ELSEVIER

Evolution and Human Behavior 28 (2007) 48–54

Evolution and Human Behavior

Color 3D bodies and judgements of human female attractiveness

Kathryn L. Smith, Piers L. Cornelissen, Martin J. Tovée*

Psychology Division, School of Biology and Psychology, Newcastle University, Newcastle Upon Tyne, UK

Initial receipt 25 January 2006; final revision received 19 May 2006

Abstract

For a set of color video clips that depict a 360° view of the bodies of 43 young Caucasian women who are within the normal range of percentage of body fat, we show that their attractiveness to both male and female observers depends strongly on their percentage of body fat and their level of skin tanning, but is not significantly related to their cardiovascular fitness (a key health measure). Although evolutionary psychology suggests that physical health should play a role in determining attractiveness judgements, it appears that cardiovascular fitness may be a weak cue, at least in bodies not undergoing cardiovascular exercise. Instead, it seems that more salient cues, such as body mass and skin tanning, are the primary determinants of attractiveness judgements.

Keywords: Female attractiveness; Female health; Cardiovascular fitness

1. Introduction

What drives attractiveness judgements? Some evolutionary psychologists have postulated that humans selected mates based on the display of certain physical cues that honestly signalled one mate to be more desirable (i.e., healthier and with a better reproductive potential) than another (Buss, 1987). Mate selection, in this way, would have enhanced their chances of successful reproduction, as mates with desirable cues are suggested to have been healthier and to have possessed greater reproductive potential. If attractiveness is an honest signal of a better-quality mate, it is, therefore, possible to suggest that an attractive body should also be a healthy body.

To explore this putative relationship in a contemporary population, we measured cardiovascular fitness (i.e., the capacity of the cardiovascular and respiratory systems to supply oxygen to working muscles during sustained physical activity). This is a key index of physical well-being and a strong predictor of longevity (Blair et al., 1996; Farrell, Braun, Cheung, Barlow, & Blair, 2002; Mora et al., 2003). For example, among women of reproductive age (15–44 years old) who are residing in England and Wales, cardiovascular diseases (such as heart disease, stroke, high blood pressure, etc.) are one of the highest causes of mortality (Barlow, Kohl, Gibbons, & Blair, 1995; Office of National Statistics, 2005), suggesting that cardiovascular fitness has a significant impact on health and mortality. Cardiovascular fitness has also been found to provide some protective benefits against cancer (e.g., Batty & Thune, 2000). Thus, although cardiovascular fitness only addresses a component of phenotypic quality, it is a component that should be related to judgements of attractiveness, if attractiveness is indeed related to actual health (Weeden & Sabini, 2005). To test this hypothesis, we took digital video clips of a set of volunteers' bodies of known cardiovascular fitness, which were then rated for attractiveness.

Previous studies have used 2D photographs or simple 3D wire-frame images of real women as stimuli (e.g., Fan, Liu, Wu, & Dai, 2004; Tovée, Hancock, Mahmoodi, Singleton, & Cornelissen, 2002). However, such images fail to capture the full range of potential cues that an observer in real life might use to judge attractiveness. Therefore, in the current study, we used color video clips of bodies that were twice rotated through 360° (see Fig. 1). Within the context of a laboratory setting, this provided

* Corresponding author. Psychology Department, The Henry Wellcome Building, Newcastle University, Framlington Place, NE2 4HH Newcastle Upon Tyne, UK. Tel.: + 44 0191 222 8616; fax: +44 0191 222 5622.

E-mail address: m.j.tovee@ncl.ac.uk (M.J. Tovée).

1090-5138/07/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.evolhumbehav.2006.05.007

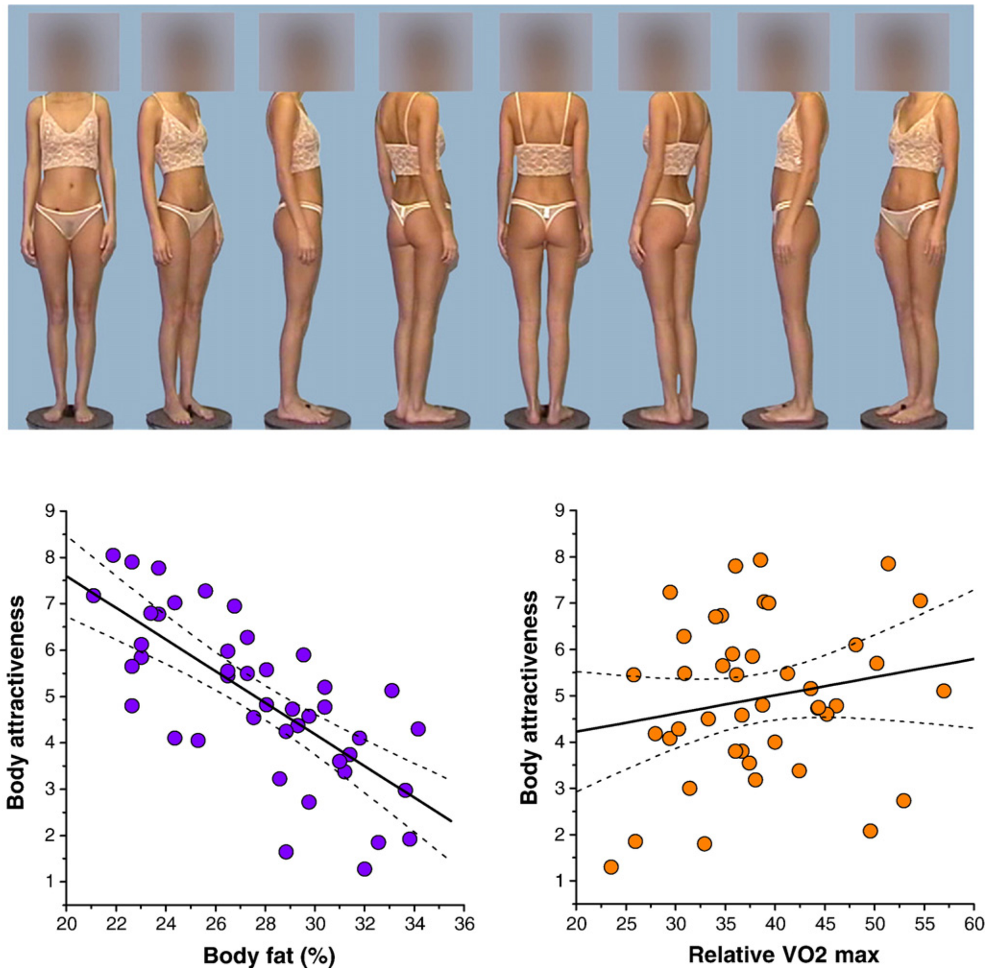

Fig. 1. Top: A subset of frames extracted from one video sequence, as the volunteer rotates through 360° counterclockwise. Lower left: A scatterplot showing the relationship between body attractiveness and the percentage of body fat. Lower right: A scatterplot showing the relationship between body attractiveness and r-predicted VO_2 max. In both plots, solid lines represent linear regressions between *x*- and *y*-variables; dotted lines represent their 95% confidence limits.

observers with many more cues than previously available with grey-level 2D images.

2. Methods

Forty-three Caucasian female volunteers, primarily university undergraduates, (age: mean=20.72 years; S.D.=2.19) consented to be filmed to provide stimuli for this study. Video clips were taken of our volunteers using a digital camcorder (with 625-horizontal-line resolution and an 800,000-pixel charge-coupled device chip) as they were twice rotated through 360° on a turntable (see Fig. 1). Clothing was standardized, with all women wearing an unsupportive flesh-colored vest and briefs. All volunteers adopted a standard pose, with feet slightly apart and arms on the sides. The volunteers' faces were then blurred in video sequences to ensure anonymity.

To measure their cardiovascular fitness, we used a standard 6-min submaximal cycle ergometry test from which a reliable maximal oxygen consumption score (predicted VO_2 max) can be estimated using the Astrand nomogram (Astrand & Rhyming, 1954). A VO_2 max test measures the maximum rate at which oxygen can be absorbed and utilized by an individual during exercise and can quantify an individual's aerobic work capacity. This test can be administered in either maximal or submaximal form. Maximal tests involve exercising individuals up to their maximal heart rate (HR) or until they reach a plateau of oxygen consumption; submaximal tests work up to only 80% of maximal HR. The former therefore provides an absolute measure of cardiovascular functioning, while the latter does not. However, using the Astrand nomogram and the norms from maximal testing, scores from submaximal tests can be used to reliably predict absolute cardiovascular functioning. A submaximal test was used here as it is less stressful on the individual and can be performed on anyone, regardless of fitness level.

Predicted VO_2 max scores can also be reported in different absolute or relative terms, with the latter taking into consideration the body weight of the individual, allowing for easier comparison across participants. Here we chose to report relative predicted VO_2 max (r-predicted

VO_2 max) scores to ensure that body weight did not influence test scores. For our set of volunteers, r-predicted VO_2 max ranged from 23.52 to 56.98 ml/min/kg (mean=38.43; S.D.=8.09). This is representative of the range of values found in the female population at large, and includes individuals from low to high cardiovascular fitness (e.g., Blair et al., 1989; DeVires & Housh, 1994). For American college students, the norm for women aged 20–29 years ranged from 28 ml/min/kg at low fitness to 49 ml/min/kg for high fitness; the 0 and 100 percentiles were 18 and 65 ml/min/kg, respectively (DeVires & Housh, 1994). We also calculated HR recovery time, as this has also been found to be independently predictive of mortality (Cole, Foody, Blackstone, & Lauer, 2000) and, therefore, may also provide a link between health and attractiveness. HR recovery time was defined as the change from peak HR during exercise to that measured after 2 min of recovery (Cole et al., 2000).

We measured volunteers' weight and height, and then calculated their body mass index (BMI; kg/m^2). The volunteers varied in BMI from 18.41 to 26.68 kg/m^2, with an average of 22.35 kg/m^2 (S.D.=2.30). Percentage of body fat was estimated from four skinfold thicknesses [biceps, triceps, subscapular (at a 45° angle), and suprailiac (at a 45° angle)] using the protocol and calculations of Durnin and Womersley (1974).

We chose to follow the Durnin and Womersley (1974) procedure, as it has been extensively replicated (see Eston, 2003). Although the measurement of body fat obtained from skinfold measurements is usually within 3–5% of that obtained with hydrostatic weighing (McArdle, Katch, & Katch, 2000), this does leave a small margin for error. A comparison of the error rate with either three or seven skinfolds has been assessed by a number of studies (reviewed in Heyward & Stolarczyk, 1996), and it was concluded that both methods are equally good in accurately predicting body fatness. For this reason, we chose to measure fewer skinfolds for practical benefits. It has been reported that combined skinfolds from the biceps, triceps, subscapular, and suprailiac muscles are negatively correlated with body density (measured using hydrostatic weighing) at $r=-.78$ for British women (Durnin & Rahaman, 1967). Body fat is inversely correlated with body density (fat has a lower density than muscle), so these results indicate a strong correlation between body fat measured by skinfolds and body fat measured by hydrostatic weighing.

Skinfold measurements were taken with Harpenden calipers (calibrated to a constant spring pressure of 10 g/mm^2). The women in the images had a percentage of body fat (%BF)=21.10–34.16 (average=27.70; S.D.=3.65). The relatively high correlation between BMI and the percentage of body fat ($r=.67$, $p<.001$) suggested that it would have been redundant to include both measures of body fat in our analysis. In previous studies (Tovée, Hancock, Mahmoodi, Singleton, & Cornelissen, 2002; Tovée, Maisey, Emery, & Cornelissen, 1999), BMI has been used in analyses, but the percentage of body fat is a more direct measure of body fat mass. We have therefore used the percentage of body fat, rather than BMI, in the subsequent analyses of our results.

We also obtained a number of biometric ratios from the women, including circumferential waist-to-hip ratio (WHR), circumferential waist-to-chest ratio (WCR), and torso-to-leg ratio (TLR) (calculated by torso length measured from the acromioclavicular joint to the perineum, and leg length as measured from the perineum to the ankle joint). The WHR range was 0.64–0.84, with an average of 0.74 (S.D.=0.04). The WCR range was 0.72–0.89, with an average of 0.80 (S.D.=0.04). The TLR range was 0.75–1.06, with an average of 0.88 (S.D.=0.06).

The volunteers in this image set were all Caucasian. However, even within this group, skin color varies due to factors such as the degree of sun tanning. As the video clips were in (24-bit) color, it was important to quantify this variation in skin tone, as it may have been a factor that influenced observers' judgements of attractiveness. Using Corel Photo-Paint 10 (Corel, USA), the mean values for the red (R), green (G), and blue (B) channels (range, 0–255) were calculated within a standard size patch of skin on each volunteer' back (50×60 pixels). The location of the patch was chosen as it was not affected by shadows, body curvature, clothing, or hair. The mean R, G, and B values of the skin patch for each volunteer's image were then factor-analyzed to compress them to a single value related to skin tone. The optimum factor analysis was without rotation; the first factor held a communality of 2.84 and explained 94.5% of the variance. The individual R, G, and B values correlated with the resultant factor at $r>-.95$ and $p<.001$. A positive number on this factor indicated darker skin, and a negative number indicated lighter skin. We called this factor the skin tone index. It is important to note that this factor is simply a covariate related to stimulus coloration and should not be treated as a direct measure of skin color.

The stimulus images were then presented to 20 male (age: average=21.45 years; S.D.=1.50) and 20 female (age: average=20.70 years; S.D.=1.34) Caucasian observers to be rated. For rating, the films were shown in a randomized order on a 17-in. LCD monitor (1280×1024-pixel resolution); each video clip measured 280×576 pixels and took approximately 15 s to rotate twice through 360°. Each observer watched the entire set of video clips to become aware of the range of shapes, sizes, and features of our volunteers. On second viewing, observers were asked to rate the images for attractiveness on a scale of 0–9 using a numeric keypad.

3. Results

3.1. Reliability

Winer's intraclass reliability for k means showed a high degree of interrater agreement for both male and female

Table 1
Pearson correlations between attractiveness ratings and all explanatory variables

	WHR	WCR	TLR	%BF	R-predicted VO_2 max	Skin tan	HR recovery time
Attractiveness	−.28	−.43**	−.17	−.73**	.19	.56**	−.01
WHR	–	.44**	.06	.31*	−.04	−.21	−.09
WCR		–	−.30	.52**	−.11	−.33*	−.10
TLR			–	.05	.12	.06	.05
%BF				–	−.19	−.36*	−.02
R-predicted VO_2 max					–	.22	.41*
Skin tan						–	.24

* $p<.05$.
** $p<.005$.

observer groups (0.96 for male and 0.97 for female observers; 0.99 when combined). We found a similar pattern of results for Cronbach's α: .98 for both male and female observers, and .99 when combined. Since the Spearman rank correlation between male and female raters was also high ($r=.95$, $p\leq.0001$), we combined results across male and female observer groups for further analysis.

3.2. Relationship between r-predicted VO_2 max, HR recovery time, and percentage of body fat

As one might expect, we found a significant positive correlation between r-predicted VO_2 max and HR recovery time ($r=.41$, $p<.01$). A significant negative relationship had been found previously between the percentage of body fat and cardiorespiratory fitness (CRF) in a large sample of White females ($n=259$) (Janssen et al., 2004). In this study, the women were allocated into two groups based on CRF, and the group with the lower CRF had a significantly higher percentage of body fat. A similar relationship was found among this study's sample of women, whereby the percentage of body fat and the r-predicted VO_2 max shared a moderate negative correlation with each other ($r=-.19$, $p>.1$). However, possibly due to the small participant sample in this present study, the relationship was not significant. This result was not due to a narrow range in either cardiovascular fitness or percentage of body fat in our group of volunteers. As described in Methods, there was a full range of cardiovascular fitness, and the range of percentage of body fat encompassed those slightly "below average" up to the "at-risk-for-obesity" category (using standards defined by Lohman, 1992). This weak relationship between the percentage of body fat and CRF may not be surprising, since it is possible to lead a sedentary lifestyle and have poor cardiovascular fitness yet keep body weight low through dieting.

3.3. Judgements of attractiveness

The pattern of relationships between variables is illustrated in the correlation matrix in Table 1 and in scatterplots in Fig. 1. We found no significant relationship between attractiveness and r-predicted VO_2 max or HR recovery time in this sample. Neither was a relationship found between attractiveness and WHR. However, we did find significant relationships between attractiveness and the percentage of body fat, WCR, and the skin tone index.

To further quantify these results, we ran a multiple regression analysis using PROC REG in SAS (SAS Institute, North Carolina, USA). All explanatory variables were centered by converting them to z-scores (Altman, 1991), and all were entered into the starting model. Next, we used PROC REG to estimate a best-fit model by simultaneously minimizing Mallow's Cp statistic, maximizing R^2, and permitting only explanatory variables that are significant at $p<.05$. According to these criteria, the best-fit model, using centered variables, was $y=0.35x_1-0.60x_2$ (where y=attractiveness, x_1=skin tone, and x_2=percentage of body fat). This simple model explained ~63% of the total variance in ratings; individually, the percentage of body fat could explain ~51% and skin tone could explain ~32%. Given that skin tone and the percentage of body fat were moderately correlated with each other (see Table 1), we also calculated the variance inflation factor for the explanatory variables in the best-fit model; neither exceeded 1.5, which was well within the acceptable limit of 10 (Birkes & Dodge, 1993).

4. Discussion

For a set of color video clips representing a 360° view of female bodies, we found that judgements of attractiveness were best explained by the percentage of body fat. This presentation format provides as many visual cues that an observer would see in the real world but in a controlled laboratory setting as possible. Only one previous study has used images that could be rotated through 360°. However, these images were shown as wire frames, rather than as 3D solid shapes with color, texture, surface curvature, shadowing, and shading (Fan et al., 2004). Our stimuli included these features and thus provided a more realistic representation of the human body. Consistent with our results, previous studies on attractiveness judgements that used relatively impoverished stimuli showed similar effects for BMI (which was closely correlated with the percentage of body fat in this sample) and a close volumetric equivalent, the volume height index (e.g., Fan et al., 2004; Tovée et al., 2002, 1999).

The present results also show a positive link between attractiveness and skin tone, with those possessing darker-toned skin receiving higher ratings of attractiveness. Although many cross-cultural studies have suggested that paler skin is generally regarded as more attractive by human populations (Aoki, 2002; Barber, 1995; Bond & Cash, 1992; Hill, 2002; Sahay & Piran, 1997), several studies have suggested that, for Caucasian faces in the Europe and America, moderate levels of tanning are regarded as more attractive than no tanning (e.g., Fink, Grammer, & Thornhill, 2001; Sahay & Piran, 1997). A paler coloring has been linked to both youth and fertility levels during the menstrual cycle (e.g., Frost, 1988; Van den Berghe & Frost, 1986); therefore, one might have argued that paler skin on the female body should be regarded as more attractive and healthy. However, skin color may be a culturally based status symbol indicating that someone with a tan can afford the free time necessary to acquire one, just as, in the 18th century, a paler complexion indicated higher status as it showed that an individual did not have to undertake manual outdoor labor (Etcoff, 1999). If this were the case, then one might expect preference for a darker skin tone to extend to all ethnic groups within a society. However, the preference for a more tanned appearance seems to be largely specific to Caucasians in Western cultures (Bond & Cash, 1992; Hill, 2002; Sahay & Piran, 1997).

In the current study, the correlation matrix showed that WCR was linked to attractiveness, although WHR was not. It has been suggested that women with a large bust relative to waist width (i.e., a low WCR) have higher levels of estrogen, which may in turn be positively associated with a higher probability of conception (Jasienka, Ziomkiewicz, Ellison, Lipson, & Thune, 2004). Thus, WCR might be used as a proxy for estrogen levels, and the prominent positioning of the breasts on the front of the body makes WCR a comparatively easy judgement to make and thus allows it to influence attractiveness ratings. A similar role has also been suggested for WHR (Singh, 1993). However, neither torso ratio survived the multiple regression analysis in this study, most likely because both WCR and WHR were correlated with the percentage of body fat in this sample. [Note: A correlation between the percentage of body fat and WHR is to be expected since large-scale epidemiological samples repeatedly demonstrate correlations between WHR and BMI (e.g., UK Department of Health, 2003).] This suggests that the major effects of the percentage of body fat on this sample of images may have been related more to dietary habits, and any additional influence of estrogen levels on body shape was either missed through sampling effects or simply outweighed by other factors.

Some previous studies using line-drawn figures have reported a strong relationship between WHR and attractiveness judgements (e.g., Henss, 1995; Singh, 1993). However, these line-drawn stimuli covaried BMI and WHR (i.e., as the WHR increases, so does the apparent BMI of the figures); thus, the change in attractiveness rating could be due to changing WHR, BMI, or both (Tovée & Cornelissen, 1999). Studies that have used digital photographs of real women have found a strong effect of changing BMI and a much weaker effect of WHR (e.g., Thornhill & Grammar, 1999; Tovée et al., 1998, 2002). Consistent with this result, other studies that have attempted to independently manipulate shape changes found that, although WHR can be significantly correlated with attractiveness judgements, it is a much weaker cue than BMI (Tovée, Benson, Emery, Mason, & Cohen-Tovée, 2003; Smith, Tovée, Hancock, Cox, & Cornelissen, 2006; Streeter & McBurney, 2003). These findings are consistent with our own results. However, Smith et al. have also shown that there is considerable variability in the shape of torsos (synthesized from four independent features defined by principal components analysis) that are treated as equally attractive, suggesting that observers may make tradeoffs among different biometric attributes when judging attractiveness (Smith, Tovée, Hancock, Bateson, Cox, & Cornelissen, 2006). Ultimately, to elucidate this complex problem, further research is required to identify exactly what image features drive perceptual judgements of attractiveness. But we also need to understand how genetic, environmental, and physiological factors interact to modify phenotypic appearance in the first place. Only then will it be possible to map the relationships between desired quality and biometric proxies such as the percentage of body fat, BMI, and WHR.

Our study shows a comparatively weak and nonsignificant correlation of .19 between attractiveness and VO_2 max, which is of a similar order of magnitude to the correlations found in previous studies that explored the relationship between physical health and female facial attractiveness (e.g., Hönekopp, Bartholomé, & Jansen, 2004; Shackelford & Larsen, 1999). Given the strength of this relationship, how many image clips would we need to reach statistical significance? Assuming an effect size of 0.19, reaching statistical significance ($\alpha = .05$) using a two-tailed test with an acceptable power of 0.80 would require a sample size of 214 image clips. This suggests that information from a very large number of bodies would have to be assimilated in our decision-making process before we could reliably judge the relationship between CRF and attractiveness. Even if we assume that we had obtained a statistically significant result, the modest relationship between attractiveness ratings and CRF means that attractiveness will be very insensitive to big changes in CRF status. For example, as we have stated in Methods, for American college students, the norm for women ranged from 28 ml/min/kg at low fitness to 49 ml/min/kg for high fitness (DeVires & Housh, 1994). At a 0.19 correlation with attractiveness, a simple regression shows that changing one's fitness from low to high changes one's attractiveness by only one point (from 4.5 to 5.5). By comparison, changing the percentage of body fat from just underweight to just overweight alters the attractiveness rating by 5 points (from 2.5 to 7.5). This fivefold difference in effect size suggests that cues to cardiovascular health are not very salient when the body is inactive, as in our video clips.

Since CRF is a good predictor of long-term health and longevity (e.g., Batty & Thune, 2000; Blair et al., 1996; Farrell et al., 2002; Mora et al., 2003), it is plausible that it should contribute to judgements of attractiveness. We suggest that cues to CRF may only become salient when someone is undertaking an activity that places cardiovascular strain upon them. For example, Brown et al. (2005) have shown that biological motion cues (in this case from dancing) can increase the saliency of cues relevant to sexual selection. Therefore, future experiments using video clips of women moving (e.g., dancing, running, or walking) may show much stronger correlations between attractiveness and cardiovascular health. Alternatively, cardiovascular health and body fat composition may have been closely linked in our ancestral population. One could determine the former by reference to the latter, which may be more perceptually salient. However, as current findings show, in our modern industrialized society, body fat and cardiovascular fitness can be decoupled, potentially making fitness judgements less straightforward and more inaccurate.

Acknowledgments

The authors would like to thank the Northumbria University Sports Science Department for the use of their facilities and equipment. We would also like to thank Melissa Bateson for her comments on the manuscript. This research was supported by an MRC studentship to K.L.S. and Wellcome Trust funding through the JIF initiative to P.L.C.

References

Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.

Aoki, K. (2002). *Sexual selection as a cause of human skin colour variation: Darwin's hypothesis revisited*. *Annals of Human Biology, 29*, 589–608.

Astrand, P. O., & Rhyming, I. (1954). A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. *Journal of Applied Physiology, 7*, 218.

Barber, N. (1995). The evolutionary psychology of physical attractiveness: Sexual selection and human morphology. *Ethology and Sociobiology, 16*, 395–424.

Barlow, C. E., Kohl, H. W. III, Gibbons, L. W., & Blair, S. N. (1995). Physical fitness, mortality and obesity. *International Journal of Obesity, 19* (S4):S41–S44.

Batty, D., & Thune, I. (2000). Does physical activity prevent cancer? *BMJ, 321*, 1424–1425.

Blair, S. N., Kampert, J. B., Kohl, III H. W., Barlow, C. E., Macera, C. A., Paffenbarger, R. S. Jr., & Gibbons, L. W. (1996). Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. *Journal of the American Medical Association, 276*, 205–210.

Blair, S. N., Kohl, H. W. III, Paffenbarger, R. S. Jr., Clark, D. G., Cooper, K. H., & Gibbons, L. W. (1989). Physical fitness and all-cause mortality. A prospective study of health men and women. *Journal of the American Medical Association, 262*, 2395–2401.

Birkes, D., & Dodge, Y. (1993). *Alternative methods of regression*. John Wiley & Sons.

Bond, S., & Cash, T. F. (1992). Black beauty—Skin color and body images among African-American college-women. *Journal of Applied Social Psychology, 22*, 874–888.

Brown, W. M., Cronk, L., Grochow, K., Jacobson, A., Liu, K., Popoviæ, Z., & Trivers, R. (2005). Dance reveals symmetry especially in young men. *Nature, 438*, 1148–1150.

Buss, D. M. Sex differences in human mate selection criteria: An evolutionary perspective. In C. Crawford, M. Smith, & D. Krebs, (Eds.), *Sociobiology and psychology: Ideas, issues and application* (pp. 335–351) Hillsdale, NJ: Erlbaum.

Cole, C. R., Foody, J. M., Blackstone, E. H., & Lauer, M. S. (2000). Heart rate recovery after submaximal exercise testing as a predictor of mortality in a cardiovascularly healthy cohort. *Annals of Internal Medicine, 132*, 552–555.

DeVires, H. A., & Housh, T. J. (1994). *Physiology of exercise for physical education, athletics and exercise science*. (5th ed.). Madison, WI: Brown and Benchmark Publishers.

Durnin, J. V.G. A., & Rahaman, M. M. (1967). The assessment of the amount of fat in the human body from measurements of skinfold thicknesses. *British Journal of Nutrition, 21*, 681–689.

Durnin, J. V.G. A., & Womersley, J. (1974). Body fat assessed from total body density and its estimation from skinfold thickness: Measurements of 481 men and women aged from 16 to 72 years. *British Journal of Nutrition, 32*, 77–97.

Eston, R. (2003). Prediction of body fat from skinfolds: The importance of including sites from the lower limb. *Journal of Sports Sciences, 21*, 369–370.

Etcoff, N. (1999). *Survival of the prettiest*. New York: Doubleday.

Fan, J., Liu, F., Wu, J., & Dai, W. (2004). Visual perception of female physical attractiveness. *Proceedings of the Royal Society of London Series B, 271*, 347–352.

Farrell, S., Braun, L., Cheung, Y., Barlow, C., & Blair, S. (2002). The relation of body mass index, cardiorespiratory fitness, and all-cause mortality in women. *Obesity Research, 10*, 417–423.

Fink, B., Grammer, K., & Thornhill, R. (2001). Human (*Homo sapiens*) facial attractiveness in relation to skin texture and colour. *Journal of Comparative Psychology, 115*, 92–99.

Frost, P. (1988). Human skin color: A possible relationship between its sexual dimorphism and social perception. *Perspectives in Biology and Medicine, 32*, 38–58.

Henss, R. (1995). Waist-to-hip ratio and attractiveness. Replication and extension. *Personality and Individual Differences, 19*, 479–488.

Heyward, V. H., & Stolarczyk, L. M. (1996). *Applied body composition assessment*. Champaign, IL: Human Kinetics.

Hill, M. E. (2002). Skin color and the perception of attractiveness among African Americans: Does gender make a difference? *Social Psychology Quarterly, 65*, 77–91.

Hönekopp, J., Bartholomé, T., & Jansen, G. (2004). Facial attractiveness, symmetry, and physical fitness in young women. *Human Nature, 15*, 147–167.

Janssen, I., Katzmarzyk, P. T., Ross, R., Leon, A. S., Skinner, J. S., Rao, D. C., Wilmore, J. H., Rankinen, T., & Bouchard, C. (2004). Fitness alters the associations of BMI and waist circumference with total and abdominal fat. *Obesity Research, 12*, 525–537.

Jasienka, G., Ziomkiewicz, A., Ellison, P. T., Lipson, S., & Thune, I. (2004). Large breasts and narrow waist indicate high reproductive potential in women. *Proceedings of the Royal Society of London Series B, 271*, 1213–1217.

Lohman, T. G. (1992). *Advances in Body Composition Assessment*. Champaign, IL: Human Kinetics Publishers.

McArdle, W. D., Katch, F. L., & Katah, V. L. (2000). *Essentials of Exercise Physiology*. (2nd ed.) Philadelphia, PA: Lippincott, Williams & Wilkins.

Mora, S., Redberg, R. F., Cui, Y., Whiteman, M. K., Flaws, J. A., Sharrett, A. R., & Blumenthal, R. S. (2003). Ability of exercise testing to predict cardiovascular and all-cause death in asymptomatic women. *Journal of the American Medical Association, 290*, 1600–1607.

Office of National Statistics. (2005). Death registrations in England and Wales: Causes. *Health Statistics Quarterly, 26*, 62–69 (Table 2).

Sahay, S., & Piran, N. (1997). Skin-color preferences and body satisfaction among South Asian Canadian and European–Canadian female university students. *Journal of Social Psychology, 137*, 161–171.

Shackelford, T. K., & Larsen, R. J. (1999). Facial attractiveness and physical health. *Evolution and Human Behavior, 20*, 71–76.

Singh, D. (1993). Adaptive significance of female physical attractiveness: Role of waist-to-hip ratio. *Journal of Personality and Social Psychology, 65*, 293–307.

Smith, K. L., Tovée, M. J., Hancock, P. J. B, Bateson, M., Cox, M. A. A., & Cornelissen, P. L. (2006). An analysis of body shape preference based on image statistics. *Visual Cognition,* (in press).

Streeter, S. A., & McBurney, D. (2003). Waist–hip ratio and attractiveness: New evidence and a critique for a 'critical test'. *Evolution and Human Behavior, 24*, 88–98.

Thornhill, R., & Grammar, K. (1999). The body and face of woman: One ornament that signals quality? *Evolution and Human Behavior, 20*, 105–120.

Tovée, M. J., Benson, P. J., Emery, J. L., Mason, S. M., & Cohen-Tovée, E. M. (2003). Measurement of body size and shape perception in eating disordered and control observers using body-shape software. *British Journal of Psychology, 94*, 501–516.

Tovée, M. J., Hancock, P. J. B., Mahmoodi, S., Singleton, B. R. R., & Cornelissen, P. L. (2002). Visual cues to female attractiveness: Waveform analysis of body shape. *Proceedings of the Royal Society of London. Series B, 269*, 2205–2213.

Tovée, M. J., & Cornelissen, P. L. (1999). The mystery of female beauty. *Nature, 399*, 215–216.

Tovée, M. J., Maisey, D. S., Emery, J. L., & Cornelissen, P. L. (1999). Visual cues to female physical attractiveness. *Proceedings of the Royal Society of London Series B, 266*, 2111–2118.

Tovée, M. J., Reinhardt, S., Emery, J. L., & Cornelissen, P. L. (1998). Optimal BMI and maximum sexual attractiveness. *The Lancet, 352*, 548.

UK Department of Health. (2003). The Health Survey for England—2003. Available: http://www.dh.gov.uk/PublicationsAndStatistics/PublishedSurvey/HealthSurveyForEngland/HealthSurveyResults/fs/en.

Van den Berghe, P. L., & Frost, P. (1986). Skin color preference, sexual dimorphism and sexual selection. A case of gene-culture co-evolution? *Ethnic and Racial Studies, 9*, 87–118.

Weeden, J., & Sabini, J. (2005). Physical attractiveness and health in Western societies: A review. *Psychological Bulletin, 131*, 635–653.