Voice pitch preferences of adolescents: Do changes across time indicate a shift towards potentially adaptive adult-like preferences?


aUniversity of Liverpool, Liverpool, L69 7ZB, UK.
bUniversity of Glasgow, Glasgow, G12 8QQ, UK. lisa.debruine@glasgow.ac.uk; ben.jones@glasgow.ac.uk.
cUniversity of Stirling, Stirling, FK9 4LA, UK. anthony.little@stir.ac.uk; craig.roberts@stir.ac.uk.

*1present address/corresponding author: Northumbria University, Newcastle, NE1 8ST, UK. tamsin.saxton@northumbria.ac.uk. Tel: 0191 2274486. Fax: 0191 2273903

Abstract: An evolutionary approach to attractiveness judgments emphasises that many human trait preferences exist in order to assist adaptive mate choice. Here we test an adaptive development hypothesis, whereby voice pitch preferences indicating potential mate quality might arise or strengthen significantly during adolescence (when mate choice becomes adaptive). We used a longitudinal study of 250 adolescents to investigate changes in preference for voice pitch, a proposed marker of mate quality. We found significantly stronger preferences for lower-pitched opposite-sex voices in the older age group compared with the younger age group (using different sets of age-matched stimuli), and marginally increased preferences for lower-pitched opposite-sex voices comparing within-participant preferences for the same set of stimuli over the course of one year. We also found stability in individual differences in preferences across adolescence: controlling for age, the raters who had stronger preferences than their peers for lower-pitched voices when first tested, retained stronger preferences for lower-pitched voices relative to their peers about one year later. Adolescence provides a useful arena for evaluating adaptive hypotheses and testing the cues that might give rise to adaptive behaviour.

Highlights:

- We hypothesised that preferences for mate quality might arise during adolescence
- Voice pitch is a purported marker of mate quality
- We assessed voice pitch preferences twice one year apart in 250 adolescents
- Older age was linked to stronger preferences for lower-pitched voices
- Preferences for low pitch were consistent relative to peers across a year
1 Introduction

Human trait preferences are thought to exist in order to assist with adaptive mate choice, directing individuals towards potential partners of suitable genetic quality and compatibility (Roberts & Little, 2008). Some of these preferences change dynamically to improve the adaptive fit of an individual’s behaviour to the circumstances. For example, women prefer putative markers of genetic quality most at the time in the ovulatory cycle when they are able to conceive (see e.g. Jones et al., 2008a), and men who are subject to cues that food is scarce prefer heavier women (Nelson & Morrison, 2005). The study of these sorts of individual differences is useful to test whether the adaptive paradigm can help explain not just human behaviour in general, but also fine details of individual behaviour (Buss & Greiling, 1999). Following the logic of adaptive individual variation, if many human trait preferences exist to help us choose a suitable partner, we might posit an adaptive development hypothesis, where preferences arise or strengthen substantially only as partner choice becomes relevant. This might be true particularly if the trait under consideration is valuable in a mate but of neutral or negative value in other contexts, so the net payoff of affiliation with the target individual depends on whether mate quality is relevant or not (c.f. fluctuating preferences across the ovulatory cycle, reviewed in Jones et al., 2008a).

In this study, we examine preferences for sexually dimorphic voice pitch, because voice pitch is proposed to be an indicator of genetic quality used in human mate choice. Men prefer higher-pitched voices in women (e.g. Feinberg et al., 2005; Jones, Feinberg, DeBruine, Little & Vukovic, 2008b), and women tend to prefer lower-pitched male voices (e.g. Collins, 2000; Saxton, Caryl & Roberts, 2006); such preferences are stronger when judging opposite-sex than same-sex voices, supporting a degree of mate choice specialisation (Jones, Feinberg, DeBruine, Little & Vukovic, 2010b). Pitch preferences are also more pronounced when mate choice based on genetic quality may be particularly apposite, such as when women make judgments at the fertile phase of the menstrual cycle or consider a short-term relationship (Feinberg et al., 2006; Puts, 2005). Lower voice pitch corresponds to higher male reproductive success in a natural population (Apicella, Feinberg & Marlowe, 2007), and to higher testosterone levels (Dabbs & Mallinger, 1999). The female voice is shaped during puberty under the influence of estrogens and progesterone (Abitbol, Abitbol & Abitbol, 1999), and women with higher-pitched voices tend to have more attractive and more feminine faces (Feinberg et al., 2005).

While a choice for sexually-dimorphic pitch might be adaptive in adulthood, the same might not be true in childhood. In adults, higher levels of testosterone are linked to lower levels of family-oriented behaviour and higher levels of aggression (review in e.g. ...
Mazur & Booth, 1998). Indeed, infants tested on female voices preferred higher pitch (Fernald & Kuhl, 1987), and both men and women tend to raise their pitch when talking to infants (Kitamura, Thanavishuth, Burnhama & Luksaneeyanawin, 2002; Warren-Leubecker & Bohannon, 1984). Although full reproductive capacity does not emerge until well beyond childhood, retrospective reports of first sexual attraction identify the age of about 10 years (McClintock & Herdt, 1996). Accordingly, functionally-motivated research on the development of attractiveness judgments has tended to focus on individuals from late childhood to adolescence. Previous research has suggested that pitch preferences change within this age range. At age 7-10, girls who judged the attractiveness of unmanipulated adult male voices did not show any significant directional pitch preferences; the standard adult preference for lower-pitched adult male voices was only apparent from groups of adolescent (aged 12 – 15) and adult raters (Saxton et al., 2006).

We previously reported the first study of adolescents’ evaluations of manipulated voice pitch (Saxton, DeBruine, Jones, Little & Roberts, 2009). In that study, two groups of adolescents, one group aged around 11 years and one group aged around 13 years, made forced-choice preference judgments of opposite-sex voices that had been manipulated for pitch. Girls aged around 13 years demonstrated significantly stronger preferences for lower-pitched boys’ voices than did girls aged around 11 years (whose preferences did not differ from chance), and boys aged around 11 years demonstrated significantly stronger preferences for higher-pitched girls’ voices than boys aged around 13 years did. In the same study, girls who were further through puberty showed greater preferences for lower-pitched boys’ voices, which is consistent with findings that increasing testosterone levels in adulthood are linked to increases in preferences for sexually-dimorphic faces (Welling et al., 2007; Welling et al., 2008). Our previous study compared younger and older raters in a between-subjects design, and used a different set of voice stimuli for the younger and older raters to allow for use of age-matched stimuli, in order to maximize ecological validity and relevance. The present study therefore set out to clarify the trajectory of circum-pubertal changes in preferences, while controlling for possible differences arising from cohort and stimulus sets, by returning to the same adolescents in a second round of data collection approximately one year after the first.

2 Material and Methods

2.1 Participants
Pupils (n=325) were initially recruited from the first and third year of British secondary education (i.e. admitting children aged around 11, and around 13) of a set of British private schools charging similar levels of school fees. The attractiveness judgment tests were repeated with as many of the same pupils as were available plus a number of their classmates in a second data collection round between nine and 13 months after the first test.
Age data are in Table 1. Children who only took part in one of the two years because of class changes or absence are included in the analyses where possible to maximize sample size and because the hypotheses concern the effects of age and puberty in general and should be independent of the specific participants. The study was approved by the University of Liverpool Research Ethics Committee for Non-Invasive Procedures, and carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

Table 1: Participant ages. *19 children did not provide age details

<table>
<thead>
<tr>
<th>Age group:</th>
<th>First round of data collection</th>
<th>Second round of data collection*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Younger</td>
<td>Older</td>
</tr>
<tr>
<td>n (males, females)</td>
<td>84, 74</td>
<td>93, 74</td>
</tr>
<tr>
<td>Mean age (years:months), +/-SD (months)</td>
<td>11:9 +/-5</td>
<td>13:11 +/-6</td>
</tr>
</tbody>
</table>

2.2 Stimuli
Stimuli are the same as those used in the original study (Saxton et al., 2009), where full details of their creation are given. Voice stimuli were created from recordings of six male and six female native-English speakers aged 11-13 (used as stimuli for the younger participants) and six male and six female native-English speakers aged 13-15 (used as stimuli for the older participants) reciting four vowel sounds (/oʊ/ (as in “go”), /u/ (as in “soon”), /ʌ/ (as in “bar”), and /i/ (as in “see”). Recordings were filtered to reduce nonvocal noise in the file, and standardised for amplitude, vowel length, and between-vowel silence (further details in Saxton et al., 2009). Each recording was manipulated to increase fundamental frequency (perceived as vocal pitch) by 20 Hz, and paired against the same voice manipulated to decrease fundamental frequency by 20 Hz. Manipulations were carried out using the PSOLA method (Praat 4.4.24, www.praat.org). Details of the original voice recordings are in Table 2.

Table 2: Details of voices used to create stimuli

<table>
<thead>
<tr>
<th>Voices:</th>
<th>Younger</th>
<th>Younger</th>
<th>Older</th>
<th>Older</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>female</td>
<td>male</td>
<td>female</td>
<td>male</td>
</tr>
<tr>
<td>Mean +/-SD of f0, in Hz</td>
<td>216 +/-19</td>
<td>202 +/-21</td>
<td>223 +/-23</td>
<td>137 +/-9</td>
</tr>
<tr>
<td>Range of individual mean f0s, in Hz</td>
<td>168-240</td>
<td>136-231</td>
<td>211-238</td>
<td>107-171</td>
</tr>
</tbody>
</table>
2.3 Procedure
Pupils listened to pairs of opposite-sex voices which were identical except for fundamental frequency in order to measure pitch preferences, and indicated which voice in each pair was more attractive. Following previous work on adolescents’ attractiveness judgments (Saxton et al., 2006; Saxton et al., 2009), no definition of ‘attractive’ was given to the pupils. The age of acquisition for the word “attractive” is cited as 8.58 years (Kuperman, Stadthagen-Gonzalez & Brysbaert, 2012), and Connolly, Slaughter and Mealey (2004) recommended using the word ‘attractive’ when seeking judgments from older adolescents (and used it for participants aged 12 upwards) because it allows comparison with the literature on adult judgments, should be comprehensible, and has potential sexual connotations which are appropriate to the research context. The order of presentation of voices within pairs, and the order of presentation of pairs, was randomised. We calculated the number of times a child selected a higher-pitched female voice or lower-pitched male voice as more attractive than its pair. Children rated the stimuli either at an individual computer (n=223 and 191 in the first and second round of data collection, respectively) or provided pen-and-paper ratings of stimuli presented through a classroom stereo system (n=102 and 121 in the first and second round of data collection) depending on the school facilities available. The voices were presented in the same way for each child who participated in both rounds. The younger group of children rated the voices of the individuals aged 11-13 in both rounds, while the older children rated the voices of the individuals aged 13-15 in both rounds. Two children entered an unrealistic date of birth, and three demonstrated extreme side bias in the first round; their data were excluded. One participant’s score in Round 2 was calculated from 5 rather than 6 ratings because of a missing data point. If fewer than five voice judgments had been collected, the participant’s data were excluded (n=6 in Round 1 and n=13 in Round 2).

Following the rating task, the participants completed a questionnaire asking demographic data including details of pubertal development. In boys, self-report of pubertal development was scored on a 0-2 scale, with one point for self-report of body hair and one point for self-report of voice change; in girls, one point was awarded for self-report of greater physical development than their peers and one point for the attainment of menarche. Puberty data were excluded due to incomplete answers (n=48) or inconsistency in answers between the two rounds of data collection (n=12).

2.4 Data
Voice judgment data consisted of the proportion of times the participants selected in one direction (i.e. lower-pitched or higher-pitched voices). Male and female voice judgments were analysed separately because raters only rated opposite-sex voices. Data sets were non-normally distributed, but t-tests are robust to this (Subrahmaniam, Subrahmaniam &
Messeri, 1975), and sample sizes were large enough to use parametric correlations (Field, 2009). Statistical analysis was carried out in SPSS Statistics 19.

3 Results and Discussion

The older girls preferred the lower-pitched boys’ voices significantly more often than chance, and the younger boys preferred the higher-pitched girls’ voices significantly more often than chance, whereas the preferences of the younger girls and older boys did not differ significantly from chance (one-sample t-tests; Tables 3 and 4).

Next, we carried out two repeated-measures ANOVAs on the proportion of times that the lower- or higher-pitched voices had been selected, for the participants who provided data in both rounds (n=124 m, 126 f), with data collection round as a within-subjects factor and age group as a between-subjects factor, for the boys’ judgments separately from the girls’ judgments. Results are represented in Figure 1. Both boys and girls tended to select more of the lower-pitched voices in the second round of data collection compared with the first, although this was only marginally significant (boys: F1,122=3.48, p=.064, r=.17; girls: F1,124=3.52, p=.063, r=.17). That is, any pitch preference change over the course of one year was not dramatic, and effect sizes were small, even in a study which focussed on pitch by presenting stimuli that differed in pitch alone. Small biases towards different pitches could accumulate in everyday life given the large number of social interactions that take place every day, although any such biases must be integrated with biases for other vocal aspects (such as speech rate, intonation, and so on), and indeed with other interpersonal attributes which could well have greater weight in attractiveness perceptions. Overall, the older boys preferred the higher-pitched girls’ voices less than the younger boys did (F1,122=8.21, p=.005, r=.25), and the older girls preferred the lower-pitched boys’ voices more than the younger girls did (F1,124=59.39, p<.001, r=.57), although the older and younger groups judged different sets of age-matched stimuli, which might have provoked different strength responses.

Men tend to prefer higher-pitched female voices, but here boys showed an increasing preference with age for lower-pitched female voices. This seems unlikely to be caused by the stimuli having youthfully high voice pitches because previous research has shown that men prefer higher-pitched women’s voices even when those voices have been manipulated up to 280 or 300 Hz (Borkowska & Pawlowski, 2011; Feinberg, DeBruine, Jones & Perrett, 2008; Re, O’Connor, Bennett & Feinberg, 2012); the individual means of the manipulated higher-pitch female voices in the present study ranged from 188 to 260 Hz (see Stimuli), and so we might expect them to have provoked standard attractiveness judgments. Yet boys’ and girls’ voices tend to lower in pitch during the teenage years (Lee, Potamianos & Narayanan, 1999), and women perceive lower-pitched male voices to be
Table 3: One-sample t-tests to determine whether male adolescents chose higher-pitched female voices as more attractive significantly more often than chance. Numbers represent all of the participants, and differ slightly from those in Figure 1 which represents only those who took part in both rounds of data collection.

<table>
<thead>
<tr>
<th>Girls’ voices rated by</th>
<th>t</th>
<th>df</th>
<th>p</th>
<th>Mean % of higher-pitched female voices selected</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Younger boys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First round</td>
<td>5.68</td>
<td>83</td>
<td>&lt;.001</td>
<td>66</td>
<td>0.6</td>
</tr>
<tr>
<td>Second round</td>
<td>5.28</td>
<td>93</td>
<td>&lt;.001</td>
<td>62</td>
<td>0.5</td>
</tr>
<tr>
<td>Older boys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First round</td>
<td>1.43</td>
<td>92</td>
<td>.157</td>
<td>54</td>
<td>0.1</td>
</tr>
<tr>
<td>Second round</td>
<td>&lt;.01</td>
<td>48</td>
<td>&gt;.99</td>
<td>50</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4: One-sample t-tests to determine whether female adolescents chose lower-pitched male voices as more attractive significantly more often than chance. Numbers represent all of the participants, and differ slightly from those in Figure 1 which represents only those who took part in both rounds of data collection.

<table>
<thead>
<tr>
<th>Boys’ voices rated by</th>
<th>t</th>
<th>df</th>
<th>p</th>
<th>Mean % of lower-pitched male voices selected</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Younger girls</td>
<td>1.26</td>
<td>73</td>
<td>.211</td>
<td>46</td>
<td>0.1</td>
</tr>
<tr>
<td>First round</td>
<td>1.47</td>
<td>77</td>
<td>.145</td>
<td>54</td>
<td>0.2</td>
</tr>
<tr>
<td>Second round</td>
<td>11.78</td>
<td>73</td>
<td>&lt;.001</td>
<td>76</td>
<td>1.4</td>
</tr>
<tr>
<td>Older girls</td>
<td>11.57</td>
<td>90</td>
<td>&lt;.001</td>
<td>75</td>
<td>1.2</td>
</tr>
</tbody>
</table>

older (e.g. Collins, 2000) (although in fact, in our stimuli, the mean of the younger girls’ voices was slightly lower than the mean of the older girls’ voices, and the lowest individual mean voice pitch came from the younger rather than the older group of girls.) Both male and female adolescents tend to state that an ideal partner would be older than themselves (Kenrick, Keefe, Gabrieldis & Cornelius, 1996). The preferences that we saw in our adolescent participants for lower pitch in opposite-sex voices could thus represent a preference for an older (i.e. more sexually mature) partner. A preference for voices that
sound somewhat older is consistent with our general hypothesis that adaptive mate preferences start to arise during puberty. Further, on the basis that familiar stimuli are more attractive (Zajonc, 1968), the deepening voice pitches of the participants’ peers across time could lead to an apparent change with age towards a greater preference for lower voice pitch. Finally, female mate value as indexed by fecundity and residual reproductive value is highest fairly shortly after puberty, whereas male mate value as indexed by resource acquisition ability is higher later in life (e.g. Buss, 1989). The pitch of the girls’ voices in our study approximated that of women, whereas the pitch of the boys’ voices was higher than that of men; that is, female voice stimuli pitches were similar to those of women of peak reproductive value, whereas the male voice stimuli pitches in the study were not. This might contribute to the clearer preferences we found for lower-pitched male voices than for higher-pitched female voices.

Figure 1. Percentage of times boys selected higher-pitched female voices, and girls selected lower-pitched male voices as more attractive, contrasting first and second round of data collection. Here, sample includes only those who participated in both rounds of data collection. Bars = mean ± SE.
There was no significant interaction between the age group of the rater and the round of data collection, meaning that any difference in preferences between the first and second round of data collection for the younger age group was roughly similar to any preference differences between the first and second round of data collection for the older group, for boys ($F_{1,122}=0.04, p=.851$) or girls ($F_{1,124}=2.36, p=.127$).

We previously found a relationship between girls’ greater pubertal development and preference for lower-pitched boys’ voices in the first round of data collection (Saxton et al., 2009). Pubertal development scores (see Methods section) in the second round of data collection were very unequal in the older groups, and so analysis was restricted to the younger group (younger boys: $n=16, 30$ and $25$ and younger girls: $n=17, 16$ and $19$ scoring 0, 1 and 2 points respectively). Pubertal development was not linked to preference in the younger girls ($F_{2,48}=2.01, p=.145$) or younger boys ($F_{2,67}=.28, p=.754$; ANCOVA on pitch judgments with the pubertal categories as independent variable, and age in years and months as covariate). The discrepancy from the first round might arise from the smaller sample in the second round of data collection, links between pubertal development and preferences that are transient during adolescence or that interact in a complex way with age-dependent preferences, or the approximate assessment of pubertal development. Long-standing links between puberty and women’s voice pitch preferences have been demonstrated elsewhere: women who reported earlier menarche showed a stronger preference for lower-pitched male voices (Jones, Boothroyd, Feinberg & DeBruine, 2010a). We did not find this relationship in the present sample – there was no relationship between the girls’ reported age of menarche and their preference for lower-pitched voices in the younger ($r=-.189, n=26, p=.355$) or older ($r=-.038, n=52, p=.786$) girls – although our participants were younger and our sample was smaller than the 104 participants of the Jones et al. (2010a) study. Future research might investigate specific endocrine variables as predictors of preferences in adolescents.

At the individual level, there were similarities in voice pitch preference strength between the two data collection rounds for three of our four groups. That is, the younger boys, younger girls, and older girls who preferred lower-pitched voices relative to their peers in the first round tended to maintain that relatively stronger preference approximately one year later (partial correlation analyses of the number of lower-pitch voices selected as more attractive in the two rounds of data collection, controlling for age: younger boys: $r=.369, n=75, p=.001$; younger girls: $r=.378, n=62, p=.002$; older girls: $r=.263, n=58, p=.043$). The relationship was not significant for the older boys (the smallest sample size; older boys: $r=.101, n=43, p=.511$). This indicates that in adolescence, as in adulthood (see e.g. Jones et al., 2008a; Roberts & Little, 2008), there may be individual differences that systematically and consistently affect pitch preferences and that remain constant across
time. Some systematic predictors of preference rely on an evaluation of one’s own mate value (e.g. Vukovic et al., 2008), and future research might investigate whether this is apparent even in adolescents.

The ontogeny and long-term stability of individual variation in judgments of attractiveness remain an important topic for future research. Firstly, it can help determine the extent to which we should consider individual variation to be adaptive. Some individual differences are clearly adaptive (c.f. e.g. Buss, 2009), but at a certain level of detail, we will want to stop saying that individual differences have adaptive explanations, and start talking of nonadaptive or even maladaptive differences (Buss & Greiling, 1999). The study of detailed individual differences and their ontogeny provides two dimensions of variation (one between participants and one within participants), and as such can help us understand the extent to which the adaptive paradigm is usefully predictive of differences, and of the fine details of human behaviour. Secondly, it can help us understand the specific proximate mechanisms that provoke adaptations. For example, external environmental cues (e.g. Little, DeBruine & Jones, 2011) and internal developmental processes (Little et al., 2010; Vukovic et al., 2009) have both been linked to preferences for facial sexual dimorphism. Thirdly, human psychology has been shaped by past regularities (Daly & Wilson, 1999). In contrast, adolescence as we know it in modern Western societies, with for example a long period of extensive financial reliance on parents, education rather than work, and lack of marriage, is not a consistent feature of human societies (Schlegel & Barry, 1991). Accordingly, changes during adolescence provide an insight into whether our preferences are sufficiently robust and/or flexible to remain adaptive (see e.g. Smith, Mulder & Hill, 2001) despite the novelties of the environment.

4 Acknowledgements

The authors thank Andrew Boardman, Kristina Gilbertson, Gideon Gluckman, Shelly Kemp, Robert Kennedy, Jenny Saxton, Thom Scott-Phillips, Sue Toole (help with data collection), Dave Feinberg (advice on voice stimuli creation), and the schools and social groups who provided much-valued participants. ACL is supported by a Royal Society University Research Fellowship. TKS was supported by a University of Liverpool Studentship and Postdoctoral Research Fellowship PTA-026-27-2007 from the Economic and Social Research Council.

5 References


