The effects of acoustic variability on absolute pitch categorization: Evidence of contextual tuning

Stephen C. Van Hedger,[a] Shannon L. M. Heald, and Howard C. Nusbaum
Department of Psychology, The University of Chicago, 5848 South University Avenue, Chicago, Illinois 60637, USA

(Received 19 February 2015; revised 28 April 2015; accepted 13 June 2015; published online 21 July 2015)

Absolute pitch (AP) is defined as the ability to label a musical note without the aid of a reference note. Despite the large amounts of acoustic variability encountered in music, AP listeners generally experience perceptual constancy for different exemplars within note categories (e.g., recognizing that a C played on a tuba belongs to the same category as a C played on a piccolo). The present studies investigate whether AP possessors are sensitive to context variability along acoustic dimensions that are not inherently linked to the typical definition of a note category. In a speeded target recognition task, AP participants heard a sequence of notes and pressed a button whenever they heard a designated target note. Within a trial the sequence of notes was either blocked according to note-irrelevant variation or contained a mix of different instruments (Experiment 1), amplitude levels (Experiment 2), or octaves (Experiment 3). Compared to the blocked trials, participants were significantly slower to respond in the mixed-instrument and mixed-octave trials, but not the mixed-amplitude trials. Importantly, this performance difference could not be solely attributed to initial performance differences between instruments, amplitudes, or octaves. These results suggest that AP note identification is contextually sensitive.
[http://dx.doi.org/10.1121/1.4922952]

[JFC] Pages: 436–446

I. INTRODUCTION

Absolute pitch (AP) is defined as the rare ability to name any musical note without the aid of a reference note (e.g., Deutsch, 2013). The genesis and true nature of AP ability has been the subject of much debate over the past century (e.g., Mull, 1925; Neu, 1947; Bachem, 1937; Baharloo *et al.*, 2000; Levitin and Rogers, 2005; Ross *et al.*, 2005), partly because the operationalization of AP ability has not been standardized across studies and paradigms.

At the heart of the debate surrounding AP ability is the nature of the category representations of AP possessors. AP note categories correspond to the 12 distinct pitch classes (also known as "chroma") found in Western music, which can be defined independently of overall pitch height given that pitch is heard in a helical fashion with note chroma repeating at the octave (cf. Ueda and Ohgushi, 1987). These 12 pitch categories are also defined independently of other acoustic attributes, such as timbre and overall amplitude. In theory, an AP possessor should be able to accurately extract the note chroma from (and thus accurately provide a note category label for) any pitched sound within the frequency range of pitch perception for humans (e.g., Atteneave and Olson, 1971; Ward, 1954), regardless of acoustic variations such as timbre, overall pitch height, and amplitude.

However, more and more research has supported the idea that the auditory nature of AP category representations vary across individuals, and are at least in part dependent on one's particular experiences with certain timbres, notes, and frequency ranges. Indeed, both systematic and idiosyncratic limitations in AP identification have been described over the past several decades, lending support to the notion that category representations for AP possessors are grounded in specific experiences. AP possessors tend to be worse (i.e., slower and less accurate) at identifying "black key" notes (C#/Db, D#/Eb, F#/Gb, G#/Ab, A#/Bb) versus white key notes (C, D, E, F, G, A, B), with the notes "C" and "G" being the easiest to classify (Miyazaki, 1990). Furthermore, the accuracy profiles of AP possessors appear to be akin to an experiential fingerprint, reflecting an individual's particular experiences with specific note ranges, instruments, and tonalities (Bahr *et al.*, 2005). "Absolute piano," an extreme but illustrative example of timbre-specific note categories, refers to an ability in which an individual can *only* pass tests for AP when the test stimuli are taken from a specific timbre, in this case, the piano (Ward and Burns, 1982). In fact, the individual and systematic performance differences in AP ability have led many researchers to view AP as a continuous, multidimensional trait, rather than an "all-or-none" clearly delineated perceptual ability (e.g., Takeuchi and Hulse, 1993; Bermudez and Zatorre, 2009).

The evidence that particular experiences with certain timbres and frequency ranges affect AP possessors' ability to correctly label a note suggests that such information may not be discarded in the process of pitch perception but rather is a part of a rich and highly detailed pitch representation system. However, it is unclear what exact role these acoustic characteristics play in AP perception, especially since many AP possessors display perceptual constancy (e.g., recognizing that a C played on both a tuba and a flute belong to the same

[a]Electronic mail: shedger@uchicago.edu

category). One possibility is that such acoustic characteristics only matter in note chroma identification insofar as an individual has experience with particular instruments or octaves. In other words, the greater the experience with a particular acoustic attribute, the stronger the note category representation is for more frequently experienced timbres and octaves. Another possibility is that such non-canonical acoustic characteristics are not solely incorporated into note category representation, but rather provide a perceptual reference or framework by which note chroma is recognized. According to this second possibility, the acoustic *context* surrounding the fundamental frequency of a note (either simultaneously as in note timbre or sequentially as in prior notes) may play a role in the recognition process. This contextual sensitivity has been discussed in *contextual tuning theories*, used to explain perceptual constancy in speech (e.g., Gerstman, 1968; Nusbaum and Morin, 1992). If *contextual tuning theories* of speech can equally be applied to note perception, then it is possible that AP note identification may be sensitive to the variation in such characteristics across notes, as individuals may maintain the framework for a previous note to understand the next. In other words, a sequence of notes may establish perceptual expectations that constrain the possible identities of subsequent notes. For example, variation in the instrument or octave across a set of notes may affect recognition performance for a given note within the set, as the uncertainty caused by the variation will first have to be resolved.

The present studies assess whether AP recognition performance—via a simplified note recognition task (in which participants monitor a string of musical notes for a pre-specified target)—is affected by acoustic variation in instrument (Experiment 1), amplitude (Experiment 2), or octave (Experiment 3) for a given set of notes. If listeners' performance suffers when target notes are presented with uncertainty about the instrument, amplitude, or octave, then it suggests that AP note representations are context sensitive to these acoustic dimensions. Of course, this inference depends entirely on performance when recognizing target notes from a single instrument, amplitude, or octave. Given that the definition of AP note recognition is that such notes are identified without reference to a contextual note (which would be relative pitch and not AP), note recognition performance on a blocked presentation should be equal to a mix of putatively chroma-irrelevant acoustic attributes. If performance in the blocked setting (single instrument, amplitude, or octave) is better than in a mixed setting (variable instrument, amplitude, or octave), then reduced performance in the mixed setting cannot be explained simply in terms of differential performance previously reported (as in absolute piano). This is the first hypothesis tested in the current set of studies.

If, on the other hand, slower note recognition in a mixed-acoustic setting can be completely explained by pre-existing performance differences in a single instrument, amplitude, or octave tested in isolation (the blocked setting), then it suggests that the dimensions of instrumental timbre, amplitude, and octave range only matter in note chroma identification insofar as an individual has experience with particular instruments, amplitudes, or octaves (or insofar as the fundamental frequency is audible). This alternative possibility would still support the idea that AP categories are experience dependent, though it would not suggest that the *immediate* acoustic context is important in making an AP category judgment. Rather, it would suggest that long-term experience shapes the category strength of particular instruments, amplitudes, and octaves, but AP note perception cannot be explained in terms of *contextual tuning theories* (e.g., Nusbaum and Morin, 1992). This is the second hypothesis that is tested.

Finally, it is possible that AP possessors might show no performance differences between instruments, amplitudes, or octaves, either presented in isolation (blocked setting) or mixed together (mixed setting). Indeed, this alternative seems plausible given the simple nature of the task in narrowing the focus of recognition on a specific single target note, especially if the different instruments, amplitude levels, octave, and instruments are commonly found in one's listening environment. This type of result would not support the idea that acoustic dimensions other than fundamental frequency are attended to in AP note categorization, since listeners would presumably be ignoring all perceptual variability that is not related to note chroma identification. These three hypotheses are tested in the present studies.

II. EXPERIMENT 1

On the surface, AP note categories seem to be grounded in specific instrumental experiences (e.g., Bahr *et al*., 2005). This suggests that timbral cues are meaningful sources of acoustic variability that have become tied to note representations for AP possessors through experience. In other words, an individual's expertise with a particular timbre seems to affect the strength of AP note categories, such that frequently heard timbres can be more quickly and more accurately identified than less frequently heard (or completely novel) timbres (e.g., Brammer, 1951).

As such, it is possible that even in the context of a simplified note recognition task—in which participants are monitoring a string of notes for a pre-specified target note—individual experiences with particular timbres will affect target recognition performance (e.g., pianists might show better performance for piano target notes compared to violin target notes and vice versa). This kind of finding would support the idea that timbral familiarity facilitates AP note recognition, which has been previously claimed (e.g., Bahr *et al*., 2005; Brammer, 1951).

In addition to any initial differences found between instrumental timbres as a function of experience, it is also possible that increasing the timbral variability (and thus increasing the timbral uncertainty) within each trial in the same note recognition task will impair performance. One reason to suspect this pattern of results is that in speech, listeners are faster to recognize target words among familiar talkers compared to unfamiliar talkers, though listening for a target word among two familiar talkers impairs listeners' performance in a comparable manner as listening for a target word among two unfamiliar talkers (Magnuson *et al*., 1995). This kind of result suggests that listeners need to recognize target words in the context of the talker. By extension, if timbral variability is treated similarly in AP categories as in speech categories, this suggests that AP possessors need to recognize target notes in the context of the instrument. Thus, even if participants show faster note identification for their own instruments, we would predict that increasing the variability (thus introducing uncertainty about the instrument of the target note) would impair performance—regardless of the target note's timbre—for all individuals.

A. Methods

1. Generic methods

The following methods were common to Experiments 1, 2, and 3 and hence are not repeated in further sections. All participants self-reported as AP possessors. Moreover, given the experimental paradigm, successful performance would have been extremely difficult to achieve for persons not possessing AP. Participants' musical and tonal language experience is listed in Table I. Participants were paid for their participation in the experiments.

The same general paradigm was used across all experiments. On each trial, one of four pre-specified notes served as the target note. On any given trial, the three unused targets as well as the eight remaining notes served as distractors. All notes were 500 ms in duration and were created with Reason music production software (Propellerhead, Stockholm, Sweden). All audio stimuli were digitized at a sampling rate of 44.1 kHz and a bit depth of 16 bits. Participants listened to the stimuli through Sennheiser HD280 headphones (Sennheiser, Wedemark, Germany) and the experiment was coded in E Prime 2.0 (Psychology Software Tools, Inc., Sharpsburg, PA).

TABLE I. List of music and language experience measures for participants across all experiments. TL = tonal language fluency, AMO = reported age of beginning musical instruction, PME = primary instrument musical experience (in years), PMI = primary musical instrument, SME = secondary instrument musical experience (in years), SMI = secondary musical instrument.

Participant	TL	AMO	PME	PMI	SME	SMI
Experiments 1 and 2						
1	No	8	15	Piano	0	N/A
2	No	5	16	Piano	9	Violin
3	No	2	17	Voice	11	Piano
4	No	4	15	Violin	8	Viola
5	Yes	5	8	Piano	1	Violin
6	No	6	17	Violin	8	Oboe
7	No	3	16	Piano	2	Flute
8	No	2	14	Violin	6	Piano
9	No	4	14	Piano	8	Cello
10	No	4	16	Piano	7	Flute
11	No	4	16	Piano	9	Cello
Experiment 3						
1	No	8	15	Piano	8	Horn
2	No	7	16	Piano	15	Clarinet
3	No	5	18	Piano	1	Voice
4	Yes	4	15	Piano	8	Flute
5	No	4	15	Violin	9	Viola
6	Yes	5	13	Piano	2	Clarinet
7	No	5	14	Piano	12	Violin
8	No	6	21	Violin	0	N/A
9	No	1	37	Voice	18	Drums
10	Yes	4	15	Piano	1	Harp
11	No	9	10	Viola	10	Bass
12	Yes	5	13	Piano	10	Violin

We used a speeded-target monitoring task, similar to those used to study talker variability processing in speech perception (e.g., Magnuson and Nusbaum, 2007). Participants were instructed to press the spacebar as quickly as possible whenever they heard the target note. The instructions, which occurred prior to each trial type, clearly stated whether participants would be listening for the target note in one or two instruments (Experiment 1), one or two amplitude levels (Experiment 2), or one or two octaves (Experiment 3). At the beginning of each trial, the target note was orthographically represented in the center of a computer display for 2000 ms. After 2000 ms, the screen went blank and participants then heard 16 notes in rapid succession, with 250 ms of silence between each note (making the stimulus onset asynchrony 750 ms). Four of the 16 notes were the target note, and targets were presented pseudo-randomly (never in the first or last position, and never back-to-back). The 12 distractors were randomly selected from all notes—excluding the trial's target—from 1 or 2 instruments (Experiment 1), 1 or 2 amplitude levels (Experiment 2), or 1 or 2 octaves (Experiment 3), depending on the trial type (blocked or mixed) of the experiment. Figure 1 shows a sketch of the experimental design.

The design for all experiments was completely within subject. Participants completed four runs of the experiment

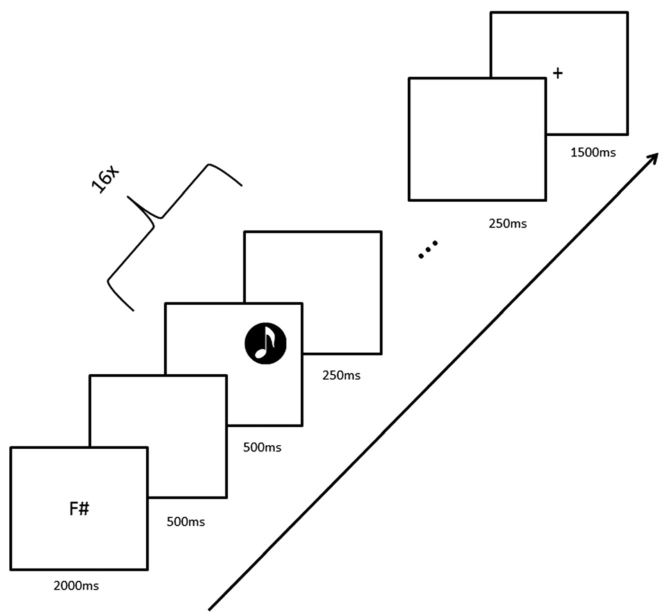

FIG. 1. Sketch of the general experimental design used across all experiments. On each trial, participants would see the target note name represented in the center of the screen for 2 s. This was followed by a blank screen (500 ms), which was then followed by a sequence (repeated 16 times) in which participants heard a note (500 ms in duration) followed by 250 ms of silence. Nothing was visually depicted on the screen during this loop. Participants had to press the spacebar as quickly as possible if the note matched the target note name at the beginning of the trial. After the loop, participants were given 1500 ms to prepare for the next trial. Each run consisted of 12 trials (of which the last 10 were analyzed), and each experiment consisted of 4 runs (2 blocked acoustic parameter runs and 2 mixed acoustic parameter runs).

(Experiment 1: one piano-only run, one-violin only run, and two mixed piano/violin instrument runs; Experiment 2: one low-amplitude run, one high-amplitude run, and two mixed low/high amplitude runs; Experiment 3: one lower octave run, one upper octave run, and two mixed lower/upper octave runs). Run order was counterbalanced across participants. Each run consisted of 12 trials, though the first 2 trials of each block were discarded to allow participants to get acclimated to the task. This left a total of 40 targets and 120 distractors per run.

We measured response time (RT) as well as target recognition accuracy (hit rates) for each run of the experiment. We only included correct responses (hits) in our RT averages. Responses to notes that were faster than 150 ms were interpreted as responses to preceding stimuli, therefore setting the correct response window at 900 ms (500 ms target stimulus + 250 ms interstimulus interval + 150 ms of the next stimulus).

2. Participants

Twelve people with AP participated in the experiment {age 20.46 ± 2.51 [mean ± standard deviation (SD)], range 18–26 yrs, 5 males}. One participant's data was omitted because the experimental script broke halfway through the experiment, leaving 11 participants in the final analyses. A description of the music and language experience of all participants can be found in Table II.

3. Stimuli and materials

The four target notes for the experiment were B, C#, E, and F#. The notes (including both target notes and distractor notes) ranged from A[3] (the A below middle C) to G#[4]. All notes were synthesized with both a piano and violin timbre. All of the stimuli were root-mean-square (rms) normalized and presented to participants at 75 dB sound pressure level (SPL).

B. Results

Overall, participants were highly accurate at the task, correctly responding to 88.3% of target notes. Given this high level of accuracy, we looked for effects of timbral variability on target note recognition within the domain of RTs to target notes.

1. Response times

First, we assessed whether participants showed an RT advantage for piano notes compared to violin notes in the blocked trials. While a particular advantage for piano notes might not be an intuitive prediction, especially since violin notes are also a complex instrumental timbre that is commonly experienced, the majority of our participants (7 of 11) listed the piano as their primary instrument, while only 3 of 11 participants listed the violin as their primary instrument. The remaining participant, who listed voice as their primary instrument, additionally listed 11 years of piano experience (and 0 yrs of violin experience). Thus, based on the piano instruction bias observed in our participants ($M = 5.6$ yrs more piano than violin instruction), in line with the possibility that participants would show faster RTs to more familiar timbres, we examined whether there was any evidence for a piano-timbre advantage. Response times for piano target notes [mean ± standard error of the mean (SEM): 419.3 ms ± 24.4 ms] did not appear to be different than RTs for violin target notes (mean ± SEM: 426.9 ms ± 27.1 ms, [$t(10) = -0.63$, $p = 0.54$]). Even when parsing the participants by their primary instrument, we failed to find any evidence that piano or violin experience was related to response speed for either piano or violin targets (see Sec. II B 2 for more details). Thus, we failed to find any evidence that our participants were initially better at identifying target notes in a piano timbre versus a violin timbre.

We then constructed a 2 × 2 repeated-measures analysis of variance (ANOVA), with instrumental timbre (piano versus violin) and trial type (blocked versus mixed timbre) as repeated factors. If AP categories need to be contextually tuned, then we would expect a main effect of trial type with the mixed timbre trials incurring slower RTs than the blocked timbre trials. Indeed a significant main effect of trial type was found: participants took an average of 423.1 ms (SEM: 25.0 ms) to respond to target notes in the blocked trials, but RTs significantly increased to 465.4 ms (SEM: 22.9 ms) when responding to the same instrumental timbres in the mixed trials [$F\ (1, 10) = 23.37$, $p = 0.001$]. There was no significant main effect of instrumental timbre [$F\ (1, 10) = 1.55$, $p > 0.24$], meaning we did not find any evidence that participants were overall faster or slower at identifying piano notes compared to violin notes, regardless of whether they were presented in isolation (blocked trials) or mixed together (mixed trials). Additionally, the interaction between instrumental timbre and trial type was not significant [$F\ (1, 10) = 0.11$, $p > 0.74$]. These results are plotted in Fig. 2.

While it is possible that the increased variability in timbre affected the response latency to target notes through a mechanism of contextual tuning, it is also possible that participants were simply able to memorize the single exemplar in the blocked condition. If participants engaged in the latter strategy, we would still expect an overall main effect of trial type (blocked versus mixed timbre). However, we would also expect a significant decrease in RT from the first to the second target in the blocked condition, presumably because participants would know the exact acoustic structure of the target and thus could use non-chroma cues in note identification (cf. Palmeri *et al.*, 1993). Additionally, we would expect to find comparable RTs for the first target across both blocked and mixed conditions, presumably because participants have not yet experienced the acoustic structure of the

TABLE II. Correlation of overall RT to target notes, RT differences between piano and violin notes, and RT differences between blocked and mixed trial target notes as a function of the AMO, total years of musical instruction (YME), and years of piano instruction subtracted from years of violin instruction (P − V). No individual difference measure was significantly correlated with any performance measure.

	AMO	YME	P − V
Overall (RT)	−0.05	0.11	−0.31
Piano vs violin (RT)	0.10	0.22	−0.43
Blocked vs mixed instrument (RT)	0.34	−0.09	0.34

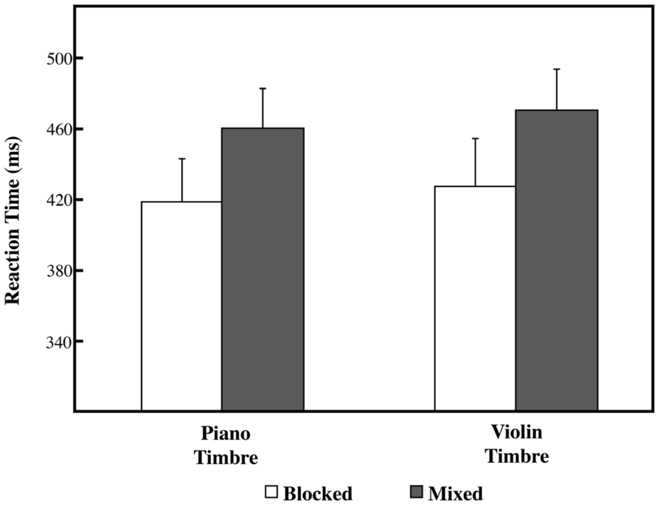

FIG. 2. Mean RTs in milliseconds for piano and violin targets when presented in blocked- and mixed-instrument trials. Error bars represent ±1 SEM.

target in order to create a holistic trace that might improve recognition speed.

We specifically tested between these two strategies of handling perceptual variability by constructing another 2×2 repeated measures (ANOVA), with target position (first versus second) and trial type (blocked versus mixed) as repeated factors. If the increased uncertainty of the target note's timbre in the mixed timbre trials was responsible for the observed RT differences, then we would expect to find a main effect of trial type (with mixed trials being slower than blocked trials), but crucially no interaction between target position and trial type (as an interaction might suggest that participants became significantly faster between the first and second targets in the blocked trials, but not in the mixed trials). Evidence of a main effect of target position could support the exemplar memorization strategy if also accompanied by a significant interaction, though a main effect without a significant interaction term would likely suggest that participants in both the blocked and mixed trials became faster from the first to the second trial.

Results of the secondary ANOVA model supported the idea that timbral uncertainty in the mixed trials—rather than holistic memorization in the blocked trials—better explains the main effect of trial type found in the original ANOVA model. Specifically, we observed a significant main effect of trial type [F (1, 10) = 11.25, $p = 0.007$], a significant main effect of target position [F (1, 10) = 12.88, $p = 0.005$], but failed to find any evidence for an interaction between trial type and target position [F (1, 10) = 0.01, $p = 0.92$]. A *post hoc* test further demonstrated that there was already a significant difference between RTs to the *first* target between the blocked (*M:* 434.8 ms, SEM: 26.2 ms) and mixed (*M:* 472.6 ms, SEM: 29.4 ms) trials [t (10) = −2.37, $p = 0.04$]. Thus, it seems unlikely that participants were simply memorizing the exact acoustic structure of targets in the blocked condition in order to respond more quickly. These results are plotted in Fig. 3.

2. Individual differences in musical instruction

Given that the instrumental timbres we tested are common primary instruments for musicians, we assessed whether

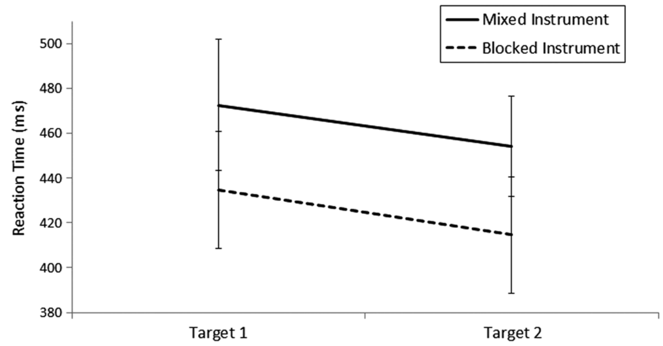

FIG. 3. Mean RTs in milliseconds to respond to the first two targets of each trial parsed by trial type (blocked- and mixed-instrument) for Experiment 1. Error bars represent ±1 SEM.

individual differences in musical instruction (specifically, piano instruction and violin instruction) were related to any performance measures, such as a speeded performance advantage for one instrument over another. All participants reported at least some experience playing the piano [*M*: 10.7 years (yrs)] or playing the violin (*M*: 5.1 yrs). Moreover, there was considerable inter-participant variability with regards to the relative amount of piano and violin experience, ranging from a 16 year advantage for piano (16 yrs of piano experience, 0 yrs of violin experience) to a 15 year advantage for violin (0 yrs of piano experience, 15 yrs of violin experience). Given that musicians show enhanced perceptual processing, even at the level of the auditory brainstem, for sounds that come from their primary instrument (Strait *et al.*, 2012), we assessed whether piano or violin experience was related to overall performance on a particular timbre, as well as target recognition performance when listening for a target within a blocked-timbre or a mixed-timbre setting. No musical measure we used [age of music onset (AMO), overall musical instruction, or instrumental expertise] seemed to be related to target recognition speed. The correlational analyses are printed in Table I.

C. Discussion

Experiment 1 was designed to explore the ways in which instrumental timbre information might affect note categorization judgment for AP possessors. We specifically tested whether the categorization of pre-specified target notes—in the context of increased variability—would impair note categorization performance, measured through RT latency. Additionally, based on prior research suggesting that AP categories are grounded in specific timbral experiences (e.g., Bahr *et al.*, 2005), we tested whether individual differences with particular instruments (specifically piano and violin) would create initial performance differences between piano and violin notes tested in isolation. We found evidence that switching between instruments—while keeping the fundamental frequency of the target note the same (i.e., the same octave)—slows target note recognition time, which was specifically seen in comparing RTs from the mixed instrument trials to the blocked instrument trials. Moreover, it is unlikely that the RT differences observed between the blocked and mixed trials can be solely attributed to memorizing the exact acoustic features of targets, as there was an initial RT difference between the blocked and mixed trials for recognition of the first target. While we did find evidence that participants became significantly faster from the first to the second target note, which can be interpreted in a musical repetition effect framework (Hutchins and Palmer, 2008), we did not find evidence for an interaction between trial type and target location, suggesting that the faster RTs from the first to the second target in each trial was comparable regardless of whether the target notes were being selected from one or two instruments.

We failed to find any support for the first hypothesis outlined in Sec. I—that initial performance differences between violin and piano notes would exist, and would be explained by individual differences in playing experiences. Indeed, individual differences in piano and violin instruction did not appear to be related to any performance differences whatsoever. In this sense, our results mark an important distinction between the previous research demonstrating performance advantages for an AP possessor's primary instrument (e.g., Brammer, 1951; Bahr *et al.*, 2005; Whipple, 1903). We are not claiming, however, that our findings are incompatible with the instrument-specific advantages found in AP possessors for their primary instruments. Perhaps the simple nature of the task (identifying a pre-specified target note) was not sensitive enough to detect inherent differences between pianists and violinists given that it was designed to maximize accuracy in order to measure differences in RT. Moreover, our operationalization of instrument experience might have been too broad. For example, specific practice estimates (e.g., number of hours) for each instrument might have provided a better approximation for instrumental expertise. Additionally, assessing the age at which participants began *each* instrument they listed (rather than reporting the age at which participants began musical instruction overall) could have been informative, as it is possible that overall instrumental experience might interact with the age at which the instrument is learned, especially given the notion of a critical period in AP (e.g., see Levitin and Rogers, 2005).

Nevertheless, the support for our second hypothesis outlined in Sec. I (that instrumental timbres provides a perceptual reference or framework by which note chroma is recognized) likely speaks to the *general* nature of increased timbre variability on note identification for AP possessors. Even in a highly simplified task, where initial differences in timbre identification do not manifest, we still find reliable evidence that increasing the timbral variability of the listening environment slows recognition time for a pre-specified target. This suggests that AP categories are contextually sensitive to variation in instrumental timbre.

III. EXPERIMENT 2

While AP listeners show poorer performance in mixed-timbre trials compared to blocked-timbre trials, it is still unclear whether an increase in any form of acoustic variability not inherently related to perception of note chroma will hurt recognition performance. From one theoretic perspective, only acoustic variability that meaningfully influences note recognition (i.e., based on prior consistent experience) should result in worse performance. Instrumental timbre has been shown to influence AP performance based on individual experiences (e.g., Bahr *et al.*, 2005), suggesting that these acoustic dimensions are important in the formation and maintenance of AP note representations. That being said, we failed to find any evidence of an initial performance difference between piano and violin timbres, even when accounting for individual experiences with these timbres. This argues that it is the acoustic variation in immediate context that is influencing recognition speed. As such, it is unclear whether *any* increase in acoustic variability (and therefore uncertainty) in context is sufficient to slow target recognition.

To make sure that the results from the first experiment could be explained in terms of learned integrality (cf. Garner, 1976) of timbre in chroma perception (and therefore that timbral variability affects recognition), we tested whether increasing the variability of a different acoustic parameter—amplitude—would engender similar effects. Since amplitude does not vary the spectral properties of a note, nor does it systematically vary with note chroma, we predicted that performance should be comparable across blocked- and mixed-amplitude trials. If, however, performance is slower in the mixed- versus blocked-amplitude trials—similar to the pattern of observed results for timbre in the first experiment—then it suggests that a simple explanation of distraction (from increasing variability along any acoustic dimension) could explain the results from the first experiment. These two hypotheses are tested in the current study.

A. Methods

1. Participants

The same participants that participated in Experiment 1 participated in Experiment 2. The presentation order was randomized, so that some participants received Experiment 1 first, while others received Experiment 2 first. There did not appear to be any significant differences between participants who received Experiment 1 versus 2 first (all $ps > 0.23$).

2. Materials

The four target notes were A#, C, F, and G. All notes (including both targets and distractors) spanned from A[3] (the A below middle C) to G#[4]. All notes were synthesized with a clarinet timbre. The reason we used a clarinet timbre is because we wanted to choose an instrument that was sufficiently different than the piano and violin (e.g., through an emphasis on odd-ratio harmonics), so as not to potentially influence the results of either Experiment 1 or 2 due to unwanted carryover effects. The low-amplitude stimuli were rms normalized and presented to participants at 70 dB SPL, while the high-amplitude stimuli were rms normalized and presented to participants at 85 dB SPL. The reason we chose these levels is because the high-amplitude stimuli could be easily distinguishable from the low-amplitude stimuli, yet both groups of stimuli were within a comfortable listening range.

B. Results

Overall, participants were highly accurate at the task, correctly responding to 89.1% of target notes. Given this high level of accuracy, we looked for effects of amplitude variability on target note recognition using RT to targets as intended.

1. Response times

Mean RTs for the low-amplitude trials (mean ± SEM: 416.1 ms ± 26.6 ms) were not significantly different than RTs for the high-amplitude trials (mean ± SEM: 417.4 ms ± 24.9 ms [t (10) = −0.11, $p > 0.9$]). This demonstrates that the difference in amplitude, on its own, does not affect recognition speed. To test whether increased acoustic variability would slow recognition performance, as was the case in Experiment 1, we carried out a 2 × 2 repeated-measures ANOVA, with target note amplitude (low- versus high-amplitude) and trial type (single- versus mixed-amplitude) as repeated factors. Unlike Experiment 1, we failed to find a main effect of trial type [F (1, 10) = 0.37, $p > 0.5$]—that is, participants were not significantly slower to respond in the mixed-amplitude trials [mean ± SEM: 421.3 ms ± 24.7 ms] compared to the single-amplitude trials [mean ± SEM: 416.8 ms ± 25.1 ms]. The interaction between amplitude and trial type was also not significant [F (1, 10) = 3.56, $p = 0.09$]. These results are plotted in Fig. 4.

Since the above finding rests on a null result, to test the idea that timbre cues are treated differently than amplitude cues, we constructed another 2 × 2 repeated-measures ANOVA model with trial type (blocked versus mixed) and acoustic cue type (timbre versus amplitude) as repeated factors. We were able to treat acoustic cue type as a within-subjects factor since the same participants who participated in Experiment 2 also participated in Experiment 1. In support of the idea that amplitude behaves differently from timbre depending on trial type, we found a significant trial type-by-acoustic cue type interaction [F (1, 10) = 14.30, $p = 0.004$].

2. Discussion

The results from Experiment 2 demonstrate that not all discriminable acoustic variability is sufficient to impair recognition of note targets. For targets in mixed-amplitude trials, recognition speed was equivalent to performance in blocked-amplitude trials. In contrast, the same participants were significantly slowed by 40 ms when listening for a target note among two instruments compared to one instrument. This significant interaction between acoustic variability type (instrument or amplitude) and trial type (single or mixed) suggests that note chroma perception is more affected by timbral changes than amplitude changes, highlighting the way in which consistent experience has structured chroma representation. As such, it is highly unlikely that the difference between the blocked-timbre trials and the mixed-timbre trials from the first experiment can be explained in terms of general uncertainty or distractibility brought on by increased acoustic variability. This suggests that spectral variability in preceding context may affect note recognition for listeners with AP rather than amplitude variability. Given that the chroma of a note is not formally specified by its spectral properties beyond the fundamental, this raises the question of whether other spectral variability within a chroma category affects recognition. To test this hypothesis we investigated octave variability that maintains chroma identity but varies spectrum.

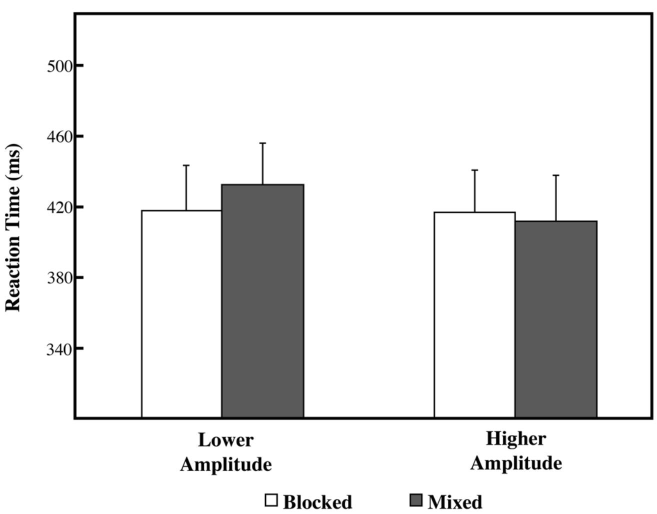

FIG. 4. Mean RTs in milliseconds for low- and high-amplitude targets when presented in blocked- and mixed-amplitude trials. Error bars represent ±1 SEM.

IV. EXPERIMENT 3

AP listeners, while accurate in identifying note chroma, are not any more accurate in identifying the octave from which a note is taken compared to non-AP possessors with comparable music experience (Lockhead and Byrd, 1981). This suggests that AP note categories are specific only to the chroma of a note. Thus, in assigning a particular pitch a musical label, the octave from which a note is selected should be irrelevant for classification, insofar as the overall frequency range is not too high (generally cited as over 5 kHz) as to degrade the overall perception of pitch (e.g., Atteneave and Olson, 1971; Ward, 1954).

On the other hand, Bahr *et al.* (2005) found evidence that AP performance in particular pitch ranges (low, middle, and high) was related to an individual's experience playing instruments in those ranges. This suggests that frequent experience with a particular pitch range—similar to frequent experience with particular instruments—might result in AP note categories that are grounded in particular registers (e.g., a tuba player might have an enhanced AP for lower notes compared to a flautist).

Given these seemingly contradictory findings (that AP possessors do not have an enhanced ability to distinguish octave register, yet show differential octave register effects for AP performance based on experience), the question becomes whether irrelevant variation in octave affects the speed of note recognition. Given that irrelevant spectral variability (in timbre) slows note recognition, if unexpected spectral variability generally affects note recognition, a similar pattern of results would be predicted for octave variability. This means that similar to instrumental timbre variability, AP listeners will show worse performance in judging a target note when presented across multiple octaves compared to a single octave—even though the target note label is not putatively dependent on any particular octave register.

Additionally, it is possible that individuals will show initial differences between recognizing target notes across different octave registers, even when blocked by single octaves. Individual experience with particular octave ranges could enhance note representations in specific pitch registers. If, however, the pre-existing differences between any octaves can completely explain any performance differences between blocked- and mixed-octave trials, then this would suggest that it is not increased variability or uncertainty in octave that impairs AP categorization.

Finally, it is possible that participants will not show any performance differences between octaves. This third possibility is likely given the fact that AP possessors are not particularly good at explicitly stating the octave range of an isolated note—even if they are remarkably good at identifying tone chroma. We explicitly test these three hypotheses in the current study.

A. Methods

1. Participants

Thirteen people participated in the experiment [age 22.6 ± 5.4 (mean ± SD), range 18–37 yrs, 6 males]. One participant was excluded from analysis due to the inability to reliably distinguish target notes from distractor notes. This left 12 participants in the final analysis. A description of the music and language experience of all participants can be found in Table II.

2. Stimuli and materials

The four target notes were C#, D, D#, and E. All notes were synthesized with a piano timbre. The lower blocked octave run spanned from A[3] (the A below middle C) to G#[4], and the upper blocked octave run spanned from A[4] to G#[5]. The two mixed octave runs were identical, spanning from A[3] to G#[5]. All of the stimuli were rms normalized and presented to participants at 75 dB SPL.

B. Results

Overall, participants were highly accurate at the task, correctly responding to 92.3% of target notes. Given this high level of accuracy, we looked for effects of octave variability on target note recognition within the domain of RTs to target notes.

1. Response times

We compared the mean RTs between the two different blocked-octave runs (lower-octave-only trials versus higher-octave only trials), to assess any target recognition differences between octaves. Unlike Experiments 1 and 2, participants were significantly slower responding to target notes in the lower octave (mean ± SEM: 409.5 ms ± 25.7 ms) compared to the higher octave (mean ± SEM: 392.7 ms ± 28.4 ms [t (11) = 2.56, p = 0.03]). To test whether increased acoustic variability would impair individuals' performance, as was the case in Experiment 1, we constructed a 2 × 2 repeated-measures ANOVA, with target octave (low versus high) and trial type (blocked- versus mixed-octave) as repeated factors. Similar to Experiment 1, participants were significantly slower to respond in the mixed octave trials [mean ± SEM: 443.6 ms ± 23.7 ms] compared to the single octave trials [mean ± SEM: 401.1 ms ± 26.9 ms, F (1, 11) = 37.4, $p < 0.001$]. In addition, target recognition was significantly slower in the lower octave (mean ± SEM: 433.4 ms ± 25.7 ms) than the higher octave [mean ± SEM: 411.4 ms ± 25.5 ms, F (1, 11) = 4.73, p = 0.05]. The interaction between octave and trial type, however, was not significant [F (1, 11) = 0.81, p = 0.39], suggesting that the difference in RTs between the blocked- and mixed-octave trials was not being solely driven by just the lower or higher octave. These results are plotted in Fig. 5.

Similar to Experiment 1, we examined whether there were any target repetition effects that might reflect use of auditory memory for the first target token. We carried out an additional 2 × 2 repeated measures ANOVA, with target position (first versus second) and trial type (blocked versus mixed) all within subject. There was a significant main effect of trial type [F (1, 11) = 25.71, $p < 0.001$], a non-significant main effect of target position [F (1, 11) = 3.87, p = 0.08], as well as a non-significant (though marginal) interaction between trial type and target position [F (1, 11) = 4.42, p = 0.06]. A *post hoc* test demonstrated that target recognition was significantly faster for the *first* target in the blocked octave (*M:* 422.1 ms, SEM: 27.6 ms) compared to the mixed octave conditions (*M:* 441.7 ms, SEM: 26.7 ms) trials [t (11) = −2.54, p = 0.03]. This indicates that even before any auditory memory for the target could be used, listeners were slowed in target recognition by octave variability. Figure 6 plots the RTs to target notes as a function of target position and trial type.

2. Individual differences in musical instruction

Similar to Experiment 1, we tested whether musical experience (particularly with the piano) was related to the

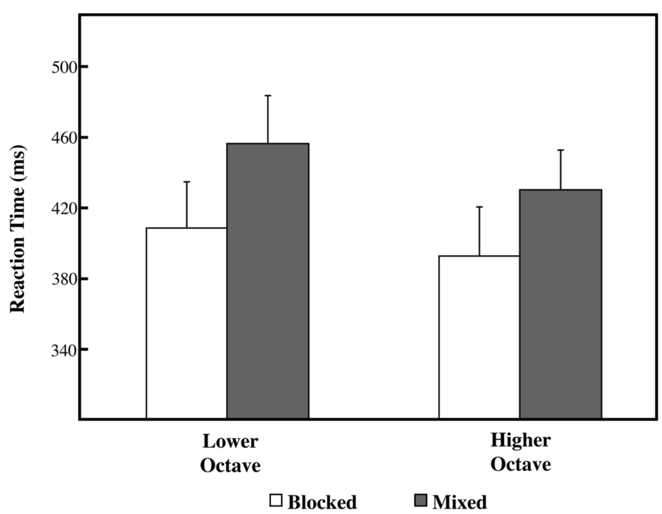

FIG. 5. Mean RTs in milliseconds for lower-octave and higher-octave targets when presented in blocked- and mixed-octave trials. Error bars represent ±1 SEM.

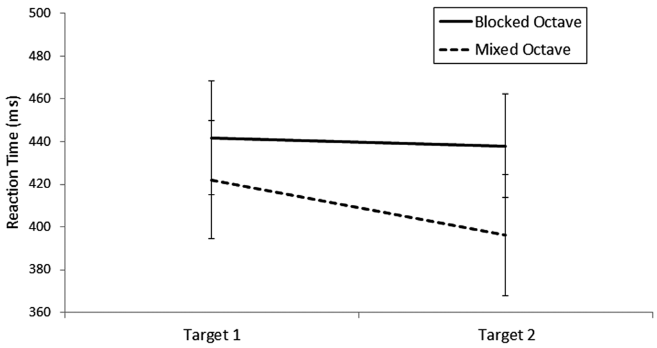

FIG. 6. Mean RTs in milliseconds to respond to the first two targets of each trial parsed by trial type (blocked- and mixed-octave) for Experiment 3. Error bars represent ±1 SEM.

magnitude of the difference between the blocked and mixed trials. None of the musical experience measures we used (AMO, overall musical instruction, or years of piano instruction) were significantly correlated with target RT, with the exception of piano experience, which accounted for the difference in RTs between the upper octave targets and lower octave targets. Specifically, since piano experience was bimodal in our population (all participants had either no piano experience or over 13 yrs of piano experience), we were able to construct an independent samples *t*-test, comparing pianists ($n = 8$) and non-pianists ($n = 4$) RT differences to the lower and upper octave. Non-pianists were 101.5 ms slower to respond to the lower octave targets compared to the upper octave targets, while pianists were just 15.2 ms slower to respond to the lower octave targets compared to the upper octave targets, which, despite the relatively small sample size, was a significant difference [$t\ (10) = 2.42$, $p = 0.04$]. Table III shows the correlation coefficients for musical instruction and target recognition speed.

C. Discussion

Experiment 3 demonstrates that AP possessors' target note identification is slower when target notes are randomly selected from two octaves. This performance difference occurs despite the fact that participants are explicitly told to expect notes from two octaves in the mixed-octave trials. The results of this study thus clearly support the conclusion that spectral variability in the immediate context of targets slows recognition of targets, even when that variability is in dimensions that are not part of the formal definition of the note chroma, which was the basis for designating the targets.

An unexpected finding from the present experiment was the initial RT difference between the lower octave trials and the higher octave trials. This finding, however, can likely be explained by frequency of experiencing particular octave ranges (cf. Bahr *et al.*, 2005). Specifically, the higher octave included commonly experienced notes, including [C5] (middle C) and thus the initial RT differences between the higher and lower octaves could have been confounded by overall listening experience, which has been shown to influence AP categories (e.g., Miyazaki, 1989). Nevertheless, the significant effect of trial type (blocked- versus mixed-octave) suggests that regardless of the initial performance differences between octaves, switching between multiple octaves incurred an additional performance reduction.

Even though all notes in the current experiment were synthesized with a piano timbre, explicit piano instruction was not related to target recognition performance, with the exception of piano instruction reducing RT differences between lower and upper octave target notes. The four participants who reported having no piano experience were all violinists, and the violin cannot play any of the lower octave target note pitches (C#, D, D#, and E below middle C) with standard tuning. The main effect of lower octave target notes being slower than upper octave target notes found for these few subjects is consistent with prior research showing that AP listeners show differential effects of their musical experience as in different note labeling performance for based on particular pitch ranges (Bahr *et al.*, 2005).

Musical instruction was not related to RT differences observed between target note recognition in a blocked octave versus a mixed octave setting. All 12 participants—regardless of individual differences in music experience—were slower to respond to a target note when in a mixed octave compared to a blocked octave setting, which is consistent with a more general effect of how spectral variability affects note chroma identification. In this sense, the present results go beyond the previous literature demonstrating that AP category strength reflects accrued experiences with particular pitch ranges (e.g., Bahr *et al.*, 2005), by providing evidence that the immediate context of prior spectral variability due to octave influences AP categorization.

D. General discussion

Across renditions of a piece of music we encounter many sources of variability that arise for a wide range of reasons. The overall pitch range, the attack and decay of a sound, the harmonic spectrum, and the dynamics of a sound, are all properties that can change based on one's listening context. Yet, for individuals with AP, any pitched sound within a certain pitch range (i.e., under 5 kHz) should be classifiable with a note label (e.g., F#), regardless of these sources of variability. Indeed, except for extreme cases, most AP possessors can accurately label notes that span a wide variety of instruments, amplitudes, and octaves.

TABLE III. Correlation of overall RT to target notes, RT differences between upper and lower octave notes, and RT differences between blocked and mixed trial target notes as a function of the AMO, YME, and number of years of piano instruction (P). The only significant relationship was between piano experience and the difference between RTs to targets from the upper octave versus targets from the lower octave. Specifically, piano experience mitigated the RT differences between the upper and lower octaves, while a lack of piano experience resulted in slower RTs to lower octave piano notes relative to upper octave piano notes. * $p < 0.05$.

	AMO	YME	P
Overall (RT)	−0.17	0.02	−0.10
Upper vs lower octave (RT)	−0.02	0.08	0.60*
Blocked vs mixed octave (RT)	0.16	0.11	−0.12

Despite this ability to accurately label notes across a variety of acoustic dimensions, a growing body of research is supporting the idea that the representations of note categories for AP possessors are not the result of some fixed, abstract note-frequency mapping, but rather are plastic and are influenced by the specific context of acoustic variability that one has experienced in the past (e.g., Bahr *et al.*, 2005) and is experiencing in the present (Hedger *et al.*, 2013). In the current set of experiments, we provide evidence for a general mechanism for why certain kinds of acoustic variability might influence AP note categories. Even when performing a simplified note recognition task, meant to promote highly accurate note categorization, participants still showed consistent evidence of context sensitivity for acoustic variability across instrumental timbre and octave, but not amplitude. This is consistent with *contextual tuning theories* found in speech (e.g., Nusbaum and Morin, 1992), and is suggestive of the contextual information that AP listeners typically attend to in resolving note chroma. From these data, differences in overall pitch and harmonic structure (in terms of octave and timbre differences) appear to be perceptually relevant to note category labeling, while amplitude is not. Ostensibly this is because octave range and instrumental timbre are fixed properties of an instrument (e.g., it is impossible to play a C[1] on a flute or to make a piano sound like a trumpet), while amplitude is not a fixed property for most instruments (e.g., a middle C on a piano is not systematically loud or soft). This theoretical interpretation of our results yields some interesting questions for future research. For instance, based on our interpretation of the present results, it is possible that AP possessors with extensive experience playing fixed amplitude instruments, such as the harpsichord, might show context sensitivity for amplitude differences within harpsichord notes. This is presumably because regardless of the gesture used to produce a sound (e.g., a soft key strike or a forceful key strike), the resulting amplitude of the produced sound would be the same, due to the mechanisms of producing a sound on the harpsichord.

Moreover, the RT differences observed between blocked and mixed instruments (Experiment 1) and octaves (Experiment 3) does not seem to be solely driven by an exact token matching strategy (cf. Palmeri *et al.*, 1993), since we also observed initial RT differences between the blocked and mixed trials for the *first* target note (i.e., before participants could adopt an exact token matching strategy). This initial RT difference for the first target note thus suggests that the expectation of encountering increased variability or experiencing variability in irrelevant categories (i.e., non-target, distractor notes) significantly impacts how quickly one will respond to targets. Future experiments could further test this claim by either independently varying the acoustic variability in target and distractor notes (e.g., keeping an identical acoustic structure for target notes but introducing variability in the distractor notes along a context sensitive dimension), or through manipulating task instructions while holding the target and distractor acoustics constant (e.g., telling participants to expect greater or lesser amounts of acoustic variability along a context sensitive dimension). The latter idea has already been tested in speech, with the finding that listener expectations of increased variability can be just as powerful as actually experiencing increased acoustic variability (Magnuson and Nusbaum, 2007).

The results from the current set of experiments support the notion that AP categories share similarities with speech categories. Speech target recognition performance is slowed when there is variability in the talker (similar to instrument) but not affected with changes in amplitude (Magnuson and Nusbaum, 2007). This ostensibly occurs because variability along perceptually relevant dimensions poses a challenge to the perceptual system. In speech, however, this kind of finding is traditionally described in terms of the lack of invariance problem (e.g., Peterson and Barney, 1952). Specifically, a single acoustic event can map onto multiple categories, and a single category is derived of multiple acoustic events. AP note identification certainly adheres to the second aspect of this description. A single category such as C can be made from an infinite number of permutations of harmonic spectrum, octave, amplitude, etc. However, generally speaking a single pitch should not belong to multiple note categories, meaning that AP might be best thought of as a "many-to-one" rather than a "many-to-many" mapping. In support of a many-to-many mapping in AP, Hedger *et al.* (2013) has provided some evidence that a single acoustic event might have more than one interpretation depending on context, by showing that identically tuned notes are interpreted as more or less "in-tune" depending on the preceding musical context.

Regardless of treating AP categories as a many-to-many mapping or a many-to-one mapping, it is clear that AP note categories are not represented in some direct acoustic-to-perceptual mapping system (cf. Gibson, 1979), in which there exists an invariant frequency-note mapping. Across three experiments, we demonstrate that timbre and octave variability impairs performance, suggesting that these acoustic dimensions represent important acoustic cues relevant for perceptually classifying a given note. This performance impairment between blocked- and mixed-acoustic trials was independent of any initial performance differences between acoustic dimensions tested in isolation. In fact, evidence that participants were initially better at particular instruments or octaves based on their musical backgrounds was mixed, yet we consistently found slower target recognition for mixed-versus blocked-acoustic trials in Experiments 1 and 3. The magnitude of the RT difference between blocked and mixed trials in Experiments 1 and 3, while modest, was nearly identical (43.3 ms for Experiment 1, 39.5 ms for Experiment 3), and similar to the RT differences previously observed between blocked and mixed talkers in speech (e.g., Magnuson and Nusbaum, 2007). Given these parallels between the effects of acoustic variability in AP note labeling and speech, we propose the idea that both may be better understood through a lens of general auditory categorization.

ACKNOWLEDGMENTS

This research was supported in part by a grant from the ONR Grant No. DoD/ONR N00014-12-1-0850 to UCSD.

Attneave, F., and Olson, R. K. (**1971**). "Pitch as a medium: A new approach to psychophysical scaling," Am. J. Psychol. **84**, 147–166.
Bachem, A. (**1937**). "Various types of absolute pitch," J. Acoust. Soc. Am. **9**, 146–151.
Baharloo, S., Service, S. K., Risch, N., Gitschier, J., and Freimer, N. B. (**2000**). "Familial aggregation of absolute pitch," Am. J. Human Genetics **67**, 755–758.
Bahr, N., Christensen, C. A., and Bahr, M. (**2005**). "Diversity of accuracy profiles for absolute pitch recognition," Psychol. Music **33**, 58–93.
Bermudez, P., and Zatorre, R. J. (**2009**). "A distribution of absolute pitch ability as revealed by computerized testing," Music Percept. **27**(2), 89–101.
Brammer, L. M. (**1951**). "Sensory cues in pitch judgment," J. Exp. Psychol. **41**, 336–340.
Deutsch, D. (**2013**). "Absolute pitch," in *The Psychology of Music*, 3rd ed., edited by D. Deutsch (Academic Press, San Diego, CA), pp. 141–182.
Garner, W. R. (**1976**). "Interaction of stimulus dimensions in concept and choice processes," Cogn. Psychol. **8**(1), 98–123.
Gerstman, L. J. (**1968**). "Classification of self-normalized vowels," IEEE Trans. Audio Electroacoustics **16**, 78–80.
Gibson, J. J. (**1979**). *The Ecological Approach to Visual Perception* (Houghton and Mifflin, Boston, MA), 332 pp.
Hedger, S. C., Heald, S. M., and Nusbaum, H. C. (**2013**). "Absolute pitch may not be so absolute," Psychol. Sci. **24**, 1496–1502.
Hutchins, S., and Palmer, C. (**2008**). "Repetition priming in music," J. Exp. Psychol.: Human Percept. Perform. **34**, 693–707.
Levitin, D. J., and Rogers, S. E. (**2005**). "Absolute pitch: Perception, coding, and controversies," Trends Cognit. Sci. **9**, 26–33.
Lockhead, G. R., and Byrd, R. (**1981**). "Practically perfect pitch," J. Acoust. Soc. Am. **70**, 387–389.
Magnuson, J. S., and Nusbaum, H. C. (**2007**). "Acoustic differences, listener expectations, and the perceptual accommodation of talker variability," J. Exp. Psychol.: Human Percept. Perform. **33**, 391–409.
Magnuson, J. S., Yamada, R. A., and Nusbaum, H. C. (**1995**). "The effects of talker variability and familiarity on mora perception and talker identification," ATR Human Information Processing Research Laboratories Technical Report TR-H-158.
Miyazaki, K. I. (**1989**). "Absolute pitch identification: Effects of timbre and pitch region," Music Percept. **7**(1), 1–14.
Miyazaki, K. (**1990**). "The speed of musical pitch identification by absolute pitch possessors," Music Percept. **8**, 177–188.
Mull H. K. (**1925**). "The acquisition of absolute pitch," Am. J. Psychol. **36**, 469–493.
Neu, D. M. (**1947**). "A critical review on the literature on 'absolute pitch,'" Psychol. Bull. **44**, 249–266.
Nusbaum, H. C., and Morin, T. M. (**1992**). "Paying attention to differences among talkers," in *Speech Perception, Speech Production, and Linguistic Structure*, edited by Y. Tohkura, Y. Sagisaka, and E. Vatikiotis-Bateson (IOS Press, Burke, VA), pp. 113–134.
Palmeri, T. J., Goldinger, S. D., and Pisoni, D. B. (**1993**). "Episodic encoding of voice attributes and recognition memory for spoken works," J. Exp. Psychol.: Learning, Memory, Cognit. **19**, 309–328.
Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of vowels," J. Acoust. Soc. Am. **24**, 175–184.
Ross, D. A., Gore, J. C., and Marks, L. E. (**2005**). "Absolute pitch: Music and beyond," Epilepsy Behavior **7**, 578–601.
Strait, D. L., Chan, K., Ashley, R., and Kraus, N. (**2012**). "Specialization among the specialized: Auditory brainstem function is tuned in to timbre," Cortex **48**(3), 360–362.
Takeuchi, A. H., and Hulse, S. H. (**1993**). "Absolute pitch," Psycholog. Bull. **113**(2), 345.
Ueda, K., and Ohgushi, K. (**1987**). "Perceptual components of pitch: Spatial representation using a multidimensional scaling technique," J. Acoust. Soc. Am. **82**, 1193–1200.
Ward, W. D. (**1954**). "Subjective musical pitch," J. Acoust. Soc. Am. **26**, 369–380.
Ward, W. D., and Burns, E. M. (**1982**). "Absolute pitch," in *The Psychology of Music*, 1st ed., edited by D. Deutsch (Academic Press, San Diego, CA), pp. 431–451.
Whipple, G. M. (**1903**). "Studies in pitch discrimination," Am. J. Psychol. **14**, 289–309.