This is a revised personal version of the text of:
Saxton, T. K., Kohoutova, D., Roberts, S. C., Jones, B. C., DeBruine, L.M. & Havlicek, J. (2010). Age, puberty and attractiveness judgments in adolescents. *Personality And Individual Differences, 49,* 857 – 862.
The published version can be found online at http://dx.doi.org/10.1016/j.paid.2010.07.016.

Age, puberty and attractiveness judgments in adolescents

Tamsin K. Saxton* [a], Dagmar Kohoutova [b], S. Craig Roberts [c], Benedict C. Jones [d], Lisa M. DeBruine [d], Jan Havlicek [b]

[a] Philosophy, Psychology and Language Sciences, University of Edinburgh, UK

[b] Department of Anthropology, Faculty of Humanities, Charles University in Prague, Czech Republic

[c] School of Biological Sciences, University of Liverpool, UK

[d] School of Psychology, University of Aberdeen, UK

* corresponding author. Email tamsin.saxton@ed.ac.uk. Tel: +44 0131 6502907 Fax: +44 0131 6513190

Previous work has suggested that judgments of the attractiveness of some facial and vocal features change during adolescence. Here, over 70 Czech adolescents aged 12 – 14 made forced-choice attractiveness judgments on adolescent faces manipulated in symmetry, averageness and femininity, and on adolescent opposite-sex voices manipulated in fundamental frequency (perceived as pitch), and completed questionnaires on pubertal development. Consistent with typical adult judgments, adolescents selected the symmetric, average and feminine male and female faces as more attractive significantly more often than the asymmetric, masculine and non-average faces respectively. Moreover, preferences for symmetric faces were positively associated with adolescents' age and stage of pubertal development. Unexpectedly, voice pitch did not significantly influence adolescents' attractiveness judgments. Collectively, these findings present new evidence using refined methodology that adolescent development is related to variation in attractiveness judgments.

Introduction

Much research has demonstrated the importance of physical attractiveness in human behaviour (review in e.g. Langlois et al., 2000). Attractiveness affects a diverse range of social interactions, ranging from relationship initiation to attributions of personality traits to beliefs about competence (see e.g. Eagly et al., 1991; Roberts & Little, 2008). Children are by no means exempt from the influences of attractiveness: children are aware of relative attractiveness from a young age, tend to agree with adults about relative attractiveness, and make use of perceptions of physical attractiveness in their behaviour (e.g. Cavior & Lombardi, 1973; Cross & Cross, 1971; Dion, 1973; Dion & Berscheid, 1974; Kleck et al., 1974).

Attractiveness judgments are thought to reflect mate preferences at least in part, helping individuals to identify potential partners of relatively higher biological quality and suitability (see e.g. Roberts & Little, 2008). Accordingly, attractiveness judgments might be expected to differ across the life span because mate choice is more relevant during some stages of life (e.g., following puberty) than it is during others (e.g., prior to puberty, Little et al., 2010). Stimuli can be objectively manipulated to differ in the physical parameters that are thought to provide information on the quality of a potential partner, and these manipulations have systematic influences on adults' attractiveness judgments (see e.g. Rhodes, 2006; Roberts & Little, 2008). Manipulations can be used to alter indicators of hormonal profile (e.g., sexually dimorphic shape cues, waist-to-hip ratio and voice pitch) or developmental stability (e.g., prototypicality and symmetry) (see Roberts & Little, 2008). Adults tend to give higher ratings of attractiveness to women whose waist is around one third smaller than their hips, and a study of participants who varied in age from six years old to adulthood found that this standard adult preference developed approximately linearly during childhood and adolescence (Connolly et al., 2004). Additionally, facial masculinity is preferred more by women in their reproductive years, and less by women before the completion of puberty or after the menopause (Little et al., 2010; see also Vukovic et al., 2009). Another study found that preferences for facial averageness, male facial symmetry, feminised male faces (when judged by girls but not boys), and lower-pitched opposite-sex voices each increased with age during puberty (Saxton et al., 2009). Finally, a study comparing female children, adolescents and adults found that only the ratings from the latter two groups gave rise to significant correlations between the rated attractiveness of a man's face compared to his voice, and that only the latter two groups demonstrated a preference for lower-pitched men's voices (Saxton et al., 2006).

Research on the development of adolescents' attractiveness judgments has also investigated the relationships between individual differences in face and voice preferences and the stages of normal pubertal development. This follows findings that individual differences in adult attractiveness judgments can be linked to individual differences in hormonal profile (e.g. Jones et al., 2008; Puts, 2006; Welling et al., 2007) and that adolescent biological development corresponds to levels of sexual behaviour in adolescence (Halpern et al., 1993; McClintock & Herdt, 1996; Udry, 1988; Udry et al., 1985). Controlling for possible effects of age, pubertal development in adolescents is correlated with boys' preferences for male facial masculinity and girls' preferences for male vocal masculinity (Saxton et al., 2009). In contrast, age, rather than physical development (own waist-to-hip ratio, height, weight, body mass index), is more important for variation in adolescents' preferences for women's waist-to-hip ratios (Connolly et al., 2004).

These earlier studies on attractiveness judgments and puberty used self-report measures of various facets of physical development (Saxton et al., 2009), or measurements of waist-to-hip ratio, height, weight and body mass index (Connolly et al., 2004), to capture biological development during adolescence. However, standardised measures of puberty exist, such as the Pubertal Development Scale (Petersen et al., 1988), which uses self-report of somatic markers of puberty to give an overall picture of pubertal development (Bond et al., 2006; Brooks-Gunn et al., 1987). The current study set out to investigate whether standardised measures of pubertal development during adolescence predicted individual differences in face and voice attractiveness judgments. In addition, previous studies either asked adolescents to rate adult stimuli (Little et al., 2010; Saxton et al., 2006) or contrasted older adolescents' judgments of older adolescent stimuli with younger adolescents' judgments of younger adolescent stimuli (Saxton et al., 2009), but have not yet contrasted judgments by adolescents of different ages on the same set of adolescent stimuli, which was taken up in the present study. A final subsidiary aim of the research was to explore preferences in a population that does not form the subject of much current research, namely Czech adolescents (c.f. Henrich et al., in press, who demonstrate how many of our expectations of psychological universals may be incorrect, and recommend cross-cultural testing).

Methods

Stimuli

All stimuli were taken from Saxton et al (2009), where a fuller description of the methods of stimuli creation can be found. In brief, face stimuli were created on the basis of 60 photographs of Caucasian adolescents aged 11 – 15 (equally divided between male and female; and equally divided between an age group of 11-13 and an age group of 13-15) using the specialist computer graphics software Psychomorph (Tiddeman et al., 2001). Twelve pairs of faces were created that differed only in symmetry: one face was manipulated to increase the bilateral symmetry of the facial features, and one to decrease it. Twelve pairs of faces were created that differed only in averageness: one face was made more average (that is, more similar to the average of the faces making up the group which it came from: i.e. 15 males aged 11 – 13, 15 females aged 11 – 13, 15 males aged 13-15, or 15 females aged 13-15), and paired with the matching unmanipulated face. Finally, twelve pairs of faces were created that differed only in sexual dimorphism: one face was made to look more masculine (i.e. more like the average face shape of 15 boys aged 13 – 15 and less like the average face shape of 15 girls aged 13 – 15) and one was made to look more feminine (the reverse manipulation). Examples of the stimuli manipulations are given in Figure 1. Vocal stimuli consisted of 12 pairs of opposite-sex voices (half aged 11-13 and half aged 13-15) from native English speaking individuals reciting four vowel sounds, standardised in length. Voices within each pair were identical except that one was raised and one lowered by 20 Hz in fundamental frequency (perceived as vocal pitch) using Praat 4.4.24 (Boersma, 2001).

Figure 1. Examples of image manipulation, applied to an adult base face (children's faces are not shown for reasons of consent). Top row: face has been masculinised (left) and feminised (right); middle row: face is original (left) and made more average (right); bottom row: face has been made more asymmetric (left) and more symmetric (right). Image originally published in Saxton et al., 2009.

Raters

Seventy-two raters aged 12 – 14 (12 years: *n*=14 boys/8 girls; 13 years: *n*=17/31; 14 years: *n*=2/0) were recruited from two school entry years (equivalent to the sixth and seventh grades) from three co-educational secondary state schools based in Prague, the Czech Republic, where state schools are the most common method of schooling. The sample consisted of 33 boys (of whom five did not provide voice ratings, one did not provide face ratings, and one did not provide female face ratings for reasons of time) and 39 girls (six of whom did not provide voice ratings). Agreement by schools and children, and written consent from parents, was supplied.

Tests

Tests were performed in the Czech language. Face tests were presented with a Java applet that presented all 72 pairs of faces, randomising presentation side and order. Male faces were blocked separately from female faces. Opposite-sex voices only were rated due to time constraints, as voice rating takes longer than face rating. The software package Powerpoint 2003 (Microsoft) was used to present the voices in pairs. Eight different versions of presentations (half for male voices) were created to minimise order effects. Each version had identical numbers of lower-pitched voices presented first or second within the pair. Children carried out the face and voice tests individually on one of three laptops with headphones, with a screen size of at least 26 x 16 cm and a resolution of at least 1280 x 800 pixels. All children filled out a demographic questionnaire that included questions on pubertal development (axillary hair; facial hair and voice change in boys) adapted from the Pubertal Development Scale (Petersen et al., 1988). The questionnaire also included a self-assessment of pubertal development on the basis of sex-specific line drawings (taken from Taylor et al., 2001). All children self-assessed pubic hair development, and girls self-assessed breast development. Children were assured that they did not need to answer any questions, but that answers would be kept confidential and anonymous. To assure privacy and decrease the risk of potential embarrassment, children were seated individually in a quiet corner of the room away from the other children to complete the questionnaire. No child omitted any answer. The order of the tests (male face test, female face test, opposite-sex voice test, questionnaire) was randomised according to when computers or test stations were available.

Participant data

In order to create roughly equal age and pubertal development categories for the analysis, these were converted to bivariate categories. Children were divided into younger (12 years of age) and older (13 or 14 years of age) groups. Pubertal data were categorised into early or late pubertal, following a formula adapted from Carskadon & Acebo (1993). Girls' self-assessed breast growth and body hair development were included in the formula to allocate children to the different categories, but in fact were not needed in order to distinguish between the categories: in our sample, the girls categorised as early pubertal were all pre-menarchal (n=13); all post-menarchal girls (n=26) were categorised as late pubertal. The boys categorised as late pubertal (n=10) reported facial hair growth, voice change, and axillary hair; the remaining boys (n=23) were categorised as early pubertal.

We calculated six face scores and one voice score for each rater. The six face scores were the proportion of times s/he chose the symmetric, average or feminine male or female face as more attractive than the asymmetric, non-average and masculine face respectively. The voice score was the proportion of times s/he chose the lower-pitched opposite-sex voice as more attractive. Statistical analysis was carried out in SPSS 18.0. Some of the sets of attractiveness judgment data were non-normally distributed (as indicated by the Shapiro-Wilk test, $p < .05$), but t-tests and GLM are fairly robust to violations of the assumption of normal distribution of data (Field, 2005; Stonehouse & Forrester, 1998; Subrahmaniam et al., 1975).

Results

First, we tested the participants' general preferences for each of the manipulations. Irrespective of the sex of the rated face, the raters preferred the symmetric, average and feminine faces (over the asymmetric, original and masculine faces respectively) significantly more often than predicted by chance alone (all one-sample $t > 2.5$, $p < .02$). Raters only rated opposite-sex voices; their preferences for high or low pitch did not differ significantly from chance (boys: $t(27)=.48$, $p=.637$; girls: $t(32)=.71$, $p=.484$).

We tested for effects of age, sex and pubertal development with a GLM-repeated measures analysis of the face scores (within-subjects factor: type of manipulation applied to the face, sex of rated face; between-subjects factors: pubertal category, age category, sex). There was a significant main effect of the type of manipulation ($F(2,124)=3.21$, $p=.044$), but this was modified by interactions with the sex of the rater ($F(2, 124)=3.95$, $p=.022$), and with the pubertal category and the age category ($F(2, 124)=7.56$, $p=.001$). To understand these better, we subsequently carried out three GLM-repeated measures analyses, one for each of the three types of face manipulation. The sex of the rated face was a within-subjects factor, and the pubertal category, age category and sex of the rater were between-subjects factors. There were no significant effects of any of these variables in judgments of the faces that had been manipulated for averageness. In judgments of facial femininity, there was a significant interaction between the age group and the pubertal category of the raters ($F(1,62)=9.39$, $p=.003$): in the younger age group, there was a non-significant tendency for the raters in the late pubertal group to have an increased preference for facial masculinity compared with the raters in the early pubertal group ($F(1,18)=3.53$, $p=.077$), whereas in the older age group, the raters in the late pubertal group had a significantly stronger preference for facial femininity compared with the raters in the early pubertal group ($F(1,44)=6.85$, $p=.012$). In respect of judgments of facial symmetry, raters in the late pubertal category picked more of the symmetric faces than raters in the early pubertal category ($F(1,62)=6.79$, $p=.011$), and older raters picked more of the symmetric faces than younger raters ($F(1,62)=4.56$, $p=.037$). Data are represented in Figures 2 and 3.

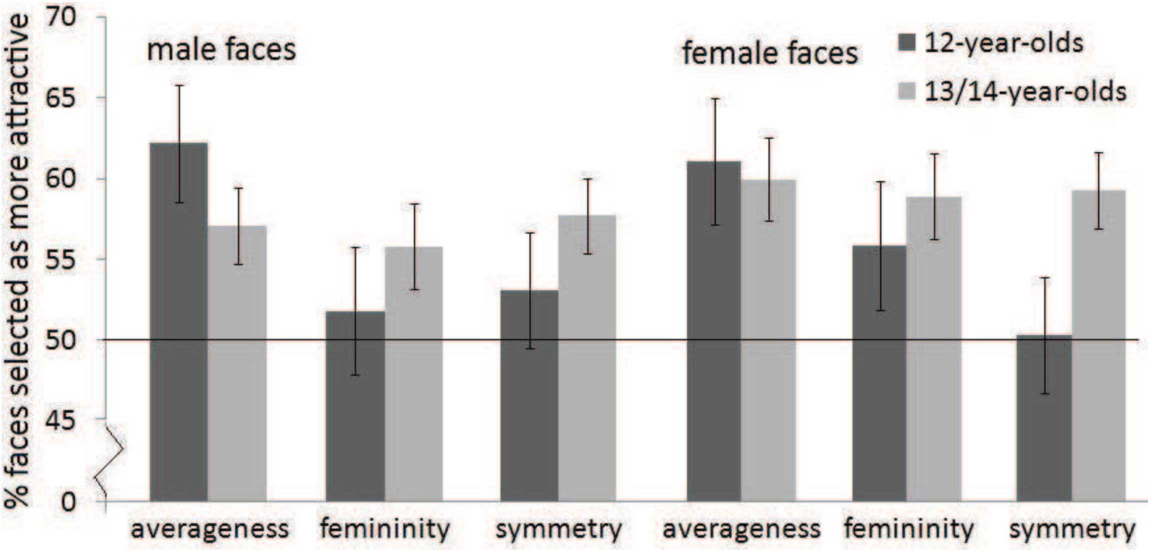

Figure 2. Percentage of average, feminine or symmetric faces selected as more attractive than original, masculine and asymmetric faces respectively in a forced choice judgment task, by pubertal category. Bars=mean ± SE. Chance levels (50%) marked.

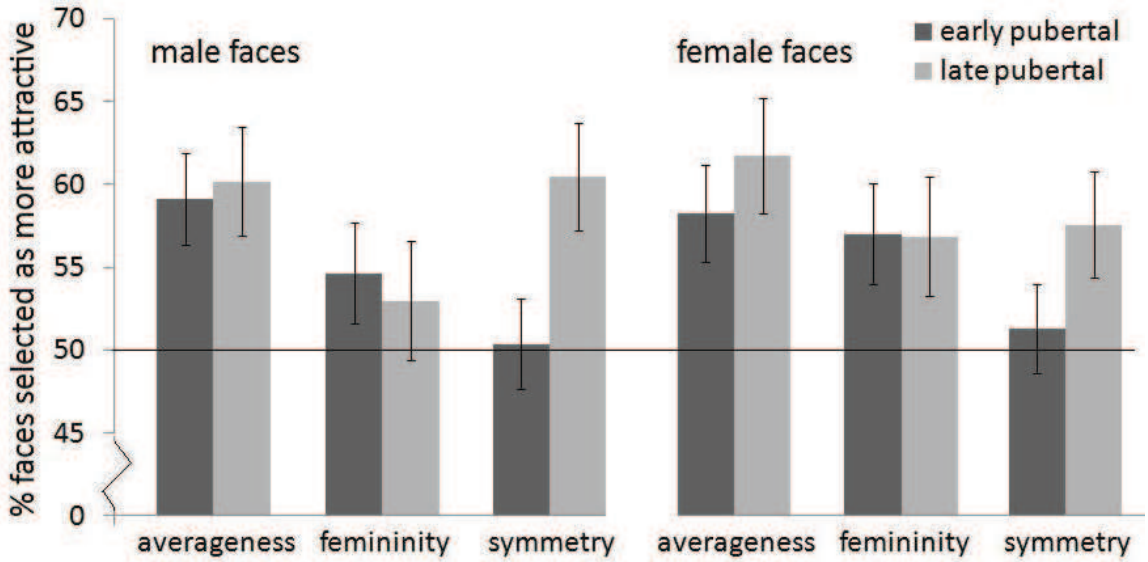

Figure 3. Percentage of average, feminine or symmetric faces selected as more attractive than original, masculine and asymmetric faces respectively in a forced choice judgment task, by age category. Bars=mean ± SE. Chance levels (50%) marked.

Next, we tested for differences in ratings of the different aged stimulus faces, with a GLM-repeated measures analysis that had the sex and age group of the rated face as the two within-subject factors. There was a significant main effect of the age group of the rated face ($F(1,69)=4.26$, $p=.043$), but this was modified by a significant interaction with the sex of the rated face ($F(1,69)=6.13$, $p=.016$) and with the type of manipulation ($F(2,138)=4.28$, $p=.016$). To understand the interaction with the sex of the rated face, we carried out two GLM-repeated measures analyses, using identical factors but separating the male and female faces. Raters were significantly more likely to select the preferred dimensions (i.e. symmetry, averageness and femininity) of the older compared with the younger faces, but only when they were rating male faces ($F(1,70)=8.63$, $p=.004$) and not when they were rating female faces ($F(1,69)=.39$, $p=.532$). To understand the interaction with the type of manipulation, we carried out three GLM-repeated measures analyses, using identical factors but separating the three types of manipulation. We found a significant effect of the age of the rated faces only among the feminised faces: raters were significantly more likely to select the feminised faces as more attractive when they were rating older compared with younger faces ($F(1,69)=8.95$, $p=.004$). Results are qualitatively identical when the age group of the raters is entered into the analysis, and there are no additional interactions between the rater age group and the age group of the rated faces.

Finally, we tested for effects of age and pubertal development with two GLM analyses (one for male and one for female raters) on the opposite-sex voice ratings (between-subjects factors: pubertal category, age category). There was no significant effect of pubertal category, or age category, on boys' judgments of pitch manipulations in girls' voices, or on girls' judgments of pitch manipulations in boys' voices (all $p > .2$). With the same GLM-repeated measures analysis, but separating ratings of the older and younger voices, there was no significant effect of the age of the stimulus voice as an additional within-subjects factor (both $p > .2$).

Discussion

Overall, the Czech adolescents judged the symmetric, average and feminine male and female faces to be more attractive than the asymmetric, less average and masculine male and female faces respectively. This is consistent with judgments made by British adolescents of the same age (Saxton et al., 2009), and broadly consistent with adults' attractiveness judgments, although adults sometimes prefer masculinity in male faces (reviews in e.g. Rhodes, 2006; Roberts & Little, 2008). The manipulations that we applied to the faces reflect dimensions that are thought to allow individuals to judge mate quality, and our results suggest that these dimensions are salient to young adolescents.

The adolescents' biological development (as measured by pubertal category, or by age) was linked to preferences for the symmetric faces. Older children, and children in the late puberty group, preferred the symmetric faces more than younger children or children in the early puberty group did. Similarly, older children had stronger preferences for feminised faces if they were further through puberty. Visual inspection of the data (Figures 2 and 3) suggests that more sensitive measures or a greater sample size might reveal more subtle relationships of a smaller effect size between development and attractiveness judgments. Although we set out to use standardised measures of puberty, we divided our sample into just two developmental groups, and found that in the girls menarche alone was so predictive of other developments that it distinguished between the two developmental groups without the need to refer to the other measurements that we took (see 'Participant data'). It remains to be tested whether late-developers eventually come to resemble early-developers in their preferences for all facial traits, but this should not necessarily be the default assumption; previous work has found that men who commence sexual relationships at a younger age have stronger preferences for feminised female faces (Cornwell et al., 2006), and that women who experience earlier menarche have stronger preferences for lower-pitched male voices (Jones et al., 2010). Although many studies have shown that menstrual cycle phase affects women's perceptions of men's faces, we did not consider menstrual cycle phase in the current study, and note that menstrual cycles are typically anovulatory and/or irregular in the age-group of circum-pubertal girls studied here (Apter & Vihko, 1985).

The raters showed no directional preference for high or low pitch, and there were no effects of biological development on pitch preferences, in contrast with previous links found between increased girls' age or pubertal development and a stronger preference for lower-pitched male voices (Saxton et al., 2009). It may be that the English voice stimuli were not processed by the Czech adolescents as familiar voice stimuli, although Apicella & Feinberg (2009) used English phrases rated by non-English-speaking Hadza people, and found effects of pitch on ratings of attractiveness.

A previous study on the judgments of British adolescents (Saxton et al., 2009) found that older adolescents had stronger preferences than younger adolescents for male and female facial averageness, male facial symmetry, and (when judged by girls but not boys) male facial femininity. However, that study asked older adolescents to rate older faces and younger adolescents to rate younger faces. While this controlled for the possible confounding effect of age differences between rater and stimulus, it did not allow effects of rater age to be distinguished from effects of the age of the stimulus face. In the present study, all of the raters rated all of the faces. In rating male faces, overall the raters were significantly more likely to select the preferred manipulation (symmetry, averageness and femininity) when they were rating older compared with younger faces. If raters pay more attention to more sexually mature male faces, or the preferred manipulation is otherwise easier to assess in the older compared with the younger male faces, this could have contributed to some of the differences in judgments of male faces between older and younger raters in the earlier study (Saxton et al., 2009). In rating male and female faces that varied in sexual dimorphism, the raters were significantly more likely to select the feminine faces when they were rating older compared with younger faces, despite the same manipulation being applied to the older and younger male face stimuli. It might be that the femininity manipulation interacted differently with the younger compared with the older male faces: for instance, masculinised male faces tend to appear older (Boothroyd et al., 2005), so it might be that feminisation of the older faces, and/or masculinisation of the younger faces, made them appear closer to the raters' ages, and that this was considered more attractive. Alternatively, facial characteristics other than masculinity-femininity, such as apparent health, can influence the extent to which exaggerated sex-typical cues are preferred in faces (Smith et al., 2009). It is possible that stimuli from the different age groups differed on such factors.

Despite our focus on mate choice, some of the appeal of facial symmetry, averageness and sexual dimorphism may well have other roots. For instance, averageness is attractive in other animate and non-animate objects, namely, dogs, birds and wristwatches (Halberstadt & Rhodes, 2000), perhaps partially because average objects are easier to process (Rhodes, 2006; Winkielman et al., 2006). Accordingly, we might expect averageness in particular to be attractive to younger children if increased ease of processing is associated with preferences at all ages (but see e.g. Rhodes et al., 2002). Further, it may be beneficial to associate oneself with attractive friends for social rather than direct mate choice benefits: attractive children tend to be more popular (Dion & Berscheid, 1974), considered more favourably by teachers (Clifford & Walster, 1973), and have higher sociometric status (Kleck et al., 1974), for instance. However, increased age of the rater, and of the rated face, both affected attractiveness judgments, and we would predict that changes in judgments of peers' faces may be particularly acute during adolescence. The developmental profile of preferences for facial characteristics from infancy to adulthood has yet to be plotted.

Acknowledgments

TKS was supported by a Liverpool University studentship and ESRC postdoctoral research fellowship PTA-026-27-2007; DK by an Owen Aldis Fellowship; and JH by grant GACR 406/09/0647. We thank our participants, and the anonymous reviewers for their helpful comments.

References

Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter–gatherers. *Proceedings of the Royal Society B: Biological Sciences, 276*(1659), 1077-1082.

Apter, D., & Vihko, R. (1985). Hormonal patterns of the first menstrual cycles. In S. Venturoli, C. Flamigni & J. R. Givens (Eds.), *Adolescence in Females* (pp. 215-238). Chicago: Year Book Medical Publishers.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341-345.

Bond, L., Clements, J., Bertalli, N., Evans-Whipp, T., McMorris, B. J., Patton, G. C., et al. (2006). A comparison of self-reported puberty using the Pubertal Development Scale and the Sexual Maturational Scale in a school-based epidemiologic study. *Journal of Adolescence, 29*, 709-720.

Boothroyd, L., Jones, B. C., Burt, D. M., Cornwell, R. E., Little, A. C., Tiddeman, B. P., et al. (2005). Facial masculinity is related to perceived age but not perceived health. *Evolution and Human Behavior, 26*(5), 417-431.

Brooks-Gunn, J., Warren, M. P., Rosso, M. P., & Gargiulo, J. (1987). Validity of self-report measures of girls' pubertal status. *Child Development, 58*, 829-841.

Carskadon, M. A., & Acebo, C. (1993). A self-administered rating scale for pubertal development. *Journal of Adolescent Health, 14*(3), 190-195.

Cavior, N., & Lombardi, D. A. (1973). Developmental aspects of judgment of physical attractiveness in children. *Developmental Psychology, 8*(1), 67-71.

Clifford, M. M., & Walster, E. (1973). The effect of physical attractiveness on teacher expectations. *Sociology of Education, 46*, 248-258.

Connolly, J. M., Slaughter, V., & Mealey, L. (2004). The development of preferences for specific body shapes. *Journal of Sex Research, 41*(1), 5-15.

Cornwell, R. E., Law Smith, M. J., Boothroyd, L. G., Moore, F. R., Davis, H. P., Stirrat, M., et al. (2006). Reproductive strategy, sexual development and attraction to facial characteristics.

Philosophical Transactions of the Royal Society B: Biological Sciences 361(1476), 2143–2154.

Cross, J. F., & Cross, J. (1971). Age, sex, race and the perception of facial beauty. *Developmental Psychology, 5*(3), 433-439.

Dion, K. K. (1973). Young children's stereotyping of facial attractiveness. *Developmental Psychology, 9*(2), 183-188.

Dion, K. K., & Berscheid, E. (1974). Physical attractiveness and peer perception among children. *Sociometry, 37*(1), 1-12.

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but... A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin, 110*, 109-128.

Field, A. (2005). *Discovering Statistics Using SPSS* (2nd ed.). London: SAGE Publications Ltd.

Halberstadt, J., & Rhodes, G. (2000). The attractiveness of nonface averages: Implications for an evolutionary explanation of the attractiveness of average faces. *Psychological Science, 11*(4), 285-289.

Halpern, C. T., Udry, J. R., Campbell, B., & Suchindran, C. (1993). Testosterone and pubertal development as predictors of sexual activity: A panel analysis of adolescent males. *Psychosomatic Medicine, 55*(5), 436-447.

Henrich, J., Heine, S. J., & Norenzayan, A. (in press). The weirdest people in the world. *Behavioral and Brain Sciences*.

Jones, B. C., Boothroyd, L., Feinberg, D. R., & DeBruine, L. M. (2010). Age at menarche predicts individual differences in women's preferences for masculinized male voices in adulthood. *Personality And Individual Differences, 48*(7), 860-863.

Jones, B. C., DeBruine, L. M., Perrett, D. I., Little, A. C., Feinberg, D. R., & Law Smith, M. J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior, 37*, 78-84.

Kleck, R. E., Richardson, S. A., & Ronald, L. (1974). Physical Appearance Cues and Interpersonal Attraction in Children. *Child Development, 45*(2), 305-310.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126*(3), 390-423.

Little, A. C., Saxton, T. K., Roberts, S. C., Jones, B. C., DeBruine, L. M., Vukovic, J., et al. (2010). Women’s preferences for masculinity in male faces are highest during reproductive age-range and lower around puberty and post-menopause. *Psychoneuroendocrinology, 35*(6), 912-920.

McClintock, M. K., & Herdt, G. (1996). Rethinking puberty: The development of sexual attraction. *Current Directions In Psychological Science, 5*(6), 178-183.

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence, 17*, 117-133.

Puts, D. A. (2006). Cyclic variation in women's preferences for masculine traits: Potential hormonal causes. *Human Nature, 17*(1), 114-127.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199-226.

Rhodes, G., Geddes, K., Jeffery, L., Dziurawiecz, S., & Clark, A. (2002). Are average and symmetric faces attractive to infants? Discrimination and looking preferences. *Perception, 31*, 315-321.

Roberts, S. C., & Little, A. C. (2008). Good genes, complementary genes and human mate choice. *Genetica, 132*, 309-321.

Saxton, T. K., Caryl, P. G., & Roberts, S. C. (2006). Vocal and facial attractiveness judgments of children, adolescents and adults: the ontogeny of mate choice. *Ethology, 112*, 1179-1185.

Saxton, T. K., DeBruine, L. M., Jones, B. C., Little, A. C., & Roberts, S. C. (2009). Face and voice attractiveness judgments change during adolescence. *Evolution and Human Behavior, 30*(6), 398-408.

Smith, F. G., Jones, B. C., DeBruine, L. M., & Little, A. C. (2009). Interactions between masculinity-femininity and apparent health in face preferences. *Behavioral Ecology, 20*(2), 441-445.

Stonehouse, J. M., & Forrester, G. J. (1998). Robustness of the t and U tests under combined assumption violations. *Journal of Applied Statistics, 25*(1), 63 - 74.

Subrahmaniam, K., Subrahmaniam, K., & Messeri, J. Y. (1975). On the robustness of some tests of significance in sampling from a compound normal population. *Journal of the American Statistical Association, 70*(350), 435.

Taylor, Whincup, Hindmarsh, Lampe, Odoki, & Cook. (2001). Performance of a new pubertal self-assessment questionnaire: a preliminary study. *Paediatric and Perinatal Epidemiology, 15*(1), 88-94.

Tiddeman, B., Burt, D. M., & Perrett, D. (2001). Computer graphics in facial perception research. *IEEE Computer Graphics and Applications, 21*(5), 42-50.

Udry, J. R. (1988). Biological predispositions and social control in adolescent sexual behavior. *American Sociological Review, 53*(5), 709-722.

Udry, J. R., Billy, J. O. G., Morris, N. M., Gross, T. R., & Raj, M. H. (1985). Serum androgenic hormones motivate sexual behavior in boys. *Fertility and Sterility, 43*(1), 90-94.

Vukovic, J., Jones, B. C., DeBruine, L. M., Little, A. C., Feinberg, D. R., & Welling, L. L. M. (2009). Circum-menopausal effects on women's judgements of facial attractiveness. *Biology Letters, 5*(1), 62-64.

Welling, L. L. M., Jones, B. C., DeBruine, L. M., Conway, C. A., Law Smith, M. J., Little, A. C., et al. (2007). Raised salivary testosterone in women is associated with increased attraction to masculine faces. *Hormones and Behavior, 52*(2), 156-161.

Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science, 17*(9), 799-806.