Psychology, Evolution & Gender 3.1 April 2001 pp. 27–46

Predicting female physical attractiveness

Waist-to-hip ratio versus thinness

Rebecca M. Puhl and Fred J. Boland
Queen's University, Kingston, Ontario

Abstract

Introductory psychology students (120 females and 120 males) rated attractiveness and fecundity of one of six computer-altered female figures representing three body-weight categories (underweight, normal weight and overweight) and two levels of waist-to-hip ratio (WHR), one in the ideal range (0.72) and one in the non-ideal range (0.86). Both females and males judged underweight figures to be more attractive than normal or overweight figures, regardless of WHR. The female figure with the high WHR (0.86) was judged to be more attractive than the figure with the low WHR (0.72) across all body-weight conditions. Analyses of fecundity ratings revealed an interaction between weight and WHR such that the models did not differ in the normal weight category, but did differ in the underweight (model with WHR of 0.72 was less fecund) and overweight (model with WHR of 0.86 was more fecund) categories. These findings lend stronger support to sociocultural rather than evolutionary hypotheses.

Keywords: gender, attractiveness, weight, waist-to-hip ratio (WHR)

Ideals of female physical attractiveness have clearly varied across time and cultures (Seid 1994; Wilfley and Rodin 1995). In recent years, in Western cultures, thinness has been stressed in various media as the ideal for women (Wiseman, Gray, Mosimann and Ahrens 1992). Consistent with this sociocultural pressure, many studies show that for females, being thin is

Psychology, Evolution & Gender
ISSN 1461-6661 print/ISSN 1470-1073 online
http://www.tandf.co.uk/journals
DOI: 10.1080/1461666011004957 3

perceived as more attractive than being average weight, plump, overweight, or obese (Brownell 1991; Fallon and Rozin 1985; Franzoi and Herzog 1987; Rozin and Fallon 1988; Silverstein, Peterson and Perdue 1986; Spillman and Everington 1989; Stice, Schupak-Neuberg, Shaw and Stein 1994). From this perspective, thinness should be a primary determinant of what males in Western cultures find attractive in females. Consider, for example, that centrefolds in *Playboy* are, on average, about 15 per cent below expected weight for height and certainly not representative of average females in Western culture (Wiseman et al. 1992). That this is a phenomenon largely restricted to Western industrialized cultures is reflected in the finding that dieting and actual eating disorders increase in females moving from non-Western to Western cultures (DiNicola 1990).

Singh (1993a) proposes a different view of what males across all cultures are predominantly attracted to in the female form. Singh presents an evolutionary perspective which proposes that the gynoid (lower body) fat distribution in females as reflected by the WHR is a critical and universal cue in signalling reproductive value to males (Singh 1993a). The WHR refers to the ratio of the circumference of waist-to-hips which is an indication of the anatomical distribution of body fat. According to Singh (1993b), a typical and healthy range for females is from 0.67 to 0.80. This reasoning stems from evidence indicating links between the WHR and mechanisms regulating female health and reproductive status (such as puberty, ovarian disease, menstrual irregularity, diabetes and mortality), which Singh suggests are critical factors of attractiveness (Singh 1993a). A WHR in the ideal range should thus be conceptualized as being most attractive to males as it signals high reproductive capability and accurately signals health in terms of absence of disease (Singh 1995a).

Singh proposes that males have evolved mechanisms to detect and rely on the WHR of females to evaluate their health and reproductive capacity. As he explains, 'WHR acts as a wide first-pass filter, which would automatically exclude women who are unhealthy or who have low reproductive capacity. A man may not even be aware of this initial selectivity' (1993a: 304). Only after this first 'culturally invariant' filter is passed do other features such as face, skin or weight, which may vary between cultures, become utilized in the final mate selection. In response to social influences of thinness such as the media and diet industry, Singh argues that studies, such as those cited above to support a sociocultural position, have focused only on body weight, ignoring body shape. Whereas the sociocultural position sees the widespread dieting practised by females in Western culture as an effort to approach the thin ideal, Singh (1994a) views it as an effort to achieve a more desirable gynoid body shape.

In several studies, Singh (1993a, 1993b, 1994a, 1994b, 1995a; see also Singh and Young 1995) consistently found that participant judgements of female attractiveness were associated with the WHR, where female stimuli with a low WHR (e.g. 0.7) were judged as most attractive in comparison to female figures with a high WHR (e.g. 0.85). In addition, Singh's results indicated that female attractiveness was judged most favourably by his participants for women of a normal body weight (with a WHR in the ideal range) rather than underweight or overweight females (1993a; 1993b; 1994b). These findings have also been demonstrated cross-culturally with Indonesian and African-American populations (Singh and Luis 1994).

While Singh presents a valid argument and gives important reasons to acknowledge the role of the WHR in judging female attractiveness, there are equally valid reasons to believe that this issue is not resolved. Methodologically, Singh's studies employed a standard within-subjects design where a number of line-drawn female figures (varying in perceived weight and size of WHR) are randomized, assigned a letter for identification purposes, and printed on 8 × 11-inch paper so that all female figures could be simultaneously examined by participants (Singh 1993a). These figures were then ranked from least to most attractive (e.g. from 1 to 12 for twelve figures). Several problems exist within this framework. First, the use of a within-subjects design where participants can freely compare figures is open to more potential bias relative to between-subjects designs since it is easier to guess the experimental manipulation and respond to the perception of what the experimenter wants (Orne 1962; Rosenthal and Rosnow 1975). A between-subjects design in which each subject remains naive and evaluates only one figure on a scale of attractiveness helps to avoid this potential bias. It could also be argued that the non-parametric technique of rank ordering a number of female figures is not as statistically powerful as parametric data derived from a scale-rating system.

Second, Henss (1995) notes that Singh's composite ranking of attractiveness was derived unconventionally by subtracting the percentage of the participants who assigned the highest ranking of attractiveness (e.g. a ranking of 12 out of twelve figures) from the percentage who assigned the lowest ranking (e.g. a ranking of 1 out of twelve figures). This statistic only focuses on the extreme ranking categories which is insufficient, and may have exaggerated the effects of Singh's WHR manipulations (Henss 1995). In Henss's (1995) own study where a weighted composite of a number of scales were used to measure attractiveness, results were opposite to those reported by Singh. That is, underweight figures were consistently judged as most attractive, followed by normal weight and overweight figures respectively (Henss 1995).

A third methodological limitation lies within Singh's use of line drawings of female figures. As Singh (1995b) self-critically indicates, the absence of detail and realism of a drawn picture can lead to artificial judgements, and more lifelike photographs may influence judgements differently to line drawings. Although Singh (1993b) once attempted to use photographs instead of his widely used drawings, his photographic images (where the WHR was slightly modified in each image) depicted only the torso and lower body of the female figure. Furthermore, the model used in his photographs was wearing an unusual outfit comprised of a black leather bikini, a silver chain belt, black wrist bands and boots – all of which may have led participants to infer certain characteristics about the figure or which may have conveyed a certain message to participants which could also affect their judgements. The findings of this particular study indicated that the normal weight figure with the low WHR was judged to be most attractive. However, it seems premature to accept these conclusions given the confounding variables associated with the figure's appearance.

Empirically, there are also reasons to question Singh's findings. First, one of Singh's own studies (1995a) comparing the WHR of 'slender' verses 'heavy' females indicated that a low WHR increased attractiveness judgements only of slender females; not heavy figures. That is, heavy female figures with a low WHR were found to be less attractive than slender figures with a higher WHR. According to Singh's theory, low WHR of a heavy female is expected to be more attractive than a slender female with a high WHR, and yet this finding led Singh to conclude that obesity had a stronger influence on judgements of attractiveness than did WHR.

A related contradiction worth mentioning pertains to the numerical range of WHR preferences of Singh's participants. Singh maintains that the WHR range of 0.67 to 0.80 represents a typical and healthy female and also reflects optimal reproductive capacity (Singh 1993b). However, in Singh's 1993 study, he found that a very low and unhealthy WHR of 0.6 was judged as being most attractive and healthy in comparison to a WHR of 0.7, 0.8 or 0.9 (Singh 1993b). Singh did not comment that this ranking was below a healthy and normal range for females and that it is outside the WHR range that he has defined for females. The remainder of Singh's studies manipulated WHR ranging from 0.7 to 1.0, which indicated a preference of WHR in the consistent order from most to least attractive of 0.7, 0.8, 0.9 and 1.0. However, Henss's (1995) replication of Singh's research resulted in a WHR of 0.8 being rated as most attractive, followed by 0.7, 0.9 and 1.0; a different pattern from Singh's findings. A second study similarly reported that mean attractiveness ratings dependent on WHR alone resulted in a WHR of 0.8 being judged most favourably (Furnham, Tan and McManus 1997).

Finally, a recent study tested the WHR hypothesis of physical attractiveness and found little evidence of an effect of the WHR (Tassinary and Hansen 1998). The stimulus set used in their study consisted of line drawings of female figures in which the weight, hip size and waist size of drawings were varied independently, and in which participants ranked in terms of physical attractiveness and fecundity. Findings indicated that weight and hip size were salient independent predictors of attractiveness and fecundity regardless of the WHR, and that the WHR had little predictive value for judgements. In addition, it was reported that attractiveness and fecundity could be either unrelated or related, positively or negatively, to the WHR depending on weight, hip size and waist size (Tassinary and Hansen 1998). Thus, although these authors employed a within-subjects design with line-drawn figures similar to those used in previous studies, their findings dis-confirm the evolutionary argument that the WHR is the most important determinant of female physical attractiveness. These results, coupled with contradictions from previous studies, question whether WHR is really the best predictor of female physical attractiveness as Singh proposes.

The purpose of the present study is to expand upon this research and to explore whether there is an influence of thinness on preferences of female physical attractiveness independent of the WHR. This can be accomplished by holding the WHR of female stimuli constant and manipulating body weight and breadth only. Taking into consideration the problems in previous research that have been mentioned, the present study makes several improvements. First, computer-generated photographic images of females (portrayed from head to toe) were used in replacement of the previously and widely used line drawings. This type of stimuli has not been previously used, but it clearly strengthens this research as participants can be presented with a real-life, full-bodied image upon which to make their judgements. Second, in order to avoid bias of attractiveness judgements that can occur when comparing a variety of stimulus figures, the present study employed a between-subjects design where each participant was shown only one female figure. In addition, factors associated with attractiveness, such as clothing, were held constant for each figure viewed. Third, a general Likert-type rating scale of attractiveness and fecundity (similar to that used by Henss (1995)) was used for participant responses which allows for more powerful data than Singh's non-parametric ranking of stimulus figures. Finally, both male and female participants were included in the present study to help clarify any gender differences or similarities that exist in perceptions of female physical attractiveness. It is hoped that, by including these components, a more confident conclusion can be drawn in determining the influence of thinness versus the WHR in predicting female attractiveness.

Three hypotheses concerning determinants of female physical attractiveness will be evaluated in this investigation. The first hypothesis follows from Singh's evolutionary theory which predicts that female physical attractiveness is determined by the WHR. According to Singh, ratings of attractiveness should be highest among females who have a WHR in the ideal range. By contrast, the second hypothesis reflects sociocultural theory suggesting that perceived body weight, particularly thinness, will account for more variance in judgements of attractiveness than the WHR. Finally, a third possible outcome is that an interaction will occur between both sociocultural and evolutionary determinants of physical attractiveness.

Methodology

Participants

The total sample consisted of 129 female and 123 male introductory psychology students, between the ages of 17 and 23 years. Each person received a 1 per cent mark credit for participating. This research received ethical approval from Queen's University.

Materials

Stimuli of female figures were derived from photographic images of actual female models. To obtain these stimuli figures, the experimenter took waist, hip, height and weight measurements of females who volunteered to pose as models for the study, and then calculated the waist-to-hip-ratio (WHR) and body mass index (BMI) for each female. Two females were then selected to model for the study on the basis of their BMIs falling in the normal or healthy range for females 20 to 25 years of age, and on the basis of their WHR measurements most accurately representing one of two different WHR sizes; one that was clearly in the ideal range according to Singh and past experimental results (Model A; BMI = 22.6; WHR = 0.72), and one that fell outside and above this range (Model B; BMI = 21.31; WHR = 0.86). The ages of Models A and B were 21.92 years and 22.41 years (M = 22.17 years), and their heights were 66 inches and 68 inches (M = 67 inches), respectively. Although these heights differed by two inches, this difference was reduced by manipulating the background of the photographs so that the models were perceived as approximately the same height. Also, Singh and other past investigators have not commented on height being a factor to consider as the WHR is the main variable being examined. All faces of models in the photographs were distorted to prevent confounding facial characteristics and to keep the identity of the female models anonymous.

Each female model was photographed against a background of chromatic blue material using a Canon Optura Digital Video Camera. All models wore the same black bathing suit for the photographs to make their waist and hips clearly visible. No other clothes or accessories were worn. Photographic images were then digitized by a plug-in developed for use with Adobe Photo Shop (Duchane 1999). The computer-generated manipulations involved changing the body weight of each model while maintaining her WHR constant. Two distinct manipulations were performed for each image where the widths of the female figures were reduced by 20 per cent and increased by 20 per cent to correspond to increases and decreases in weight. Experience suggests that this may be equivalent to a 20 per cent increase or decrease in weight. These percentages were selected using the classification of healthy weights for females aged 19 to 24 years according to the BMI, which requires females to have a BMI between 19 and 24 (Canada Expert Group on Weight Standards 1988). A 20 per cent decrease in breadth placed the BMI of both models below the normal and healthy weight requirements for their ages resulting in an underweight classification, whereas a 20 per cent increase placed the BMI of both models above the normal range, resulting in an overweight status with increased health risks. Therefore, each of the two models was depicted in three different photographic images as being underweight, normal weight and overweight, but still having the same WHR in all three manipulated images (see Figure 1).

Each image was printed in colour using a laser printer, and attached to separate sheets of 8½ × 11-inch paper so that a clear, full-bodied figure could be accurately viewed by participants. Photographs were kept approximately the same size as those drawings used in several of Singh's studies (Singh and Young 1995). Below the picture, two different seven-point scales were provided for participants to make ratings. The first stated 'Please make a mark on the scale below to indicate your judgement of the model's overall physical attractiveness', where 1 represented very unattractive and 7 represented very attractive. To be consistent with past research on the WHR, a second scale was included regarding the fecundity of the female. The instructions here stated, 'Now, using the same procedure, please judge the "fecundity" of the model in the photograph. That is, in your opinion, what is her capability of bearing children?', where 1 represented a very weak ability of bearing children and 7 represented a very strong ability to bear children.

In addition to attractiveness and fecundity ratings of the stimulus figure, several other relevant areas were assessed in the questionnaire package. For exploratory purposes, participants were asked to judge the model to be ideal, average or non-ideal on a seven-point scale. Specifically, these questions assessed the participants' judgements of the target model's overall body weight, overall body size, overall body shape, as well as the shape and size of the model's hips, waist, thighs, shoulders, stomach and breasts.

Following these ratings, demographic information was gathered including the participants' height, weight, sex and age. This allowed for any gender differences in attractiveness ratings to be examined, and allowed the BMI for each participant to be calculated to indicate those participants who

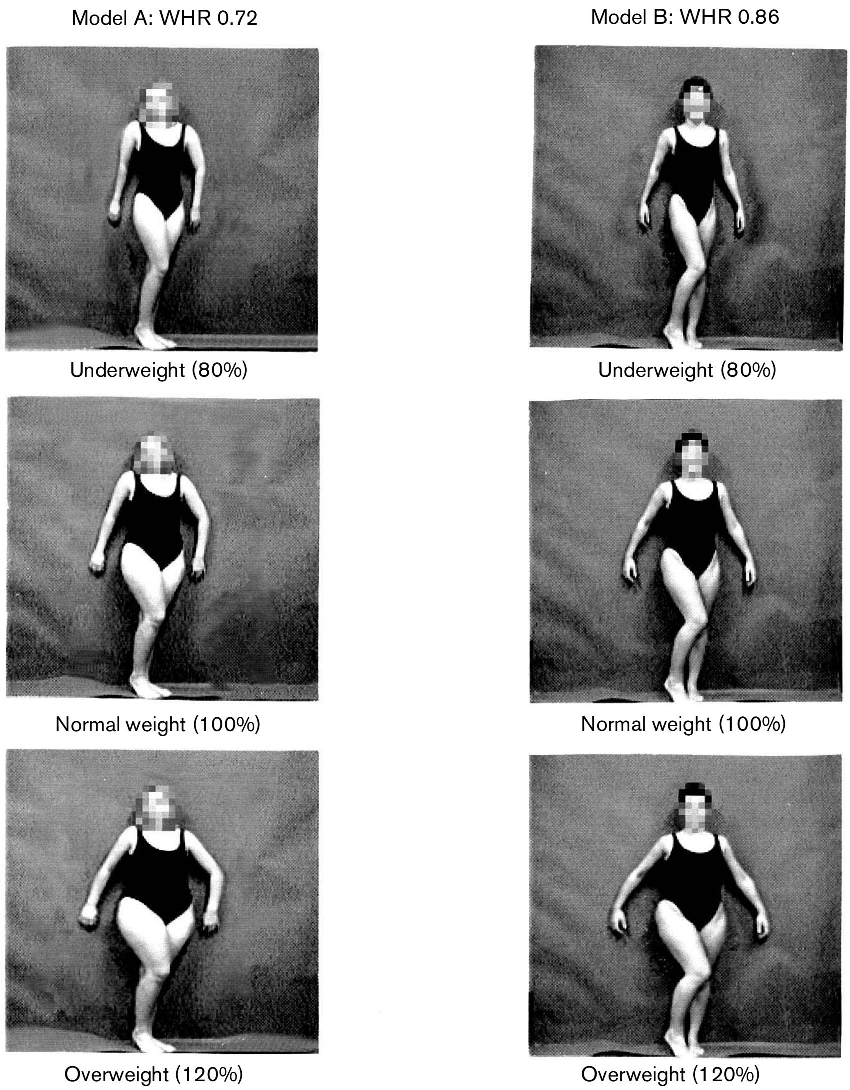

Figure 1 Computer-generated images of Model A (WHR = 0.72) and Model B (WHR = 0.86) depicted in underweight, normal weight and overweight conditions

were underweight, normal weight or overweight, and to determine whether participants' own BMIs were related to their attractiveness ratings of the stimulus figure. Second, two questions on dieting were asked to determine whether there was an association between dieting and physical attractiveness ratings. These dichotomous questions were (1) 'Are you currently dieting?' and (2) 'If you are not currently dieting, have you ever dieted in the past?' Finally, participants were asked to indicate how they feel about their own weight. Specifically, these questions were (1) 'In your opinion, are you presently (a) very underweight, (b) underweight, (c) average weight, (d) overweight or (e) very overweight?' and (2) 'Currently, how satisfied are you with your body?' This question was rated on a seven-point scale where 1 represented 'not at all satisfied' and 7 represented 'completely satisfied'. These questions were included to help determine if personal perceptions regarding one's own weight were related to perceptions of physical attractiveness ratings.

Procedure

Participants completed the questionnaires in regular classrooms. In total there were six different questionnaire packages, each containing one version of the photographic images of the two models. That is, either Model A (WHR = 0.72) or B (WHR = 0.86) was presented in either the underweight, average weight or overweight condition. These six different questionnaire packages were labelled A–, A, A+, and B–, B, B+. However, all packages still contained the identical questions and rating scales for each participant; only the stimulus picture varied. Of the original 253 questionnaires distributed, thirteen surveys were discarded due to incomplete or missing data. The final sample of 240 students allowed each of the six different stimulus figures to be evaluated by 40 students each (20 females and 20 males).

Prior to distribution of questionnaires, the order of questionnaires was randomized. Once presented with the questionnaire, all participants were asked to read and sign the consent form before responding to the survey. Other than the information provided on the consent form, which indicated to participants that they would be asked to complete a survey investigating female physical attractiveness, participants were not provided with any further details regarding the study. After the surveys had been successfully completed, the experimenter collected the questionnaires, and students were thanked for their participation and debriefed.

Results

Sample characteristics

The average height and weight among female participants was 1.66 metres (SD = 0.07) and 60.85 kilograms (SD = 10.48) respectively. BMI calculations indicated a mean of 21.94 (SD = 3.11) ranging from 16.70 to 37.92 among females. For males, the average height and weight was 1.79 metres (SD = 0.06) and 76.29 kilograms (SD = 12.37) respectively. The mean BMI for males was 23.77 (SD = 3.21) ranging from 17.17 to 34.50. The mean age of participants was 19.67 years (SD = 0.81) with a range of 17.71 to 23.67 years. Among females (n = 120) the mean age was

19.44 years (SD = 0.61), and the mean age of males (n = 120) was 19.90 years (SD = 0.61). There were no significant differences in ages of females and males.

Evaluation of main hypotheses

Overall physical attractiveness ratings were analysed using a 2 (gender) by 2 (WHR) by 3 (weight) ANOVA. Weight of the model (F (2, 238) = 165.79, $p < 0.001$), WHR (F (1, 239) = 40.56, $p < 0.001$), and gender of participants (F (1, 239) = 22.37, $p < 0.001$) emerged as main effects, with no interaction evident (see Figure 2). Comparisons of the attractiveness ratings given in the underweight, normal weight and overweight conditions revealed that the underweight figure was viewed as significantly more attractive. As Figure 2 indicates, participants rated Model B (WHR = 0.86) to be more physically attractive than Model A (WHR = 0.72) across underweight, normal weight and overweight conditions. In addition, attractiveness ratings were much more favourable for both models as their weights decreased from overweight to underweight. Thus, students clearly differentiated among the target weight manipulations of the models, and they made distinct attractiveness judgements in these various weight conditions.

Fecundity ratings of the models were also analysed using a 2 (gender) by 2 (WHR) by 3 (weight) ANOVA. A significant main effect emerged for weight on participants' ratings of fecundity (F(2, 238) = 4.23, p = 0.02). Although ratings of fecundity for Model A (M = 5.45, SD = 0.11) were slightly higher than for Model B (M = 5.30, SD = 0.11), no significant differences in WHR emerged. A significant interaction was revealed between weight and WHR conditions of the models (F(2, 238) = 8.6, $p < 0.001$). As shown in Figure 3, Model B (WHR = 0.86) was rated less fecund than Model A (WHR = 0.72) at the underweight level, and more fecund than Model A at the overweight level. However, in the normal weight conditions there were no significant differences in fecundity ratings between the two WHRs of Model A and Model B. A post-hoc analysis using the Student–Neuman–Keuls test ($p < 0.05$) confirmed that fecundity ratings in the normal weight condition significantly differed from ratings in the underweight and overweight conditions of the models. Thus, normal weight conditions of both models received the highest fecundity ratings.

Secondary analyses

Participants were asked to rate the overall body weight of the model they viewed in order to determine whether they perceived the different

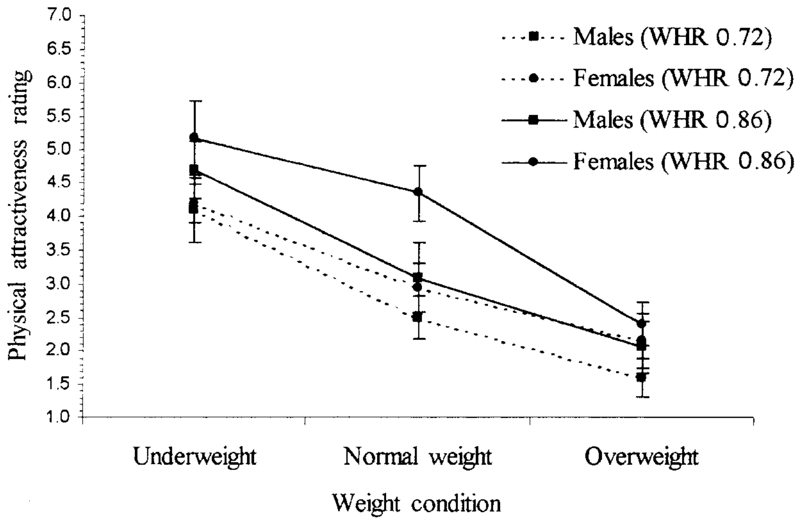

Figure 2 Mean physical attractiveness ratings across weight conditions. Vertical lines depict standard errors of means

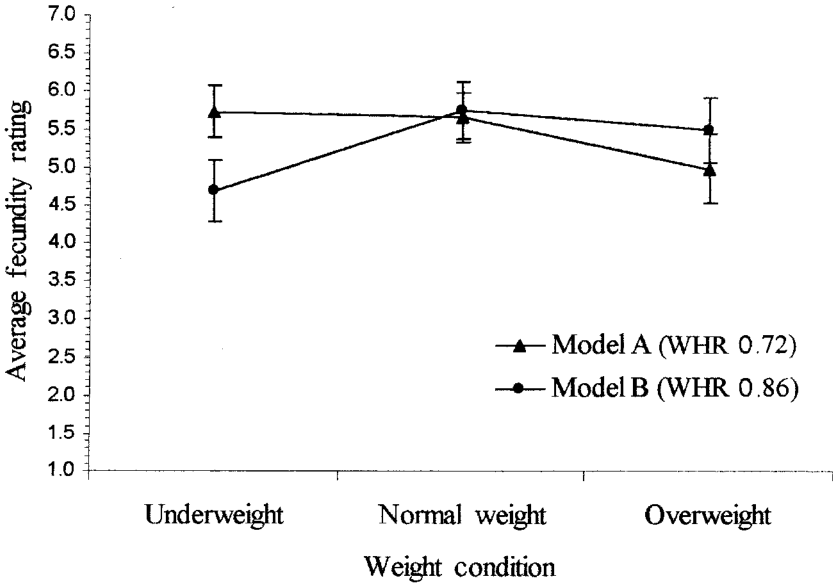

Figure 3 Mean fecundity ratings across weight conditions. Vertical lines depict standard errors of means

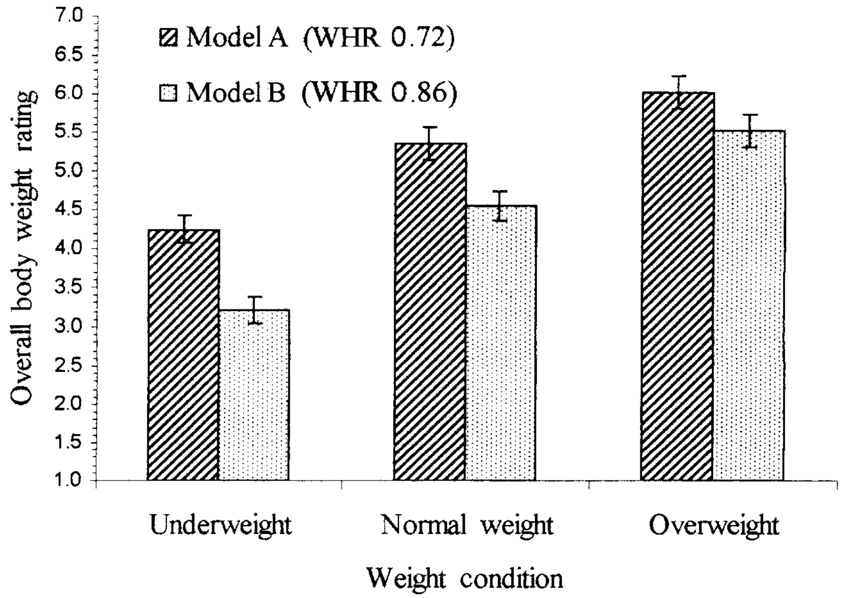

Figure 4 Mean ratings (collapsed across gender) of overall body weight across weight conditions, where 1 represents very underweight and 7 represents very overweight. Vertical lines depict standard errors of means

weight conditions of the models appropriately (see Figure 4). To reduce confusion in this analysis, ratings of overall body weight were reverse scored so that 1 represented very overweight and 7 represented very underweight. A 2 (gender) by 2 (WHR) by 3 (weight) ANOVA indicated that participants clearly judged the three weight conditions of each model to be significantly different from each other. A significant main effect for weight occurred ($F(2, 238) = 228.76$, $p < 0.001$), and a weaker main effect for WHR was reported, indicating that participants distinguished the two different WHRs of the models from each other ($F(1, 239) = 98.12$, $p < 0.001$).

The analysis of variance also revealed a small but significant interaction between weight and WHR ($F(2, 238) = 3.77$, $p < 0.001$). This interaction is barely perceptible in Figure 5, which shows that the discrepancy in ratings of overall body weight between Model A and Model B was largest in the underweight condition, but gradually declined in the normal weight and overweight conditions, where it was the smallest. This finding suggests that the effect of weight on ratings was so strong in the overweight condition that the effect of WHR was almost eliminated.

Participants additionally rated how ideal they judged the model's overall body size to be on a seven-point scale where 1 represented non-ideal and 7 represented ideal. Again, a 2 (gender) by 2 (WHR) by 3 (weight) ANOVA revealed a main effect of weight on ratings of the models' overall body sizes

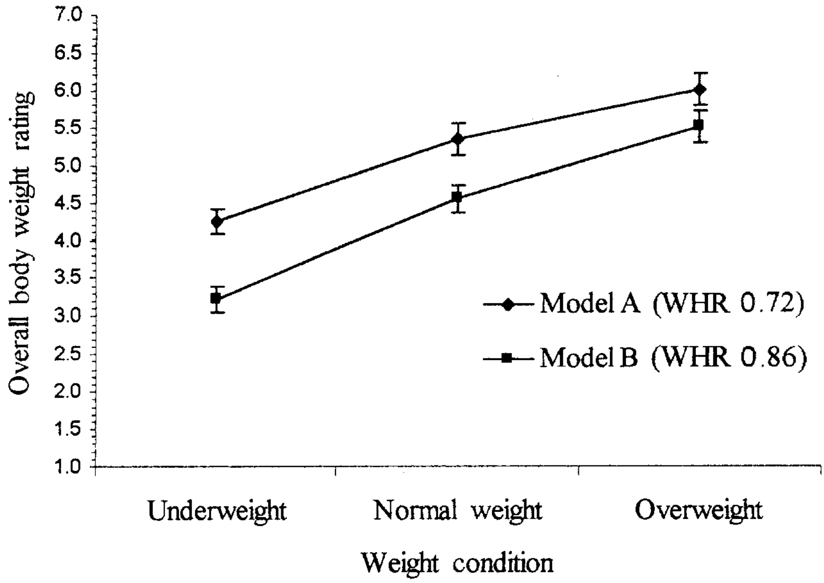

Figure 5 Interaction between weight condition and WHR for mean ratings of overall body weight, where 1 represents very underweight and 7 represents very overweight. Vertical lines depict standard errors of means

(F(2, 238) = 97.13, $p < 0.001$), as well as a main effect of WHR (F(1, 239) = 23.41, $p < 0.001$). This analysis indicated that participants' judgements of the overall body size of the model increased in favourability as the models' body weights decreased from overweight to normal weight to underweight conditions. Students also judged Model B (WHR = 0.86) to have a more ideal body size than Model A (WHR = 0.72).

Participants also rated how ideal they judged the model's overall body shape to be. Analysis of variance indicated a main effect of weight on ratings of overall body shape (F (2, 238) = 80.49, $p < 0.001$), as well as a main effect of WHR (F(1, 239) = 46.57, $p < 0.001$). Again, these results show that students' judgements of the overall body shape of the model they viewed were more favourable as the model's body weight decreased from overweight to normal weight to underweight, and participants perceived Model B to have a more ideal body shape than Model A.

Although not presented here, separate 2 (gender) by 2 (WHR) by 3 (weight) ANOVAs were also computed on participant perceptions of the perceived size and shape of the models' hips, waist, thighs, shoulders, stomach and breasts to be. These analyses of size and shape ratings of specific body parts of the models strongly corresponded to ratings of overall body size and shape, and all analyses reflected the general pattern seen in the more global ratings discussed above. Reliability analyses were also carried out on each set of size and shape ratings. An alpha reliability of 0.82 was obtained for the ratings of size of body parts, and an alpha reliability of 0.88 was reported for the shape ratings. When size and shape items were combined, the amalgamated score yielded an alpha reliability of 0.93.

Certain subject variables explored in order to determine whether individual differences among participants influenced their ratings of physical attractiveness and fecundity. First, the question arose whether the participants' own body fat (as measured by the body mass index) could have influenced their ratings of the model's physical attractiveness and fecundity. In addition, participants' satisfaction of their own bodies could also have potentially affected their ratings. To answer these questions, a 2 (gender) by 2 (WHR) by 3 (weight) ANCOVA was computed with body mass index (BMI) and participant ratings of personal body satisfaction as covariates. No significant effects of participants' BMI or body satisfaction were found for ratings of physical attractiveness or fecundity, and the initial main effects of gender, weight and WHR remained unchanged as a result of this analysis.

Discussion

The present study aimed to compare two competing hypotheses regarding determinants of female physical attractiveness. The findings support the sociocultural perspective that in industrial societies body weight is a significant predictor of female physical attractiveness. While Singh's research indicated that female stimuli of normal body weights were judged more favourably than both underweight and overweight figures, the present investigation showed that models were judged to be most attractive in the underweight conditions, followed by normal and overweight categories. Although females generally rated the models as being more attractive than did males, both female and male ratings showed a significant positive effect of thinness on judgements of attractiveness. These results support a great deal of research regarding the relationship between body weight and physical attractiveness of females (Andersen and DiDomenico 1992; Wiseman et al. 1992) and suggest that a sociocultural ideal of thinness may indeed predominate.

Although results showed that waist-to-hips ratio (WHR) also had a main effect on physical attractiveness ratings, this effect did not reflect the relationship between WHR and physical attractiveness that has been proposed by Singh (1993a, 1993b, 1994a, 1994b, 1995a). Where Singh's research has consistently reported female stimuli with an 'ideal' WHR to be judged as most attractive compared to females with a WHR above the ideal range, these findings were not replicated in the present study. Instead, both male and female participants found the model with the high WHR significantly more attractive than the model with the ideal WHR across all weight conditions. Thus, these findings clearly call into question Singh's evolutionary argument that a lower WHR should receive the highest attractiveness ratings because of its health and reproductive signals.

The fact that participants favoured the model with the straighter figure also fits with past research indicating that the ideal female body has become thinner and more tube-shaped. For example, Morris, Cooper and Cooper (1989) examined physical characteristics of female fashion models over the time period 1967–87, and showed that models' body shapes had become less curvaceous and more tubular. Another study reported similar results in an investigation of the changing body shape of female models in *Ladies Home Journal* and *Vogue* magazines from 1901 to 1987 and among thirty-eight film actresses between the 1930s and 1980s (Silverstein, Perdue, Peterson and Kelly 1986). The authors reported that the standard of bodily attractiveness among the women in the study was less curvaceous than in the past. In fact, the authors stated that 'in the mid-20's, the only other time this century when the models in *Vogue* and *Ladies Home Journal* were as non-curvaceous as they are now, an epidemic of eating disorders appeared among young women' (1986: 532). Thus, although Singh's research points to attractiveness as encompassing a curvaceous ideal body shape, findings from the present study as well as from past research suggest a thin tubular bodily standard of female attractiveness, a standard which does not reflect Singh's findings.

With respect to the present findings regarding judgements of fecundity, the absence of a significant difference between ratings of the two models further casts doubt on previous research by Singh. According to evolutionary hypotheses, the WHR is an indicator of reproductive ability and health status of females, and it is proposed that a low WHR should be conceptualized as being most fecund because it is associated with increased reproductive health. By contrast, a female with a high WHR should be less fecund because her WHR is linked to a greater risk of various diseases and reduced reproductive capability (Singh 1995a). However, if this were the case we should expect to see a consensus among participants that Model A (WHR = 0.72) was significantly more fecund than Model B (WHR = 0.86), whose WHR was clearly outside and above the ideal range proposed by Singh. This was not the case. Rather, the data indicated that participants' judgements of fecundity were influenced most by perceived body weight of the models. Using body weight, students appropriately judged females of normal weight to be more fecund than females who were underweight (who may have difficulty achieving regular menses necessary for reproduction) or overweight (which brings increased health risks). Although there is no simple explanation for the lack of a main effect for WHR, it may be that the adaptive significance of the WHR is not as strong in today's media-dominated modern environment as when it first evolved in ancestral populations as a primary factor of mate selection.

The secondary analyses carried out also provide support to the present results, showing that participants were able to distinguish appropriately all three weight conditions, and indicating that their preferences for underweight figures and for a high WHR were consistent throughout ratings of size and shape of the models' body parts. The between-subjects design employed in this investigation further strengthens the main findings, as the experimental manipulations of models' weights and their differing WHRs were not known to participants. In consideration of the present findings and previous results reported by Tassinary and Hansen (1998), it seems fair to conclude that perceived body weight has a more powerful relationship to physical attractiveness than perceived WHR.

Perhaps one of the most important findings in the present study was the striking agreement among male and female participants in their significant preferences for underweight female figures. In fact, Model B (WHR 0.86) who was judged to be significantly more attractive than Model A (WHR 0.72) across all weight categories, actually approached an anorexic weight in the underweight condition. That is, included in the diagnostic criteria for anorexia nervosa is weight loss leading to a body weight of 85 per cent of one's expected weight for height, or reaching a BMI of less than 17.5 (American Psychiatric Association 1994). In the underweight condition, the estimated BMI for Model B was 17.05, a level below the anorexic cut-off. Thus, of all six photographic images produced to depict the models in the three weight conditions, an anorexic female figure received the highest ratings of physical attractiveness by both male and female students. In addition, females require 17 per cent body fat to begin menses at puberty and 22 per cent body fat to maintain regular menstrual cycles (Mustajoki 1992). Thus, participants' preferences for Model B in the underweight condition also serve to contradict Singh's hypothesis that fecundity and physical attractiveness are highly related, as participants in the present study favoured a female figure whose weight substantially reduced her reproductive ability.

These results clearly indicate the importance of body weight in conceptualizing female attractiveness, and they point towards the larger picture of sociocultural pressures facing women in Western industrialized societies which prescribe excessively thin body ideals and which have implications for eating disorders. For example, recent research evidence of media influences on eating disorders has indicated that the increase in eating pathology in recent decades has paralleled a decrease in the body weight of ideal female images depicted in the media (Wiseman et al. 1992). In addition, endorsement and internalization of the thin ideal has been found to predict bulimia diagnoses and eating-disorder symptoms (Stice and Shaw 1994).

There are several limitations to the present study, which require some examination. Although the female figures used in this experiment reflected two distinct WHRs, the inclusion of a third model with a WHR well below the ideal range for females (such as a WHR of 0.6) may have produced a more comprehensive analysis. Because a WHR of such small numerical value is rare among women (and was too difficult to obtain for the present study), the development of computer programs which can directly manipulate the WHR of a figure should be a consideration for future research.

A related concern arises with the use of two human models in the present study. Although the utilization of photographic images was a substantial improvement upon artificial line-drawn stimuli in previous studies, the ability to select a single model and manipulate her WHR and body weight would allow full control of any extraneous factors contributing to judgements of physical attractiveness or fecundity. For example, although every effort was made in this experiment to keep the height of the female figures constant, there were slight height differences between the two models. In this case, differences in height were not expected to have any influence on student ratings as the difference was not detectable in the photographs and because the between-subjects design allowed students to view only one figure. However, potential problems like these could be avoided by making computer-generated manipulations on a single model.

A third limitation involves several findings in the present study which remain unclear and which require additional research to draw more confident conclusions. First, there is no straightforward explanation for the finding that females generally judged the models to be more attractive than the males did. Although it could be hypothesized that females gave higher ratings because they shared the same gender as the models, other interpretations could be equally plausible.

The secondary analyses indicated that a subject's own body fat, as judged by BMI, did not influence the outcome of the analyses of attractiveness or fecundity. However, we did not collect information on subjects' WHR. Conceivably this variable could have influenced the results. While this might be addressed in future studies, it seems an unlikely candidate as a major influence. BMI and WHR are highly correlated and both predict body fat and health outcomes (Willett and Manson 1995). Thus, it would be expected that if BMI did not influence outcome, neither would WHR.

Finally, a concern emerges with respect to the generalizability of the sample in the experiment, as participants were all first-year university students. It would be useful to compare a wider community sample of males and females to determine whether main effects of weight, WHR and gender hold for judgements of female physical attractiveness. It would also be of interest to conduct experiments of this nature with other cultural groups.

For example, one would not expect the present results if the study was carried out in a non-Western, less industrialized culture. However, if the study was carried out with participants who immigrated from such a culture to a Western industrialized culture, one would expect the same results as in the present study.

With regards to the concept of the WHR, there is also room for future exploration. Clearly, the role of the WHR in predicting physical attractiveness is not as straightforward as proposed by Singh, and it remains to be fully understood in this context. The WHR of a person's figure can be manipulated by increasing or decreasing the size of the waist or hips. It will be of interest for future research to determine whether there are independent effects for waist and hips in judgements of attractiveness and fecundity. Preliminary research by Tassinary and Hansen (1998) suggests that relative hip size may have a potent influence on attractiveness and fecundity over and above the effect of the WHR.

Author note

Rebecca M. Puhl, Department of Psychology; Fred J. Boland, Associate Professor, Department of Psychology, Queen's University.

Thank you to David Duchane for his expertise in constructing the computer-generated images, to the females who posed as models for the study, and to those who volunteered to participate in this project.

Correspondence concerning this article should be addressed to Rebecca M. Puhl who is now at the Department of Psychology, Yale University, Kirtland Hall, 2 Hillhouse Avenue, New Haven, CT 06520-8205, USA. E-mail: rebecca.puhl@yale.edu.

References

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, Washington, DC: American Psychiatric Association.

Andersen, A. and DiDomenico, L. (1992) 'Diet vs. shape content of popular male and female magazines: a dose response to the incidence of eating disorders?' *International Journal of Eating Disorders* 11: 283–7.

Brownell, K. (1991) 'Dieting and the search for the perfect body: where physiology and the culture collide', *Behavior Therapy* 22: 1–12.

Canada Expert Group on Weight Standards (1988) *Canadian Guidelines for Healthy Weights: Report of an Expert Group Convened by Health Promotion Directorate, Health Services and Promotion Branch*. Ottawa, ON: Health and Welfare Canada.

DiNicola, V.F. (1990) 'Anorexia multiforme: self-starvation in historical and cultural context. Part II: Anorexia nervosa as a culture-reactive syndrome', *Transcultural Psychiatric Research Review* 27: 245–86.

Duchane, D. (1999) 'A computer generated body image device', Unpublished master's thesis, Queen's University, Kingston, Ontario, Canada.

Fallon, A. and Rozin, P. (1985) 'Sex differences in perceptions of desirable body shape', *Journal of Abnormal Psychology* 94: 102–5.

Franzoi, S. and Herzog, M. (1987) 'Judging physical attractiveness: what body aspects do we use?' *Personality and Social Psychology Bulletin* 13: 19–33.

Furnham, A., Tan, T. and McManus, C. (1997) 'Waist-to-hip ratio and preferences for body shape: a replication and extension', *Personality and Individual Differences* 22: 539–49.

Henss, R. (1995) 'Waist-to-hip ratio and attractiveness. Replication and extension', *Personality and Individual Differences* 19: 479–88.

Morris, A., Cooper, T. and Cooper, P. (1989) 'The changing shape of female fashion models', *International Journal of Eating Disorders* 8: 593–6.

Mustajoki, R. (1992) 'Could mannequins menstruate?' *British Medical Journal* 305: 1575–6.

Orne, M.T. (1962) 'On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications', *American Psychologist* 17: 776–83.

Rosenthal, R. and Rosnow, R. (1975) *The Volunteer Subject*, New York: John Wiley.

Rozin, P. and Fallon, A. (1988) 'Body image, attitudes to weight, and misperceptions of figure preferences of the opposite sex: a comparison of men and women in two generations', *Journal of Abnormal Psychology* 97: 342–5.

Seid, R.P. (1994) 'Too "close to the bone": the historical context for women's obsession with slenderness', in P. Fallon, M.A. Katzman and S.C. Wooley (eds), *Feminist Perspectives on Eating Disorders*, New York: Guildford Press.

Silverstein, B., Perdue, L., Peterson, B. and Kelly, E. (1986) 'The role of the mass media in promoting a thin standard of bodily attractiveness for women', *Sex Roles* 14: 519–32.

Silverstein, B., Peterson, B. and Perdue, L. (1986) 'Some correlates of the thin standard of bodily attractiveness for women', *International Journal of Eating Disorders* 5: 895–905.

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

Singh, D. (1993b) 'Is thin really beautiful and good? Relationship between waist-to-hip ratio (WHR) and female attractiveness', *Personality and Individual Differences* 16: 123–32.

Singh, D. (1994a) 'Ideal female body shape: role of body weight and waist-to-hip ratio', *International Journal of Eating Disorders* 16: 283–8.

Singh, D. (1994b) 'Body fat distribution and perception of desirable female body shape by young black men and women', *International Journal of Eating Disorders* 16: 289–94.

Singh, D. (1995a) 'Female health, attractiveness, and desirability for relationships: role of breast asymmetry and waist-to-hip ratio', *Ethology and Sociobiology* 16: 465–81.

Singh, D. (1995b) 'Female judgement of male attractiveness and desirability for relationships: role of waist-to-hip ratio and financial status', *Journal of Personality and Social Psychology* 69: 1089–101.

Singh, D. and Luis, S. (1994) 'Ethnic and gender consensus for the effect of waist-to-hip ratio on judgements of women's attractiveness', *Human Nature* 6: 51–65.

Singh, D. and Young, R. (1995) 'Body weight, waist-to-hip ratio, breasts, and hips: role in judgements of female attractiveness and desirability for relationships', *Ethology and Sociobiology* 16: 483–507.

Spillman, D. and Everington, C. (1989) 'Somatotypes revisited: have the media changed our perception of the female body image?' *Psychological Reports* 64: 887–90.

Stice, E. and Shaw, H. (1994) 'Adverse effects of the media portrayed thin-ideal on women and linkages to bulimic symptomology', *Journal of Social and Clinical Psychology* 13: 288–308.

Stice, E., Schupak-Neuberg, E., Shaw, H. and Stein, R. (1994) 'Relation of media exposure to eating disorder symptomology: an examination of mediating mechanisms', *Journal of Abnormal Psychology* 103: 836–40.

Tassinary, L. and Hansen, K. (1998) 'A critical test of the waist-to-hip-ratio hypothesis of female physical attractiveness', *Psychological Science* 9: 150–5.

Wilfley, D.E. and Rodin, J. (1995) 'Cultural influences on eating disorders', in K.D. Brownwell and C.G. Fairburn (eds) *Eating and Obesity: A Comprehensive Handbook*, New York: Guildord Press, pp. 78–82.

Willet, W.C. and Manson, J.E. (1995) 'Epidemiological studies of health risks due to excess weight', in K.D. Brownell and C.G. Fairburn (eds), *Eating Disorders and Obesity: A Comprehensive Handbook*, New York: Guilford Press.

Wiseman, C., Gray, J., Mosimann, J. and Ahrens, A. (1992) 'Cultural expectations of thinness in women: an update', *International Journal of Eating Disorders* 11: 85–9.