ARTICLE

Diversity of accuracy profiles for absolute pitch recognition

Psychology of Music

Psychology of Music

vol 33(1): 58–93 [0305-7356
(200501) 33:1; 58–93]
10.1177/0305735605048014
www.sagepublications.com

NAN BAHR, CAROL A. CHRISTENSEN
AND MARK BAHR
UNIVERSITY OF QUEENSLAND

ABSTRACT Most advanced musicians are able to identify and label a heard pitch if given an opportunity to compare it to a known reference note. This is called 'relative pitch' (RP). A much rarer skill is the ability to identify and label a heard pitch without the need for a reference. This is colloquially referred to as 'perfect pitch', but appears in the academic literature as 'absolute pitch' (AP). AP is considered by many as a remarkable skill. As people do not seem able to develop it intentionally, it is generally regarded as innate. It is often seen as a unitary skill and that a set of identifiable criteria can distinguish those who possess the skill from those who do not. However, few studies have interrogated these notions. The present study developed and applied an interactive computer program to map pitch-labelling responses to various tonal stimuli without a known reference tone available to participants. This approach enabled the identification of the elements of sound that impacted on AP. Pitch-labelling responses of 14 participants with AP were recorded for their accuracy. Each participant's response to the stimuli was unique. Their accuracy of labelling varied across dimensions such as timbre, range and tonality. The diversity of performance between individuals appeared to reflect their personal musical experience histories.

KEYWORDS: *aural development, aural perception, perfect pitch*

Absolute pitch (AP) is the ability to label a heard tone without the use of comparison to a known referent. This ability is rare, demonstrated by less than one hundredth of 1 percent of the general population (e.g. Bachem, 1955; Bloom, 1982; Dowling, 1999). Not only is AP rare in the general population, the ability is also scarce even amongst expert musicians (Dowling, 1999). In a recent review, Ward (1999) reported that AP was highly esteemed by musicians, who often regarded it as the pinnacle of musicianship.

Stumpf first noted the existence of AP in his review of the 'talents' of great composers in 1883 (Ward and Burns, 1982). It was seen as a genetically determined skill that indicated musical giftedness (Ward and Burns, 1982). This view of AP has been enduring so that Levitin (1994) and Ward (1999) described the ability as somewhat mysterious and Costall (1985) suggested that it is often regarded almost mystically.

Innateness of AP has been fairly constantly represented in the literature over the past 100 years (Profita and Bidder, 1988; Revesz, 1953; Stark, 1985; Torff and Winner, 1994; Vernon, 1977). In particular, work by Bachem (1937, 1940, 1954) strongly promoted this view. He had a pronounced influence because he combined heredity arguments with a proposal for a classification framework for AP, so that his work became an anchor for further research in the field.

The view that those who demonstrate AP are born with the ability, and that AP can be traced back through their family (Ward, 1999) has been particularly influential in shaping research on the skill. For example, genealogical research has attempted to establish the type of genetic relationships involved in its inheritance (Revesz, 1953). However, much of the evidence used to support the innateness of AP has been anecdotal (Pruett, 1987; Vernon, 1977) and has relied heavily on autobiographical recollections or the retelling of family legends (Vernon, 1977).

The argument for the heritability of AP centres on four key issues. AP is:

- present very early in life (before the age of five) (Grebelnik, 1984);
- acquired very quickly (Burns and Campbell, 1994);
- acquired without effort (Takeuchi and Hulse, 1993); and
- runs in families (Revesz, 1953).

Until very recently, a common assumption of the literature on AP was that it could be described dichotomously. That is, an individual either possessed AP or did not. There were no shades of grey in between. This suggested that people could either demonstrate a complete and unerring ability to label tones without a referent, or they could name none (Brady, 1970; Profita and Bidder, 1988; Stark, 1985; Terhardt and Seewann, 1983). Thus, the entire population could be divided into two neat groups: those innately blessed with AP and those without it. Key researchers have argued for classification groups within AP possessors. Bachem, for example, used the terms 'genuine absolute pitch', 'quasi-absolute pitch', and 'pseudo-absolute pitch' to describe AP recognition that was in between the extremes (Bachem, 1937). However, his discussion of these classes sets 'genuine AP' apart from the others, as the only real or true AP. Baharloo et al. (1998) also divide AP possessors into subclasses, but identify one group, the 'AP-1', as being true possessors.

There is increasing evidence that AP may reflect early musical training and experience between the ages of 3 and 6 (Baharloo et al., 1998; Baharloo et al., 2000; Cohen and Baird, 1990; Lenhoff et al., 2001; Sergeant, 1969;

Sergeant and Roche, 1973; Takeuchi and Hulse, 1993). While these researchers have challenged the notion that AP is innate, they have retained the notion that there is a clear distinction between individuals with AP and those without. Thus, they have assumed a set of clear sorting criteria could be applied to individuals to distinguish those who possess AP. Thus, researchers have generally assumed the existence of an arbitrary performance 'dividing line' between those who have AP skill and those who have not.

However, this assumption has not withstood close scrutiny (Takeuchi and Hulse, 1993). Researchers have been unable to define explicit performance criteria to mark the boundary between AP and non-AP identification skills. Neither tone range, note colour (e.g. black or white), timbre or speed and accuracy of response provided a manner of effectively dividing the population into neat categories of those with AP and those without. It could be argued that the very existence of disagreement on the defining point for AP challenged the notion that there was a clear division between AP and non-AP (Oakes, 1951).

As a result of problems in defining specific criteria for AP, research often demonstrated a kind of 'sliding scale' in the dividing point between what can be accepted as AP skill and what should be considered non-AP (Jeffress, 1962). Thus, rather than suggesting a dichotomous skill, work identifying performance boundaries for AP indicated a graduated or possibly a continuum of skill, where the defining point between those who have and have not is indistinct (Takeuchi and Hulse, 1993; Terhardt and Seewann, 1983).

Researchers have generally argued that the difficulty in defining performance boundaries indicated imperfections in the measurement parameters rather than an indictment of a dichotomous assumption (Jeffress, 1962). The dichotomous approach is a key element in the argument for heredity of AP. The argument depends on the notion that the tendency for people to think in categories about sound (i.e. naming tone types) is fundamentally different and oppositional to the tendency to think relationally about sounds. That is, the two approaches are mutually exclusive of each other. If there is no gradation of pitch-labelling skill blurring the boundaries between those with AP and those without, then there is a stronger case for predisposition (i.e. hereditary) foundations to the nature of pitch perception.

The attractiveness and simplicity of this argument has diverted research that may have contradicted the assumption of dichotomy. The view of AP as innate and the assumption of a dichotomy in the distribution of AP led to the argument that AP was a unitary/unidimensional skill (e.g. Bachem, 1955; Cohen and Baird, 1990). This means that AP could be exercised in a way that was unaffected by factors such as the pitch or timbre of the note being identified. However, the robustness of AP under different conditions has not been supported by empirical research. There is substantial diversity in the labelling accuracy, consistency, speed of labelling response following a heard tone and extent of the labelling ability for people described as having AP skill (e.g.

Miyazaki, 1990, 1992, 1993, 1995). If AP was to be considered a form of acquired knowledge, then the unidimensional view would support acquisition grounded in the simple accumulation of frequency and label facts into uncomplicated schemata organized around the 12 fundamental tone labels. As such, AP performance would vary little according to manipulations of sound attributes, provided the frequencies remained clearly discernible. However, recent research has shown that AP performance does vary when attributes of aural stimuli are manipulated even when the frequency remains clearly discernible (Miyazaki, 1990, 1992, 1993, 1995; Takeuchi and Hulse, 1993; Ward, 1999).

Takeuchi and Hulse (1993) suggested AP is multidimensional, where accuracy and latency of response to a tonal stimulus vary along several parallel dimensions. For example, Miyazaki, (1990, 1992, 1993, 1995) described the attributes of AP knowledge that may be shared between individuals. His findings indicated that AP might be affected by various attributes of the sound stimuli, such as range and timbre.

Thus, Takeuchi and Hulse (1993) argued that AP is a continuous and multidimensional skill. They suggested that pitch discrimination skills vary on several dimensions, including timbre and pitch range, and that an individual's response might vary in accuracy and latency of response across these dimensions. They suggested that individual performance might be represented as a profile of AP skill. Profiles would document an individual's pitch discrimination knowledge based on attributes such as aural acuity, timbre, pitch range and temperament and accuracy. For example, an individual might be able to identify tones quickly and accurately, but only if they are of a specific timbre, or they might be able to identify tones in any timbre but only within a specific frequency range.

This study initially investigated the nature and structure of knowledge of pitch of participants with AP. This was examined by an analysis of accuracy of responses by participants with AP. Second, the study evaluated the role of familiarity with tonal stimuli in the nature and structure of knowledge for AP to examine the extent to which AP was a learned skill.

Method

PARTICIPANTS

Thirteen participants had their pitch labelling skill profiled by the AP Profiler program (see later). There were 11 AP possessors, and two RP possessors. There were eight female and three male AP participants aged from 17 to 60 years. The RP participants were both female and 37 years old. All participants were advanced and active musicians. Musicianship was defined as long-standing (> 10 years) and continuing commitment to high-demand performance. All were volunteers. All participants were residents of Queensland, Australia, and were recruited through general advertisement in magazines and circulars that are popular with the musical community of South East Queensland, or by direct approach.

Occupations for the AP participants varied: two classroom music teachers (one of whom was recently retired), three instrumental music teachers, two music students, a freelance musician, a dentistry student, a law student and a psychologist. The RP participants were both classroom music teachers.

MATERIALS

Interactive software, AP Profiler, was developed to provide comprehensive accuracy profiles of participants' labelling skill. The software mapped the accuracy of response to tonal stimuli. It played three banks of sounds in separate timbres, one note at a time. Participants were required to identify the note. Their note identification accuracy was recorded. The stimuli comprised: piano tones, sine wave tones and tones of each participant's own instrument selected from the 166 Wavetable options. All tones were delivered through the binaurally controlled stereo headphones. Responses were made on the labelled MIDI controller keyboard. Each set of stimuli included 60 notes, each heard twice (i.e. 120 tone presentations) of the standard chromatic scale over five octaves in equal temperament. The five octaves were necessary to ensure participants were tested over a broad range of stimuli to investigate the effect of pitch height on AP salience. The controller keyboard had the required five octaves of keys to facilitate the recording of octave choice along with pitch label.

There were three complete sets of tones in separate tone sets of the profiler program, one for each timbre type (i.e. a total of 360 tone presentations with six hearings per discrete pitch). Each tone set of the profiler consisted of a different random sequence of tones. Tones were heard for 0.8 seconds' duration and were separated by .5 seconds of silence and .5 seconds of white noise. The white noise acted to erase temporal memory of the previous heard tone.

A Visual Basic program was used to generate the three randomized tone sets for the stimuli. Visual Basic randomization functions were used to generate a rectangular distribution of tones, sampled without replacement. Two parallel tone sets were included as stimuli in each tone set of the study. The tone set consisted of the MIDI code for a particular note and its note name in ASCII. Participants were shown a progress bar on their computer screen which indicated how far through each tone set they had progressed at any given time. Each time the music keyboard was pressed a purple 'note' icon appeared at the base of the screen as the only confirmation of a key press. Participants did not hear an aural feedback to any key press to preclude their use of RP for tone labelling.

The order of the tone set presentation to participants was random and at the completion of all three tone sets, the Microsoft Excel radar charts generated from all three tone sets were displayed to the participant. Profiles of pitch-labelling responses by AP participants mapped their accuracy for each note across variations in colour, timbre, and range.

PROCEDURE

Initially participants were given a brief orientation to the equipment. The correct wearing of headphones, use of the control panel for personal binaural volume control, musical keyboard layout and location of the initial stimuli response device were explained. Participants were told that the program was in three parts, that they could have a short break between each section and that each tone set would take around 10 minutes to complete. They were shown an example of the type of chart that would be automatically produced for them. The example chart was generated by a relative pitch (RP) volunteer and so did not show consistent successful response to the stimuli; this was to ensure that participants did not feel unduly challenged for their own performance. If the participant had no questions, the profiling program was initiated for them.

Results

ACCURACY PROFILES

To give as full a picture as possible of the dimensions of each participant's pitch-labelling skills, performance was evaluated for:

1. correct responses;
2. responses with only octave (8va) discrepancies (either over- or under-judging the pitch by exactly an octave); and
3. responses within a semitone of the target tone.

Participants had a 15:60 or 25 percent possibility of accurately identifying the tone by chance. Comparisons of performance to chance levels were also made for responses including octave discrepancy and for responses including both octave and semitone discrepancy, or overall accuracy. Chance level for selecting the correct note in its correct octave was 1:60(1.7%). Chance level for correct responses allowing for the inclusion of over- or under-judgments by an exact octave was generally 5:60(8.3%).

In the high or low registers, the chance level for selecting the correct response including octave discrepancies was 3:60(5%), because it was beyond the range of the response keyboard to make similar over- or under-judgments as in the middle register. The chance level for correct responses including semitone and octave discrepancies, all allowable error, was 15:6(25%). This was the measure used to determine if AP skill could be accepted as having been demonstrated, as it represented all responses within a semitone of the absolutely correct note label.

Each participant's performance was analysed to consider the effects of timbre, range and note type on accuracy levels. This was only appropriate for those participants who performed significantly better than chance in their overall accuracy score, that is, for those participants accepted as having demonstrated AP. All participants who claimed AP performed significantly better than chance for their overall accuracy. In contrast RP participants included for validation of the protocol did not perform significantly better than chance (see Table 1).

Pitch-labelling performance varied for each individual participant. Many participants showed favour for particular attributes of stimulus tones. Responses within a semitone of the correct label were included in the overall accuracy score for each participant.

The difference between the overall accuracy of each participant's responses and chance was tested using a z test of the difference between two proportions.

TABLE 1. *Overall number of correct responses for all participants after three tone sets of the AP Profiler program*

	Correct	8va discrepancy	Overall accuracy (inc 8va and semi discrepancy)	z
Andrew	138	145	243	18.6*
Anthony(/240)	67	75	169	16.3*
Cathy	201	263	355	32.3*
Helen	82	116	254	19.9*
Jean	200	217	353	32.0*
Julie	194	210	243	18.6*
Kerri	261	338	356	32.1*
Leo	344	356	358	32.6*
Lesley	52	85	193	12.5*
Sonja	119	176	323	28.4*
Tina	169	296	354	32.1*
💣Clare	7	30	99	1.09
💣Margaret	22	24	86	–.05

*p = .05; 💣 RP participants.
Number of responses = 360 for all participants except Anthony, where N = 240 (Anthony heard only two tone sets).

Effects of range

Table 2 shows the differences in labelling performance shown by participants as affected by register overall. Some participants appear to make octave errors, but not uniformly across the range sets. For example, Kerri made significantly more errors in the low range than the other two ranges, and Lesley made significantly fewer errors in the mid range than in the other two ranges. This is especially noticeable in the extreme low or high registers. The nature of this effect varied by participant. Leo and Andrew tended not to make octave misjudgments and so a range effect was not apparent. The high register included the top 20 tones (E^4–B^6), the middle register, the centre 20 tones

TABLE 2 *Overall number of correct responses for the high, middle and low registers for all tone sets of the AP Profiler program*

	Range effects															
	Correct				8va discrepancy				Semitone discrepancy				Correct including all allowable discrepancy			
	High	Mid	Low	$\chi^2(2)$	High	Mid	Low	$\chi^2(2)$	High	Mid	Low	$\chi^2(2)$	High	Mid	Low	$\chi^2(2)$
Andrew	45	43	50	0.57	0	5	1	N/A	37	30	29	1.18	82	78	82	0.13
Anthony	16	34	17	9.16**	5	4	2	1.27	31	25	34	1.40	52	63	53	1.32
Cathy	57	82	61	5.41	32	1	29	28.3***	27	36	30	1.35	119	119	120	0.01
Helen	33	35	14	9.82**	17	12	4	7.81*	37	47	55	3.51	87	94	73	2.70
Jean	62	79	59	3.49	14	1	2	18.5***	42	37	53	3.04	118	117	114	0.07
Julie	63	77	52	4.91	13	8	4	4.88	0	1	2	0.33	76	86	58	5.49
Kerri	110	119	32	52.62**	7	0	70	51.5***	3	1	14	16.3***	120	120	116	0.09
Leo	109	119	116	N/A	9	1	2	N/A	0	0	2	N/A	118	120	120	N/A
Lesley	19	25	9	7.39*	15	4	14	6.73*	42	39	37	0.32	76	68	60	1.89
Sonja	34	57	28	12.53**	18	8	30	13.00**	57	47	45	1.66	109	112	103	0.39
Tina	71	58	40	8.60*	29	38	59	11.28**	18	23	17	1.06	118	118	116	0.02

*$p < .05$; **$p < .01$; ***$p < .001$.
$N = 120$ for each range for all participants except Anthony, where $N = 80$ for each range.

(G#^{2}–D#^{4}), and the low register included the lowest 20 tones of the stimulus set (C^1-G^2) (i.e. A4 = 440Hz).

Effects of note colour

Table 3 shows the effect of note colour (i.e. black vs white notes) in each range. The table is presented in three parts, one for each of the high, middle and low ranges. Note that colour is an attribute of tone labelling that derives from the convention of colouring keys black or white on a keyboard. This convention extends to general tone labelling, with black notes being the sharps or flats and featuring less frequently in music performance than white notes. The label for a white note consists simply of its alphabetic label (A, B, C etc). The label for a black note is derived from its nearest white note and is identified as being either a semitone lower than (i.e. A flat, A♭) or a semitone higher (i.e. A sharp, A#) than a neighbouring white note.

Many participants were comparatively unaffected by note colour differences over range. That is, a stimulus note that was a sharp or flat (i.e. G#) was no more or less accurately identified than the white notes. For the most part, any significant differences shown by participants were in the number of responses with semitone discrepancy. Those participants who made a response with semitone discrepancy for a black stimulus tone rounded their responses down or up to a white tone. Percentages were used for this table due to the unbalanced number of black and white notes in each register (low: black = 48, white = 72; mid: black = 54, white = 66; high: black = 48, white = 72).

TABLE 3(a) *Percentage of correct responses for each participant by note colour in the high range (E^3–B^4)*

	Colour × high range (%)								
	Correct			Octave discrepancy			Semitone discrepancy		
	Black	White	$\chi^2(1)$	Black	White	$\chi^2(1)$	Black	White	$\chi^2(1)$
Andrew	41.6	34.7	0.64	0	0	N/A	27.08	33.3	0.60
Anthony	46.9	2.1	41.32***	12.5	2.1	8.06**	3.1	62.5	54.55***
Cathy	47.9	48.6	0.01	27.1	26.4	0.00	25	22.2	0.19
Helen	33.3	25	1.37	18.7	12.5	1.58	25	36.1	1.98
Jean	45.8	55.5	0.98	8.3	13.9	1.64	41.7	30.5	1.66
Julie	54.2	51.4	0.15	14.6	8.3	2.13	0	0	N/A
Kerri	93.7	90.3	0.09	4.2	6.9	0.33	2.1	2.8	.00
Leo	95.8	87.5	0.35	0.02	0.11	N/A	0	0	N/A
Lesley	4.2	23.6	12.44***	2.1	19.4	14.72***	56.3	20.8	16.62***
Sonja	31.3	26.4	0.62	10.4	18.1	2.29	52.1	44.4	0.51
Tina	66.7	54.2	1.39	22.9	25	0.08	10.4	18.1	1.69

**p* < .05; ***p* < .01; ****p* < .001.
Percentages are used because of the unequal number of black (#/♭) and white stimulus tones in the range (white = 72, black = 48; Anthony: white = 48, black = 32).

TABLE 3(b) *Percentage of correct responses for each participant by note colour in the mid range ($G\#^1$–$D\#^3$)*

	Colour × mid range (%)								
	Correct			Octave discrepancy			Semitone discrepancy		
	Black	White	$\chi^2(1)$	Black	White	$\chi^2(1)$	Black	White	$\chi^2(1)$
Andrew	44.4	28.8	3.08	1.8	6.1	2.00	20.4	28.8	1.28
Anthony	72.2	18.2	32.40***	5.5	4.5	0.09	11.1	47.7	22.34***
Cathy	72.2	65.1	0.36	0	1.5	N/A	31.8	31.2	0.62
Helen	27.8	28.8	0.02	16.7	6.1	5.26*	29.6	45.4	3.85*
Jean	66.7	65.1	0.03	1.8	0	N/A	27.8	33.3	0.58
Julie	68.5	60.1	0.5	7.4	6.1	0.28	0	1.5	N/A
Kerri	98.1	100	0.02	0	0	N/A	1.8	0	N/A
Leo	100	98.5	0.02	0	0.02	N/A	0	0	N/A
Lesley	14.8	25.7	2.95	0	6.1	N/A	44.4	22.7	7.12**
Sonja	57.4	39.4	2.98	9.2	4.5	1.92	25.9	50	7.57**
Tina	61.1	37.9	5.34*	35.2	28.8	0.75	1.8	33.3	28.4***

*$p < .05$; **$p < .01$; ***$p < .001$.
Percentages are used because of the unequal number of black (#/♭) and white stimulus tones in the range (white = 66, black = 54; Anthony: white = 44, black = 36).

TABLE 3(c) *Percentage of correct responses for each participant by note colour in the low range (C^0–G^1)*

	Colour × low range (%)								
	Correct			Octave discrepancy			Semitone discrepancy		
	Black	White	$\chi^2(1)$	Black	White	$\chi^2(1)$	Black	White	$\chi^2(1)$
Andrew	41.7	41.7	0.01	4.2	0	N/A	22.9	25	0.08
Anthony	37.5	10.4	16.33***	3.1	2.1	0.00	25	54.2	10.65**
Cathy	47.9	52.8	0.25	20.8	26.4	0.53	31.2	20.8	1.92
Helen	10.4	13.9	0.36	4.2	2.8	0.14	50	44.4	0.26
Jean	45.8	51.4	0.26	0	2.8	N/A	52.1	38.9	1.86
Julie	56.3	36.1	4.74*	2.1	4.2	1.28	0	2.8	N/A
Kerri	22.9	29.2	0.92	60.4	56.9	0.08	12.5	11.1	0.17
Leo	97.9	95.8	0.02	0.02	0.01	N/A	0	0.03	N/A
Lesley	2.1	11.1	6.23	2.1	18.1	12.80***	35.4	27.8	0.78
Sonja	16.7	27.8	2.69	31.2	20.8	1.92	35.4	38.9	0.22
Tina	37.5	30.6	0.94	45.8	51.4	0.37	12.5	15.3	0.57

*$p < .05$; **$p < .01$; ***$p < .001$.
Percentages are used because of the unequal number of black (#/♭) and white stimulus tones in the range (white = 72, black = 48; Anthony: white = 48, black = 32).

Effects of timbre

A further series of z tests of difference between proportions were conducted to investigate the impact of note colour on participants' accuracy. Note colour appeared to affect the responses for some participants in some but not all timbres of the stimulus set. The timbres included in the stimulus set were piano, sine wave and the timbre of the participant's own instrument as declared at the start of the AP profiling program. Table 4 shows the percentage of correct responses made by each participant for black and white tones in each timbre.

Again, the effects on performance varied between individuals. Some were significantly more accurate with white tones than with black in specific timbres (e.g. Lesley), and some participants were more accurate with black notes rather than white in the same or other timbres (e.g. Tina, Leo and Andrew).

Chi squared analyses were conducted to compare accuracy of participants' responses for each of the twelve notes collapsed across octaves (see Table 5). This showed the wide diversity amongst the participants with respect to the particular types of errors they made. For example, some participants were prone to semitone errors rather than octave errors (e.g. Anthony and Helen).

RP VALIDATION TRIAL PROFILES

For comparison, RP participants were also interviewed after their responses to the stimulus set. Their comments reinforced that the activity was not suited to the use of RP for note labelling. Below is an example.

Clare

The accuracy of Clare's responses did not vary significantly from chance and so she could not be accepted as having demonstrated AP skill. Clare did make comments after the profiling trials that showed she was attempting to use RP strategies for labelling stimuli, and that the program, as designed, did not allow her to be effective in the use of her strategies. She was in fact quite distressed at her inability to perform:

> I found it just incredibly frustrating because firstly I couldn't, I didn't have a starting pitch. Whereas I could hear a lot of intervals between the notes each time the pitch changed, yeah so I could hear every single one, major third, minor third, perfect fourth, perfect fifth, but I didn't have a reference point to start with, so that I found really frustrating . . .
>
> If you would have checked the three phases I reckon I would have got worse as time went on because by the end of it I was just completely desperate }
>
> . . . but it's also frustrating not being able to hear the sound that you play. You are really just looking for that reference point . . . that's frustrating because I'm used to identifying pitches relatively within a tonal system . . . it was very disorienting. It was like going through a bit of a, it was like leaving gravity, not

TABLE 4 *Percentage of correct responses for participants by note colour (black/white) in the different timbres of the AP profiling program*

	Black notes vs white × timbre											
	Overall			Piano timbre			Sine timbre			Own instrument's timbre		
	Black	White	z	Black	White	z	Black	White	z	Black	White	z
Andrew	42.6	35.2	1.42	34	31.4	0.30	38	35.7	0.26	56	38.6	1.89*
Anthony	51	11.4	3.84*	50	11.4	4.70*	52	11.4	4.89*	N/A	N/A	N/A
Cathy	56.7	55.2	0.28	48	38.5	1.04	70	68.6	0.16	52	58.6	0.72
Helen	23.3	22.4	0.20	26	24.3	0.21	20	18.6	0.19	24	24.3	–0.04
Jean	53.3	57.6	–0.81	58	45.7	1.33	46	60	–1.52	56	65.7	–1.07
Julie	60	49.5	1.99*	68	67	0.12	28	15.7	1.66	84	65.7	2.24*
Kerri	72.7	72.4	.063	76	72.8	0.39	70	71.4	–0.16	72	72.8	–0.09
Leo	98	93.8	1.87	98	100	1.18	96	91.4	0.99	100	90	2.30 *
Lesley	7.3	19.5	–3.20 *	4	20	–2.54*	6	18.6	–2.01*	12	20	–1.17
Sonja	36	30.9	1.01	28	18.6	1.22	46	34.3	1.29	34	40	–0.67
Tina	55.3	40.9	2.69*	40	32.8	0.81	64	38.6	2.75*	62	51.4	1.15

*$p < .05$; **$p < .01$; ***$p < .001$.

Percentages are used because of the unequal number of black (#/♭) and white notes in the timbre sets (black = 50, white = 70 in each timbre set; black = 150, white = 210 overall for all except Anthony where black = 100, white = 140).

TABLE 5 *Chi squared (χ^2) analysis results of comparisons of responses by note for each participant*

	Comparison by note ($\chi^2(11)$)			
	Correct	8va discrepancy	Semitone discrepancy	Correct including all allowable discrepancy
Andrew	7.74	N/A	4.04	13.76
Anthony	31.03***[b]	13.47*[a]	48.61***[e]	9.28
Cathy	29.29**	19.23	22.17*	0.92
Helen	23.66*	18.41*[e]	32.66***	46.02***
Jean	123.32*	2.00[a]	58.36***	1.23
Julie	4.68	6.31	12.85[e]	3.81
Kerri	10.15	6.06	9.00[d]	0.34
Leo	1.01	N/A	N/A	0.15
Lesley	33.38**	9.82[c]	46.79***	29.01*
Sonja	86.64**	10.35[e]	40.5***	4.30
Tina	38.77**	64.09***	51.93***	0.45

*$p < .05$; **$p < .01$; ***$p < .001$.
[a]d.f. = 6; [b]d.f. = 7; [c]d.f. = 8; [d]d.f. = 9; [e]d.f. = 10.
Figures 3.25–3.31 illustrate specific note comparisons for each significant outcome marked in this table. Obtained χ^2 with 11 d.f. are given for all cells except where indicated.

> having any gravity. There's no place within that pitch change to touch base because you didn't have a reference point, and you've got no finishing thing. So it's like doing something without a map, with no North.

AP PROFILES

Many participants demonstrated a preference for one or a small subset of specific notes over other specific notes for complete accuracy, for over- or under-judging by exactly an octave or for semitone discrepancy.

Example graphs of all the significant note effects are included for Jean and Tina (Figures 1 and 2). The scores seemed to indicate that note effects varied among the participants. The effects were not simply confined to participants with the lowest overall performance accuracy. Jean, for example, had an overall accuracy of 353 out of a possible 360(98%) (refer to Table 5) and her performance accuracy was clearly and consistently different depending on the actual stimulus notes. Andrew, by comparison, had a lower overall accuracy, 243 out of a possible 360(67.5%), but was largely consistent in his performance accuracy for all notes.

Table 6 presents the performance for each participant by timbre. As for the results of the note colour response comparison, it is clear that the timbre of the stimulus tone influenced the ability of some participants to note label (e.g. Tina, Sonja, Cathy and Lesley). This influence varied in magnitude and

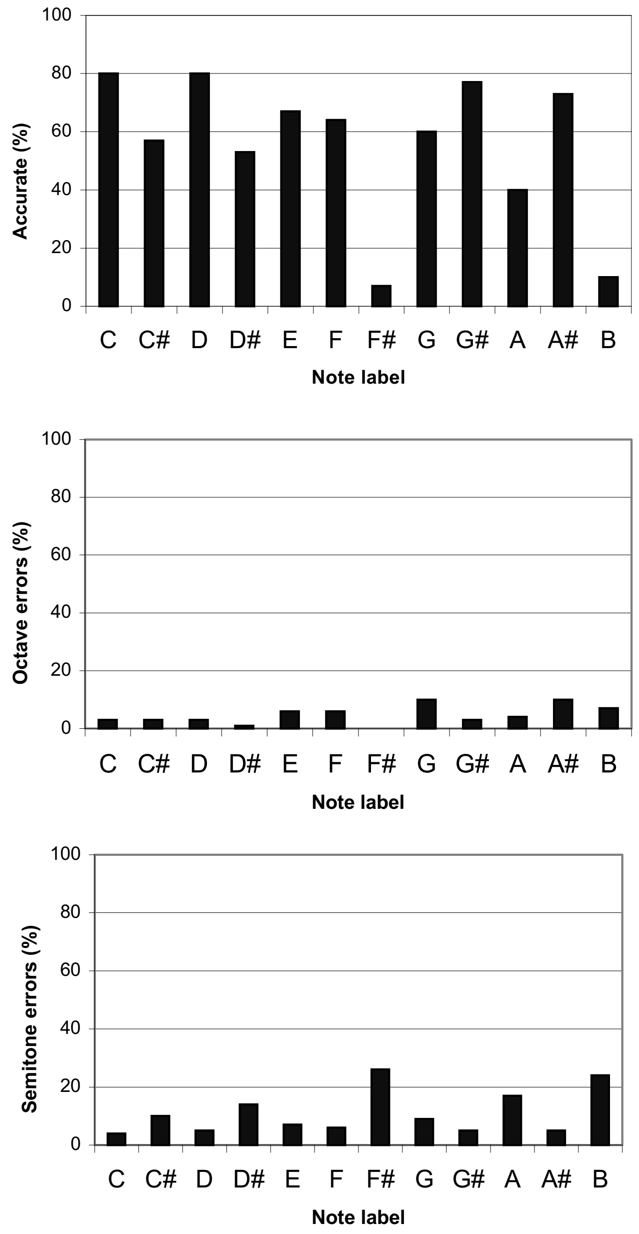


FIGURE 1 *Jean's note-labelling performance by note.*

by type for different participants. These effects are explicated for the relevant participants in their case profiles.

Table 7(a), (b) and (c) examines the effects of timbre in each register of the stimulus set (high, mid, low). It can be seen that timbrel effects, where they

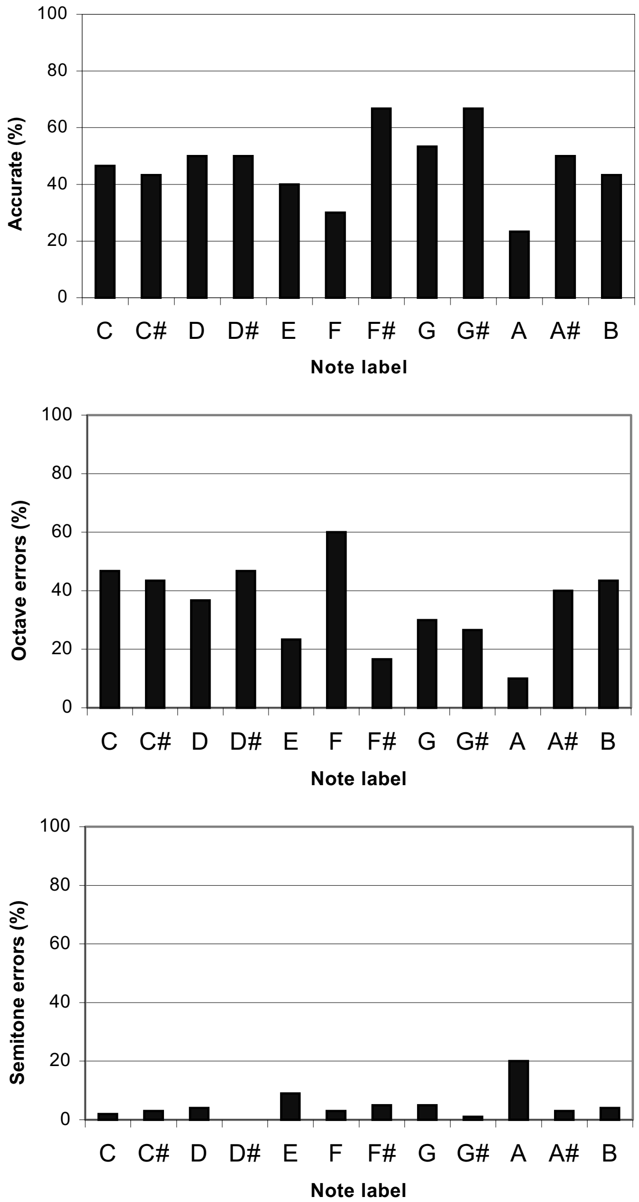


FIGURE 2 *Tina's note-labelling performance by note.*

occur, were limited to the extreme ranges (e.g. Lesley, Andrew and Helen in the high register; Tina, Sonja, Cathy and Lesley in the low register).

Figure 3 illustrates the percentage of all correct responses with 8va discrepancies for each AP participant. Tina, Kerri and Cathy made a comparatively higher proportion of octave-discrepant responses than the other

TABLE 6 *Comparison of the number of correct responses, octave and semitone discrepancies for each participant in each timbre*

	Timbre effects (N = 120 for each timbre set)											
	Number of correct responses				Number of responses with 8va discrepancy				Number of responses with semitone discrepancy			
	Piano	Sine	Own instrument	$\chi^2(2)$	Piano	Sine	Own instrument	$\chi^2(2)$	Piano	Sine	Own instrument	$\chi^2(2)$
Andrew	39	44	55	2.91	3	3	1	1.14	27	22	49	12.63**
Anthony	33	34	N/A	$\chi^2(1) = 0.01$	6	5	N/A	$\chi^2(1) = 0.09$	41	50	N/A	$\chi^2(1) = 0.89$
Cathy	51	83	67	7.64*	26	16	20	2.45	43	20	29	8.76*
Helen	30	23	29	1.05	10	13	11	0.27	35	51	52	10.98**
Jean	61	65	74	1.33	5	6	6	0.12	54	46	36	3.59
Julie	81	25	88	36.87***	7	8	1	5.37	1	1	1	0.00
Kerri	89	85	87	0.06	25	28	24	0.34	6	5	7	0.33
Leo	119	112	113	0.25	0	7	5	N/A	0	0	2	N/A
Lesley	16	16	20	0.62	7	22	4	16.91***	31	42	35	1.72
Sonja	27	47	45	6.12*	15	18	24	2.39	61	46	40	14.05***
Tina	43	59	67	5.30	52	46	29	6.27*	24	15	19	2.10

*$p < .05$; **$p < .01$; ***$p < .001$.

TABLE 7(a) *Number of correct responses in the high range (B^4–E^3) as affected by timbre for each participant*

	Timbre effects in the high range (*N* = 40 in each timbre set)											
	Correct responses				8va discrepancy				Semitone discrepancy			
	Piano	Sine	Own	$\chi^2(2)$	Piano	Sine	Own	$\chi^2(2)$	Piano	Sine	Own	$\chi^2(2)$
Andrew	13	16	16	0.40	0	0	0	N/A	10	5	22	12.4**
Anthony	6	10	N/A	$\chi^2(1) = 1.00$	2	3	N/A	$\chi^2(1) = 0.20$	15	16	N/A	$\chi^2(1) = 0.03$
Cathy	17	25	15	2.94	10	9	13	0.81	13	4	10	4.67
Helen	14	4	15	6.72*	3	10	4	5.06	8	19	10	5.56
Jean	20	21	21	0.03	3	5	6	1.0	16	14	12	0.57
Julie	26	8	29	12.28**	6	6	1	3.84	0	0	0	N/A
Kerri	36	36	38	0.07	3	4	0	0.14	1	0	2	.33
Leo	39	34	36	0.35	0	5	4	N/A	0	0	0	N/A
Lesley	7	7	5	0.42	4	10	1	8.40*	12	17	13	1.00
Sonja	7	11	16	3.59	4	9	5	2.33	22	18	17	0.73
Tina	21	24	26	0.54	10	9	10	0.07	8	7	3	2.33

$^*p < .05$; $^{**}p < .01$; $^{***}p < .001$.
Each timbre set had 40 stimuli in the high range.

TABLE 7(b) *Number of correct responses in the middle range (G#[1]–D#[3]) as affected by timbre for each participant*

	Timbre effects in the middle range (N = 40 in each timbre set)											
	Correct responses				8va discrepancy				Semitone discrepancy			
	Piano	Sine	Own	$\chi^2(2)$	Piano	Sine	Own	$\chi^2(2)$	Piano	Sine	Own	$\chi^2(2)$
Andrew	10	14	19	2.83	3	2	0	N/A	9	8	13	1.40
Anthony	17	17	N/A	$\chi^2(1) = 0.00$	3	1	N/A	$\chi^2(1) = 1.00$	10	15	N/A	$\chi^2(1) = 1.00$
Cathy	24	33	25	1.78	0	1	0	N/A	16	6	14	4.67
Helen	11	14	10	0.74	7	2	3	3.50	12	16	19	1.57
Jean	25	27	27	0.10	0	0	0	N/A	15	12	10	1.03
Julie	33	10	34	14.36***	2	5	1	3.25	0	1	0	N/A
Kerri	40	39	40	0.02	0	0	0	N/A	0	1	0	N/A
Leo	40	40	39	N/A	0	0	1	N/A	0	0	0	N/A
Lesley	7	6	12	2.48	1	2	1	0.50	17	13	9	2.46
Sonja	17	20	20	0.31	2	3	3	0.25	19	15	13	1.19
Tina	16	23	19	1.27	15	12	11	0.68	9	5	9	1.39

*$p < .05$. **$p < .01$. ***$p < .001$.
Each timbre set had 40 stimuli in the middle range.

TABLE 7(c) *Number of correct responses in the low range (C^0–G^1) as affected by timbre for each participant*

	Timbre effects in the low range (N = 40 in each timbre set)											
	Correct responses				8va discrepancy				Semitone discrepancy			
	Piano	Sine	Own	$\chi^2(2)$	Piano	Sine	Own	$\chi^2(2)$	Piano	Sine	Own	$\chi^2(2)$
Andrew	16	14	20	1.12	0	3	0	N/A	7	9	13	1.93
Anthony	10	7	N/A	$\chi^2(1) = 0.53$	1	1	N/A	$\chi^2(1) = 0.00$	15	19	N/A	$\chi^2(1) = 0.47$
Cathy	10	25	26	7.90*	16	6	7	6.27*	14	9	7	2.60
Helen	5	5	4	0.14	0	1	3	1.00	15	17	23	1.89
Jean	16	17	26	3.08	2	0	0	N/A	20	20	13	1.84
Julie	22	7	23	9.27**	1	2	1	0.50	1	0	1	N/A
Kerri	13	10	9	0.81	22	24	24	0.11	5	4	5	0.14
Leo	40	38	38	N/A	0	2	0	N/A	0	0	2	N/A
Lesley	3	3	3	0.00	2	10	2	9.14*	10	14	13	0.70
Sonja	2	16	9	10.9**	9	6	15	4.20	19	17	9	3.73
Tina	6	12	22	9.80**	26	25	8	10.4**	7	3	7	1.88

*$p < .05$; **$p < .01$; ***$p < .001$.
Each timbre set had 40 stimuli in the low range.

participants. Figure 4 shows the percentage of correct responses with semitone discrepancy for each participant. Leo made very few semitone-discrepant responses, but then there was little room for discrepancy in his profile because of his extremely high proportion of completely accurate responses (correct label in the correct octave). Jean, Sonja, Helen and Anthony made a comparatively high proportion of responses with semitone discrepancy. Table 1 and Figures 3 to 6 show that AP skill is variable and not a unitary phenomenon.

Table 8 is a simplification and overview of all evaluated effects for all 11 participants. Dimensions of the stimulus tones that appear to significantly affect the performance accuracy of most participants was range on octave misjudgments (number affected = 7), and specific note preferences (number affected = 6). Generally though, this table shows a variable spread of influence on performance accuracy for the participants.

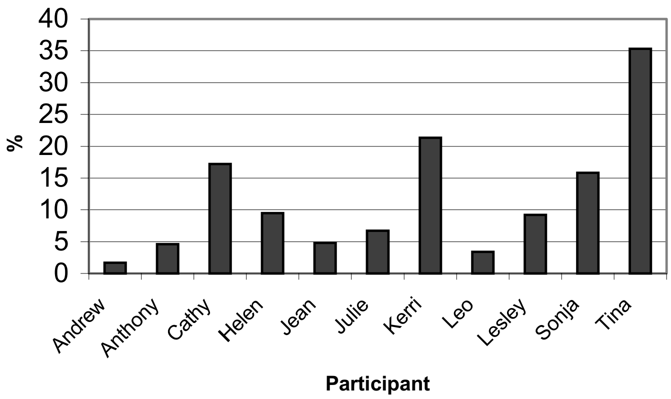


FIGURE 3 *Comparison of total percentage of octave errors by participant.*

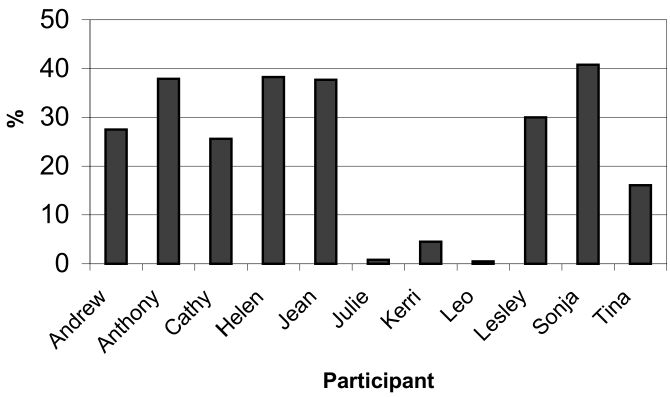
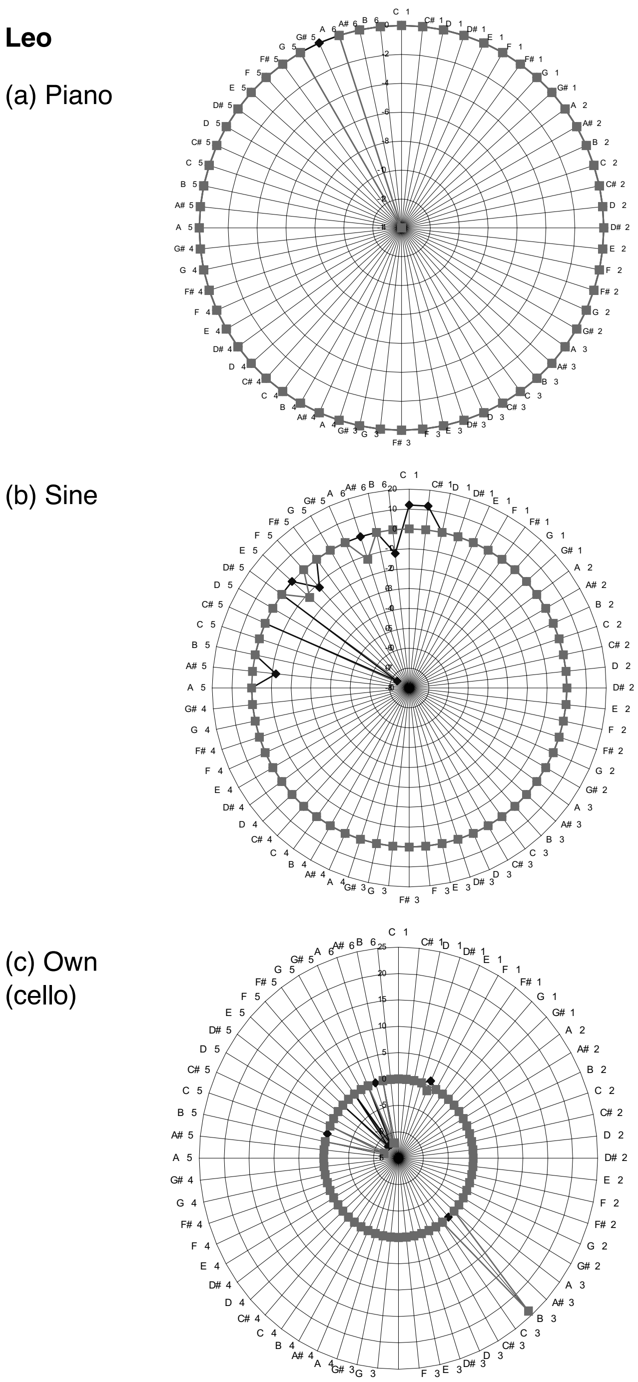


FIGURE 4 *Comparison of total percentage of semitone errors by participant.*

TABLE 8 *Number of AP participants whose note-labelling performance was affected by dimensions of the tones in the stimulus sets*

	Number affected ($n = 11$)
by note colour (black vs white) overall	3
by note colour in:	
piano timbre	2
sine timbre	3
own timbre	3
by note colour in:	
high range	2
middle range	5
low range	3
by note overall	7
by range on:	
accuracy	6
8va error	7
semitone error	1
by timbre in:	
high range	4
mid range	1
low range	5
by timbre on:	
accuracy	3
8va error	2
semitone error	4

Figures 5 and 6 are example radar plots generated from the AP profiling program for Leo and Tina for each timbre set. The x axis (circumference) shows each individual pitch ranging from C^0 (the lowest C) to B^4 (the highest B). The radial y axes extend from the centre point like wheel spokes and represent the discrepancy of the participant's responses from completely accurate (for example a score of +5 would indicate that the participant's response over-judged the pitch by five semitones). Therefore, correct responses would circle around the centre point of these charts along the zero line. Correct responses with 8va discrepancy would also circle the centre point but at the +12 or –12 line. Leo's piano chart (Figure 5(a)), for example, reflects a response set featuring absolutely no discrepancies. Tina's sine chart (Figure 6(b)) for example, shows a propensity for octave over-judgments in the low range.



Since each tone was heard twice in each stimulus

FIGURE 5 *Leo's response discrepancies (in notes) for each timbre.*

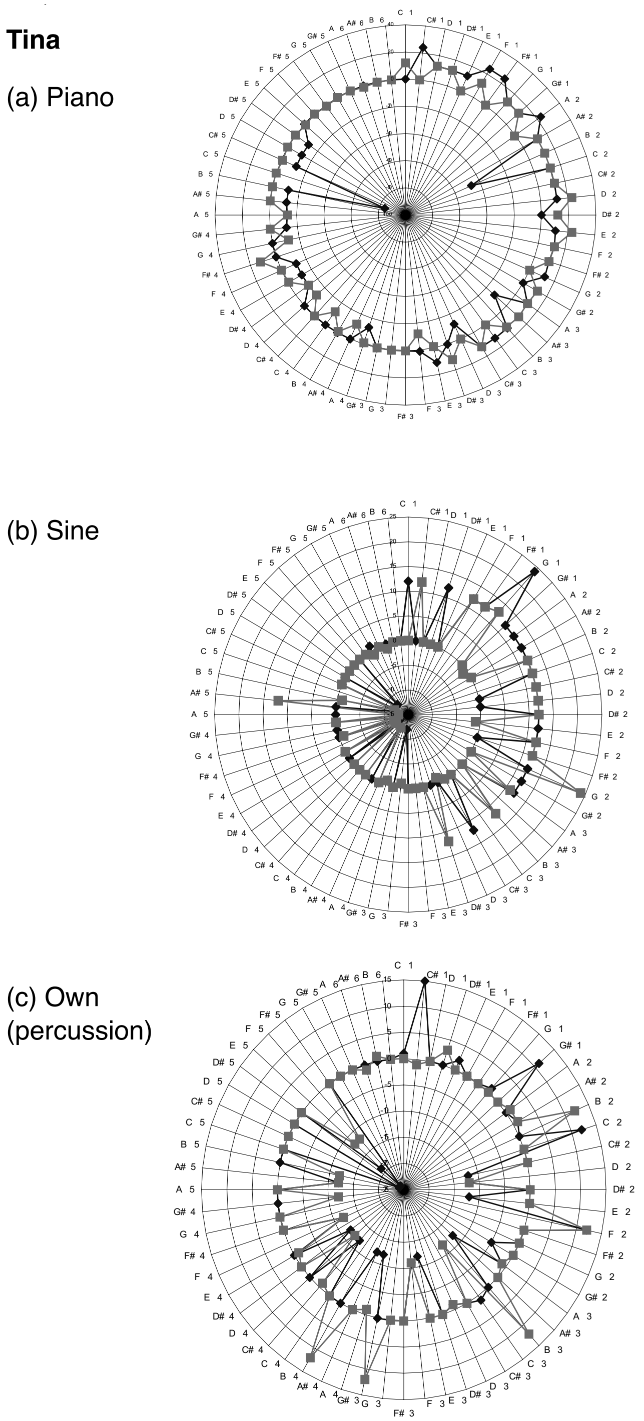


FIGURE 6 *Tina's response discrepancies (in notes) for each timbre.*

set, the radar charts show two line plots. If both responses were identical, the second line masks the first, if the responses differed for the same note each time it was heard in a particular stimulus set for a participant, then both points along that note's *y* axis can be seen.

By comparison, the RP charts (e.g. Figure 7) appear to indicate random responses. Clare's profile for piano timbre is somewhat different from her profile for sine timbre. For sine she appears to be able to find roughly the octave of the stimulus, whereas for piano tones, she shows a tendency to put all her responses in the middle range regardless of the actual register of the stimulus. Thus, it would appear that RP participants are not simply bereft of tone-labelling skill, rather, their note-labelling attempts reflect the utilization of an inadequate strategy when faced with labelling tones without a reference.

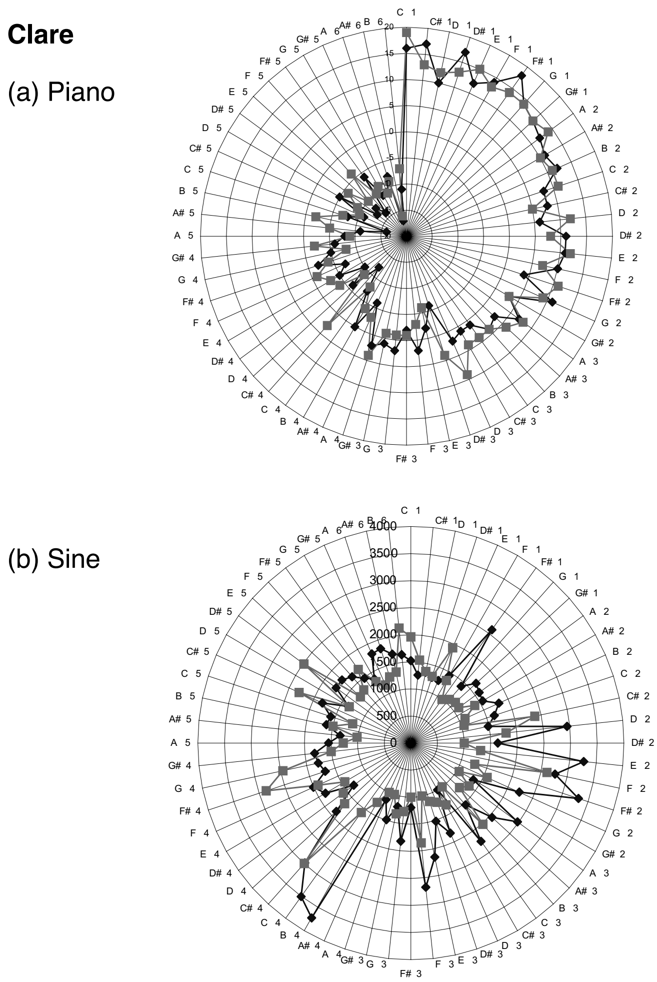


FIGURE 7 *Clare's (RP) response discrepancies (in notes) for each timbre.*

The post-profiling interviews provided insight to the role of familiarity in the development of an AP profile. Below are some examples.

Andrew

Andrew's ability to label tones without a reference was significantly greater than chance and so he can be accepted as having demonstrated AP skill. Andrew's performance was affected by timbre, with responses for black notes significantly more accurate than for white tones in the timbre of his own instrument (French horn). Andrew made extremely few octave misjudgments, and so comparison of this type of discrepancy over timbre, range, and note colour was inappropriate. His responses varied with respect to semitone discrepancy by timbre however, with most occurring for tones of his own instrument's timbre, and least for piano tones. A range effect was also evident by timbre. Andrew made significantly fewer semitone misjudgments in the high register for sine tones than for piano and own instrument timbres. Andrew's performance did not vary by note. This absence of a note effect is counter to what would be expected when considering Andrews comments following the profiling activity:

> What I found was that a few notes that I know, and when I hear them, like G is one, G is one that I know and so I'd be going through not quite sure if I was choosing the right notes and there'd be a G and then I'd either have to adjust my thinking from the previous notes or I'd know that I was on the right track.

Helen

Helen's ability to label tones without a reference was significantly greater than chance, so she can be accepted as having demonstrated AP skill. In Helen's post-profiling interview she indicated that the sine wave timbre set was difficult for her: 'I've always had trouble recognizing computer sounds.' This effect was apparent in her profile for only tones in the high range, with significantly fewer accurate responses occurring for the sine timbre than for own and piano timbres.

Helen also commented on the effect of range on her performance: 'So I felt that the lower register I couldn't hear very well.' It is unclear whether this was a volume control or simple identification problem for Helen. Accuracy was affected by range with Helen performing significantly less accurately in the low range than for the middle and high.

Note-labelling performance varied by stimulus note. She was most accurate for F and A# (B♭), and least accurate for C# and B. However, she made more octave errors for F, A# (B♭) and C#, and although G# was only moderately accurate, she made absolutely no octave errors. Including all allowable errors, Helen was significantly better able to label D and A than any other notes. The most semitone errors were for B, E and A.

Julie

Julie's ability to label tones without a reference was significantly greater than chance, and so she can be accepted as having demonstrated AP skill. There was no significant overall range effect, but range did affect responses according to note colour. Julie responded significantly more accurately for black tones than for white, especially in the low register. This was particularly true in tones of her own timbre (piano), and was not evident in the other timbres.

Julie's responses did not vary significantly by note, but there was a significant timbre effect with responses for the sine tones significantly less accurate than the other timbres, across all ranges.

Leo

Leo's tone-labelling ability was significantly greater than chance and so clearly demonstrated AP skill. Leo made so few errors that it was inappropriate to test for any differences by range, timbre or note type. He appeared unaffected by most aspects of difference in the stimulus sets. In his post profiling interview he said: 'I had no problems, no.' Leo's accuracy was, however, significantly affected by note colour in his own instrument's timbre with black significantly more accurate than white tones. Leo was an accomplished cellist and the cello is an instrument noted for its broad range. The range of the cello exceeds the range of the tones in the tone set for this study.

Lesley

Lesley's ability to label tones was significantly greater than chance and so it can be accepted that she demonstrated AP skill. Her comments in the post-profiling interview focused on the effects of timbre, especially on the likelihood of semitone misjudgments because of her tendency to think in transposed pitch to cater for performance on clarinet:

> I found it difficult with different timbres . . . it depended on what sound it was, I wasn't sure whether I should be saying it in B flat pitches or C pitches.

A timbre effect was found for most aspects of her performance.

Her accuracy differed significantly for white and black tones with greatest accuracy for white. This effect was consistent for both the piano and sine timbres, but she demonstrated no significant preference for note colour in tones of her own instrument's timbre. Her accuracy for white tones was significantly greater than for black tones in the high and low registers. She made significantly more semitone misjudgments for black tones than white in the high and middle ranges.

There was also a significant note effect on Lesley's performance for accuracy, octave and semitone discrepancy. Her most accurate notes were A and C, and this trend was consistent for accuracy including 8va misjudgment. She responded with significantly more semitone discrepancy for A# (B♭) than for other notes. Responses for the note D were significantly less accurate than for any other note overall. Lesley's responses also varied significantly by range, demonstrating least accuracy in the low range, and least 8va discrepancy in the middle register. She showed a significant timbre effect on 8va discrepancy in the high and low registers, making significantly fewer misjudgments in the piano and own instrument timbres.

Tina

Tina's tone-labelling ability without a reference tone was significantly greater than chance and so Tina can be accepted as having demonstrated AP skill. In the interview following the profiling she said: 'You just follow your instinct.' This suggests that she was not aware of any decision action prior to her responses. Tina's labelling accuracy varied significantly for black and white notes overall and for the sine timbre, with black notes most accurate. There were significant note differences for accuracy, 8va and semitone discrepancy, with A as the least accurate tone and attracting significantly more semitone errors than as for the other notes.

Tina's overall responses differed by range for accuracy and for octave misjudgment. She was most accurate in the high range and made most octave misjudgments in the low range. There was also a significant timbre effect on octave discrepancy with most occurring for tones from the piano timbre set, and the fewest for her own instrument's timbre set (marimba). This timbre effect was not apparent for complete accuracy or semitone discrepancy. Timbrel differences were apparent by range. Tina's responses were affected in the low range for accuracy and octave discrepancy, with her own timbre most accurate and piano least, most octave misjudgments occurred for piano tones and the least occurred for tones from her own instrument's timbre.

In the middle range, Tina's responses showed significant difference for black and white notes for accuracy and semitone discrepancy. In this range Tina favoured black notes as the most accurate. This effect did not extend to the high or low ranges and did not vary by timbre.

Discussion

AP AS A DIVERSE SKILL

The profiles indicate that AP is a very diverse skill. The profile for each participant was distinctly different. Moreover, manipulation of sound stimuli impacted on AP performance in a diversity of ways. This suggests that AP knowledge does not simply involve sound frequencies. AP appears to depend on multiple attributes of sounds including tonality, timbre and range. Thus, it appears that these multiple attributes of sound are key elements in the instantiation of pitch knowledge into schemata. This is in keeping with Takeuchi and Hulse's (1993) multidimensional model, which allows for multiple dimensions of variation in AP performance.

However, the findings of this study challenge Takeuchi and Hulse (1993) and other researchers (Miller and Clausen, 1997; Miyazaki, 1988, 1989) who have sought some general AP trends in pitch identification. These researchers have suggested that AP responses were generally faster and more accurate in the middle region (Miyazaki, 1988, 1989; Rakowski and Morawska-Bungeler, 1987), for sine tones, for white notes (Miller and Clausen, 1997; Ward, 1999), and for common tones (Miyazaki, 1989). AP participants in this study did not show general regularities in their responses. Rather, the participants' responses were quite distinct from each other.

For example, where some participants, such as Lesley, responded significantly more accurately for white key tones than for black, others such as Anthony responded more accurately for black key tones than white. Still others such as Cathy showed no difference in their AP identification responses for black and white key tones.

This diversity of response between participants was evident for each attribute of the sound stimuli (timbre, range and tonality). Some responses appeared to reflect combination effects of one attribute of the sound stimuli with another. For example, Tina's responses were significantly affected by note key colour, but only for sine tones.

Points of distinction between individuals reflected the impact of experience and musical competence on the nature and structure of pitch identification knowledge. This finding was consistent for people with AP or RP. AP and RP emerged as different kinds of knowledge for pitch, one founded on absolute features of sound (AP) and the other on relational features (RP). Any commonalities in the accuracy of pitch-labelling responses across individuals appeared to reflect shared or similar musical experiences, rather than any overriding template for AP.

Researchers who have explored variations in pitch identification skills have argued that people with AP responded in consistent ways. They argued that there were particular hallmarks of the skill that could be identified (e.g. Bachem, 1937; Miyazaki, 1992; Takeuchi and Hulse, 1991, 1993). Some writers (Baird, 1917; Miller and Clausen, 1997; Miyazaki, 1989; Rakowski and Morawska-Bungeler, 1987; Sergeant, 1969; Simpson and Huron, 1994; Takeuchi and Hulse, 1991; Ward, 1999) have sought to detail these general attributes for AP. For example, Baird (1917), one of the earliest researchers, reported a range effect concentrating on judgment errors in the extremely high and low registers as a general feature of AP. Later, Sergeant (1969) described a note colour effect impacting on semitone errors. He argued that a hallmark of AP was the tendency to mistake black notes (e.g. A#) with their white note neighbours. More recently, Miyazaki (1989) described a pitch class effect impacting on octave errors. He argued that participants with AP did not encode octave information with their knowledge for discrete pitches, rather, there was a general tendency to disregard octave and register placement of tones.

The variation between the responses of individuals in this study challenges the idea that AP may be described as a unidimensional skill with consistent general attributes. Where other researchers reported on the apparent similarities between participants, this study found diversity and points of distinction between individuals. Participants were variably influenced by attributes of the sound stimuli. Each individual's pattern of responses was unique and distinct from all other participants. Instead of supporting the argument that attributes of sound may impact on pitch-labelling performance in a way that was similar for different individuals, the data presented evidence that the AP participants' responses were quite different from each other.

Diversity in range effects

Researchers such as Miyazaki (1989) and Rakowski and Morawska-Bungeler (1987) suggested that AP responses were more accurate in the mid pitch region. This was apparent from the profiles for some participants (Anthony and Sonja). However, not all participants showed any range effect at all, and those that did, did not necessarily perform best in the mid range. For example, Tina's accuracy was significantly better in the upper region than in the low and mid region. She explained this in terms of her familiarity with the range of the marimba, 'because normally the marimba doesn't actually go beyond two octaves below middle C } I couldn't hear it. I know it's the lower octave but I just have to guess the note.'

For Andrew and Leo, pitch region had limited effect on accuracy. Both Andrew's instrument (French horn) and Leo's (cello) have extremely wide ranges and so they would have been reasonably familiar across the whole range of the stimulus set. Both the French horn and the cello have an effective range of around four octaves.

There also appeared to be an interaction between pitch region and timbre for some participants. For example, Julie was much less accurate in the low range for sine tones than for other timbres in the same range, and yet Cathy was much more accurate for sine tones than piano tones in the same range. The consistent element that appeared to support pitch region variations in AP performance was musical experience and specific instrumental competence. For example, Cathy's preferred instrument was the violin, which can sound similar to sine wave tones. Her familiarity with the violin may have assisted her to label the pitches of sine wave tones, but because of the distinct differences between piano and sine wave sounds, may not have assisted her to label piano tones in the same register, hence explaining her poorer performance.

Diversity in note colour effects

Takeuchi and Hulse (1991) conducted an extension and replication of Miyazaki's (1989) study, and along with Miller and Clausen (1997) suggested that AP performance was more accurate for white notes than for black. The earlier study by Miyazaki was criticized on the basis that the physical configuration of the keyboards utilized in his study might have exaggerated the effect. Takeuchi and Hulse's (1991) investigation controlled for this. However, it showed the same AP pattern of responses for black and white notes. Miller and Clausen (1997) and Ward (1999) argued that performance might be less accurate for black notes compared to white, because black notes have longer names and are labelled relative to their white note neighbours.

This study controlled for the criticisms of Miller and Clausen (1997) and Ward (1999) by requiring participants to make a common physical response at first recognition of each presented tone, black or white, before selecting the correct key in their own time (up to five seconds). The data indicated that the expected more accurate responses for white tones compared to black reported by these other researchers were not universally true for participants with AP.

While some participants showed note colour effect, others did not. Those who did were not consistent in favouring white note responses, as Miyazaki (1989, 1990, 1992) would predict. For example, Lesley was much surer with white notes than she was with black, but this was only evident in the high range. Anthony was more accurate for black notes than for white, and Cathy showed no difference in performance. Tina was significantly affected by note colour in sine timbre but not any other. Contrary to other research findings, she performed more accurately with black than white tones. This diversity of performance profiles challenged the concept of general tendency in AP responses.

Diversity in timbre effects

In general it was not possible to accurately predict an individual's pitch-labelling responses to one stimuli set on the basis of another. Overall accuracy and error types were quite variable from one timbre set to another. For example, Leo's pitch-labelling accuracy with piano tones would have constituted a clear attribution of AP according to Bachem (1937), but his performance with sine tones would not have been so clear. Additionally, not all participants found sine waves easier as Ward and Burns (1982) would predict, or harder to label as Balzano (1984) would predict. Responses to the sine wave and piano tones were particularly variable between participants.

Diversity in note class effects

Individuality and diversity were also demonstrated in accuracy of identifying tones which varied by note class. Myazaki's (1989) study noted a particular configuration of response accuracy to individual note classes for the seven participants in his study. He predicted that principal tones of C major (tonic C and dominant G) would generally attract accurate AP responses. Miyazaki argued for the role of tonal familiarity in this effect. However, he considered this to be influenced by overriding societal musical contexts on general AP skill. He ignored the idea that differences might result from individual's distinct experiences.

AP was much less consistent across individual performance profiles for note class than Miyazaki's research would predict. Not all participants showed a note class effect, and those who did at times made octave errors for predominantly specific notes and not all 12 tones. For example, Anthony was significantly more accurate for G# and A#, Cathy and Sonja were significantly more accurate for C. Helen's best notes were F, B♭, and A. For these individuals, the particular notes of greatest accuracy reflected the scales and instruments of their personal musical experience. As a trumpeter in a brass band, Anthony would be particularly familiar with B♭ (A#) as the tuning note. B♭ is also the fundamental tone of the trumpet and is the first tone usually learnt. Cathy reported a high level of familiarity with the key of C: 'I suppose the flats and sharps might be a little bit harder, but you know C scales probably are a bit easier yeah', which may have assisted her in identifying C more accurately and quickly. Sonja reported the same familiarity with C: 'C used to be the easiest of all } it's a piano based thing } C is pretty easy to identify.' Helen reported both A and B♭ to be common tuning notes in her regular ensembles and F was the fundamental sounding tone for the French horn.

Consistency of AP accuracy

It emerged that consistency of response accuracy was not necessarily tied to overall AP performance. Some participants with lower overall response accuracy were more generally consistent in their error types than participants with higher overall accuracy. For example, Jean made consistent semitone errors across the three timbres of the stimuli and yet her ability to identify tones was generally much surer than Lesley's, who was not as consistent in her type of response error. Consistency of errors is not an attribute of AP that has been explored in the past. This is perhaps because most researchers have disregarded incorrect responses as being not demonstrative of AP. This approach has limited the range of data considered regarding the nature of knowledge for AP and may not assist in the development of a clearer picture of AP skill.

Impact of dimensions of sound such as timbre, range and tonality on the nature and structure of knowledge for AP

Not only do aspects of sound such as timbre, range and tonality appear to provide points of distinction and diversity between individuals; the participants' response data supported the notion that these central attributes of sound form the anchors for AP development. Seemingly extraneous sound attributes such as timbre had a pronounced effect on the ability of participants to effectively identify pitch.

For example, Julie found sine tones to be extremely difficult to identify. She seemed to be generally unable to label tones in that timbre and yet her AP performance in response to piano tones was very accurate. The sine tones were unfamiliar to Julie and did not correspond to the anchors for her pitch schemata.

Bachem (1937) suggested that extraneous aural features would not influence genuine AP. He argued that as long as the frequency of a tonal stimulus was identifiable, then pitch labelling should not be affected. The response data showed that the performance of all participants with AP was affected to some extent either in their labelling accuracy by competence and attributes of the stimuli. The data suggest a multidimensional model for AP skill. While there were distinct differences in profiles between participants with RP and those with AP, AP is not effectively described using a simple dichotomous model, which would consider AP to be either present or not present.

Takeuchi and Hulse (1993) proposed a measure of AP that was continuous along several dimensions. However, Takeuchi and Hulse's multidimensional model, which considers the impact of the various attributes of sound on AP skill, appears to go only part way in assisting in the description of AP and the better modelling of the nature and structure of knowledge for AP. It does not fully consider the combination effects that appeared to occur in the present study between the various dimensions of sound for tonal identification.

A MULTIDIMENSIONAL VIEW OF AP

Results from this study support a broad view of AP definition. The multidimensional and individual nature of AP that emerged in the data suggested that it would be inaccurate to place a meaningful dividing line between 'AP' and 'not AP' on the basis of accuracy criteria. The broad definition should simply accept as AP any ability to label a tone without a reference. AP is like other skills where the dividing line between skilled and unskilled is blurred. Thus, it is unreasonable to define AP in terms of its quality, but certainly the skill of labelling a tone without a reference can be described by using a quality indicator such as accuracy.

Each participant in the study showed the influence of experience on their ability to identify pitch. Responses were consistently more accurate for tones that were familiar. For example, sine tones are generally unfamiliar as musical sound. Except for some computerized music, sine tones are not often specifically scored in music compositions. As a result most musicians have limited familiarity with sine tones even though some electronic organ settings, notes from the high register of a pipe organ, or some sounds produced on a bowed stringed instrument are quite similar. The data indicate a large majority of participants performed less accurately when presented with sine tones. A notable exception was Kerri. Kerri reported that the sine tone stimuli sounded similar to her violin. Even though some researchers suggest that sine tones are easy to recognize given their lack of 'confusing' overtones (Ward and Burns, 1982) the data indicate that experience with sounds that share similarities with sine tones most directly influenced pitch labelling performance of participants.

It has been argued that pitch identification knowledge is 'imprinted'. For example, Jeffress (1962) argued that the construction of pitch knowledge is based on single event experiences, as long as they occurred as the first experiences with pitch. He believed that once a sound's frequency had been cognitively coded or tagged with a pitch label, musicians with AP would be unlikely to confuse these labels when exposed to timbrel variants, especially if the frequencies of tones were clear.

The data in this study indicate that knowledge of pitch is acquired incrementally over time as the result of on-going experience, Participants showed confusion when they were presented with timbrel variants especially where the nature of the stimuli suggested a transposed set of pitch labels. For example, Lesley, Andrew and Anthony found the stimuli of a transposing instrument to evoke a transposed set of labels rather than their true concert pitch names. Lesley tried to convert tones of clarinet timbre to B♭ transposed pitch: 'I found it difficult with different timbres . . . it depended on what sound it was, I wasn't sure whether I should be saying it in B flat pitches or C pitches.' Andrew attempted to label tones that were French horn timbre according to F transposed pitch. Anthony was confused by sine tones, which he felt sounded like trumpet tones. He attempted to identify them with B♭ transposed labels (as would be appropriate for trumpet tones) and as a result made more semitone errors for sine timbre stimuli than for piano timbre.

For Lesley, Andrew and Anthony, the frequency of the tones was not the only key to appropriate pitch labelling. The timbre also provided a vital cue for their identification of tones. For these people, a given frequency was labelled differently depending on the apparent timbre. The frequencies were not permanently locked to a particular note name as suggested by an imprinting model. Rather, it appeared that attributes of sounds presented cues used by the participants to actively decide on an accurate label from an array of possible choices. Given that each of them began learning music with a non-transposing instrument (piano or electronic keyboard), it was clear that Jeffress's 'imprinting' process did not explain their AP knowledge and its application to labelling tasks. It was more likely that individuals developed their knowledge for pitch over time using the various attributes of sounds.

Jeffress's (1962) model suggested that an individual's first musical experiences would be the most influential on their responses. However, participants in this study showed a marked preference for sounds that reflected the instrument with which they had the most intensive and/or current engagement. This was the case even if the sounds were not of the same type of instrument as their first musical experiences.

Several participants showed the effect of recent musical experiences on their pitch labelling. Andrew commented on the effects of his recent focus on the French horn. This was reflected in his preference for stimuli in the range of the French horn, even though the first instrument he learnt was piano. Helen's responses were also affected by range, reflecting her recent performance focus on high woodwind instruments. Another example was Tina, who performed best with tonal stimuli whose attributes most closely aligned to the marimba, her best and most current instrument. This reflected her currency with the aural attributes of the marimba's timbre rather than her first musical experiences, which were with the piano. Tina's knowledge for pitch had changed and elaborated since her first musical experiences. Tina reported that she had spent little time with the piano in recent years, and so by comparison to her current familiarity with the marimba, the piano had become less familiar. If she had acquired AP through Jeffress's 'imprinting' process, she would not have demonstrated any difference between her AP labelling of marimba tone stimuli and piano tones (her first instrument). Alternatively, her responses to piano tones would have been surer than her responses to marimba tones. These new experiences appeared to add detail to each individual's earlier notions of pitch.

In summary, this study has clearly demonstrated that AP is a multidimensional skill characterized by unique performance for individuals tied to their learning histories and musical experience.

REFERENCES

Bachem, A. (1937) 'Various Types of Absolute Pitch', *Journal of the Acoustical Society of America* 9: 146–51.

Bachem, A. (1940) 'The Genesis of Absolute Pitch', *Journal of the Acoustical Society of America* 11: 434–9.

Bachem, A. (1954) 'Time Factors in Relative and Absolute Pitch Determination', *Journal of the Acoustical Society of America* 26: 751–3.

Bachem, A. (1955) 'Absolute Pitch', *Journal of the Acoustical Society of America* 27: 1180–5.

Baharloo, S., Johnston, P.A., Service, S.K., Gitschier, J. and Freimer, N.B. (1998) 'Absolute Pitch: An Approach for Identification of Genetic and Nongenetic Components', *American Journal of Human Genetics* 62: 224–31.

Baharloo, S., Service, S.K., Risch, N., Gitschier, J. and Freimer, N.B. (2000) 'Familial Aggregation of Absolute Pitch', *American Journal of Human Genetics* 67: 755–8.

Baird, J.W. (1917) 'Memory for Absolute Pitch', in E.C. Sanford (ed.) *Studies in Psychology, Titchener Commemorative Volume*, pp. 43–78. Worcester, MA: Wilson.

Balzano, G.J. (1984) 'Absolute Pitch and Pure Tone Identification', *Journal of the Acoustical Society of America* 75: 623–5.

Bloom, B.S. (1982) 'The Role of Gifts and Markers in the Development of Talent', *Exceptional Children* April: 510–22.

Brady, P.T. (1970) 'Teach Yourself Absolute Pitch in 365 Hard Lessons', *Journal of the Acoustical Society of America* 48: 88.

Burns, E.M. and Campbell, S.L. (1994) 'Frequency and Frequency Ratio Resolution by Possessors of Absolute and Relative Pitch: Examples of Categorical Perception?', *Journal of the Acoustical Society of America* 96: 2704–19.

Cohen, A.J. and Baird, K. (1990) 'Acquisition of Absolute Pitch: The Question of Critical Periods', *Psychomusicology* 9: 31–7.

Costall, A.P. (1985) 'The Relativity of Absolute Pitch', in P. Howell, I. Cross and R. West (eds) *Musical Structure and Cognition*, pp. 189–208. London: Academic Press.

Dowling, W.J. (1999) 'The Development of Music Perception and Cognition', in D. Deutsch (ed.) *The Psychology of Music*, pp. 603–25, 2nd edn. London: Academic Press.
Grebelnik, S.G. (1984) 'Training Preschoolers to Develop Absolute Pitch', *Voprosy Psikhologii* 2: 90–8.
Jeffress, L.A. (1962) 'Absolute Pitch', *Journal of the Acoustical Society of America* 34: 987.
Lenhoff, H.M., Perales, O. and Hickok. G. (2001) 'Absolute Pitch in Williams Syndrome', *Music Perception* 18(4): 491.
Levitin, D. (1994) 'Absolute Memory for Musical Pitch: Evidence from the Production of Learned Memories', *Perception and Psychophysics* 56: 414–23.
Miller, L. and Clausen, H. (1997) 'Pitch Identification in Children and Adults: Naming and Discrimination', *Psychology of Music* 25: 4–17.
Miyazaki, K. (1988) 'Musical Pitch Identification by Absolute Pitch Possessors', *Perception and Psychophysics* 44: 501–12.
Miyazaki, K. (1989) 'Absolute Pitch Identification: Effects of Timbre and Pitch Region', *Music Perception* 7: 1–14.
Miyazaki, K. (1990) 'The Speed of Musical Pitch Identification by Absolute Pitch Possessors', *Music Perception* 8: 177–88.
Miyazaki, K. (1992) 'Perception of Musical Intervals by Absolute Pitch Possessors', *Music Perception* 9: 413–26.
Miyazaki, K. (1993) 'Absolute Pitch as an Inability: Identification of Musical Intervals in a Tonal Context', *Music Perception* 11: 55–72.
Miyazaki, K. (1995) 'Perception of Relative Pitch with Different References: Some Absolute Pitch Listeners Can't Tell Musical Interval Names', *Perception and Psychophysics* 57: 962–70.
Oakes, W.F. (1951) 'An Alternative Interpretation of "Absolute Pitch"', *Transactions of the Kansas Academy of Sciences* 54: 396–406.
Profita, J. and Bidder, T.G. (1988) 'Perfect Pitch', *American Journal of Medical Genetics* 29: 763–71.
Pruett, K.D. (1987) 'A Longitudinal View of the Musical Gift', *Medical Problems of Performing Artists* 2: 31–8.
Rakowski, A. and Morawska-Bungeler, M. (1987) 'In Search of the Criteria for Absolute Pitch', *Archives of Acoustics* 12: 75–87.
Revesz, G. (1953) *Introduction to the Psychology of Music*. London: Longmans Green.
Sergeant, D. (1969) 'Experimental Investigation of Absolute Pitch', *Journal of Research in Music Education* 17: 135–43.
Sergeant, D. and Roche, S. (1973) 'Perceptual Shifts in the Auditory Information Processing of Young Children', *Psychology of Music* 1: 39–48.
Simpson, J. and Huron, D. (1994) 'Absolute Pitch as a Learned Phenomenon: Evidence Consistent with the Hick-Hyman Law', *Music Perception* 12: 267–70.
Stark, E. (1985) 'Perfect Pitch: All in the Family?' *Psychology Today* 19: 10.
Stumpf, C. (1883) *Tonpsychologie*. Leipzig: Hirzel.
Takeuchi, A.H. and Hulse, S.H. (1991) 'Absolute Pitch Judgements of Black and White Key Pitches', *Music Perception* 9: 27–46.
Takeuchi, A.H. and Hulse, S.H. (1993) 'Absolute Pitch', *Psychological Bulletin* 113: 345–61.
Terhardt, E. and Seewann, M. (1983) 'Aural Key Identification and its Relationship to Absolute Pitch', *Music Perception* 1: 63–83.
Torff, B. and Winner, E. (1994) 'Don't Throw the Baby Out with the Bath Water: The Role of Innate Factors in Musical Accomplishment', *The Psychologist* 7: 361–2.

Vernon, P.E. (1977) 'Absolute Pitch: A Case Study', *British Journal of Psychology* 68: 485–9.

Ward, W.D. (1999) 'Absolute Pitch', in D. Deutsch (ed.) *The Psychology of Music*, pp. 265–98, 2nd edn. London: Academic Press.

Ward, W.D. and Burns, E.M. (1982) 'Absolute Pitch', in D. Deutsch (ed.) *The Psychology of Music*, pp. 407–47. London: Academic Press.

NAN BAHR is a lecturer in Education at the University of Queensland. Nan teaches and researches in the field of educational psychology and has published in the areas of music perception and cognition, educational technologies and pedagogy, and learning.
Address: The University of Queensland, Ipswich Campus, School of Education, 11 Salisbury Road, Ipswich, Queensland 4305, Australia.
[email: n.bahr@mailbox.uq.edu.au]

CAROL CHRISTENSEN is a senior lecturer teaching educational psychology at the University of Queensland. Her research focuses on students' learning. She is particularly interested in understanding why students fail to learn and what can be done to address their difficulties. She has published widely on student learning from both psychological and social perspectives.
Address: The University of Queensland, Ipswich Campus, School of Education, 11 Salisbury Road, Ipswich, Queensland 4305, Australia.

MARK BAHR is a cognitive psychologist with interests in the acquisition of skill and research methods. He has worked on two of the largest educational reform projects in Queensland.
Address: The University of Queensland, Ipswich Campus, School of Education, 11 Salisbury Road, Ipswich, Queensland 4305, Australia.