 2 |-----------------------------|-----------------------------|

[…]

Pair 10 |-----------------------------|-----------------------------|

Statistical analysis for the old/new guesses (Paris part 3)

For each of the 7 violins, the numbers of correct guesses over the total number of guesses are presented along with two intervals, the first one being the 95% Clopper-Pearson interval and the second one the Bayesian Credible Interval with Jeffreys’prior:

violin 1 (old): 7/45 [6.49%; 29.46%],[7; 28.13%]
violin 2 (new): 11/46 [12.59%; 38.77%], [13.42%; 37.57%]
violin 3 (N): 27/45 [44.33%; 74.30%], [45.43%; 73.33%]
violin 4 (N): 26/45 [42.15%; 72.34%], [43.24%; 71.35%]
violin 5 (O): 24/43 [39.88%; 70.92%], [41.00%; 69.88%]
violin 6 (N): 22/44 [34.56%; 65.44%], [35.62%; 64.38%]
violin 7 (O): 23/46 [34.90%; 65.10%], [35.92%; 64.08%]

What if the guesses were governed entirely by chance, as if the subjects were in fact tossing coins? All guesses would then be independent one from each other – both from one subject to the next, and among each subject’s guesses. In this case, the number of correct guesses would be distributed as a binomial distribution with parameter *n* being the total number of guesses (273), and parameter $\pi = 0.5$ being the probability of a correct guess. Then the 95% Clopper-Pearson confidence interval would be [38.7%; 50.8%]. It includes 50%, meaning that the hypothesis that chance entirely governed the guesses cannot be ruled out. Nonetheless, given these data it could be stated that $Pr^*(40.9\% < \text{true proportion} < 48.6\%) = 0.80$, just as if the subjects were using a biased coin with less than 50% chance of success! But that rests on the assumption of every subject performing independent guesses which is far from plausible.

Explanations provided by the listeners in New York to justify their preferences

Only the explanations related to projection and loudness were taken into account here. The terms that we considered associated with projection are the following:

project(ion), loud, quiet, power(ful), muted, big sound, small sound, direct sound, carrying power, strong, weak, present, full. Clear was added to this list for the listeners who selected the '"Which comes through to your ears more clearly" as a possible definition for projection.