ELSEVIER

Evolution and Human Behavior 24 (2003) 88–98

Evolution and Human Behavior

Waist–hip ratio and attractiveness New evidence and a critique of "a critical test"

Sybil A. Streeter, Donald H. McBurney*

Department of Psychology, University of Pittsburgh, 4429 Sennott Square, Pittsburgh, PA 15260, USA

Received 18 March 2002; accepted 20 August 2002

Abstract

An evolutionary model of mate choice predicts that humans should prefer honest signals of health, youth, and fertility in potential mates. Singh and others have amassed substantial evidence that the waist–hip ratio (WHR) in women is an accurate indicator of these attributes, and proposed that men respond to WHR as an attractiveness cue. In response to a recent study by Tassinary and Hansen [Psychol. Sci. 9 (1998) 150.] that purports to disconfirm Singh's hypothesis, we present evidence showing a clear relationship between WHR and evaluations of attractiveness. We evaluated responses to a range of waist, hip, and chest sizes, spanning the 1st through 99th percentiles of anthropometric data. Waist, hip, and chest sizes were altered independently to give WHRs of 0.5, 0.6, 0.7, 0.9, and 1.2. We replaced line drawings with more realistic computer-manipulated photographs. The preferred WHR was 0.7, concordant with the majority of previous results. By asking participants to estimate weight in each stimulus figure, we were able to statistically control for the effects of weight on attractiveness judgments; the effect of WHR remained.

Keywords: Waist–hip ratio; WHR; Female attractiveness; Evolutionary psychology

1. Introduction

An evolutionary model predicts that humans should prefer honest signals of health, youth, and fertility in potential mates (Buss, 1989). Waist–hip ratio (WHR) in women is

* Corresponding author.

1090-5138/03/$ – see front matter © 2003 Elsevier Science Inc. All rights reserved.
PII: S1090-5138(02)00121-6

an indicator of these attributes (e.g., Singh, 1993a; Wass, Waldenstrom, Rossner, & Hellberg, 1997; Zaadstra et al., 1993). Before puberty, boys and girls both have a WHR of about 1.0. At puberty, female WHR declines to 0.7 and rises again at menopause (Ley, Lees, & Stevenson, 1992). Increased WHR is related to decreased fertility, and can be the result of conditions such as pregnancy, menopause, and polycystic ovary syndrome (PCOS) (Pirwany, Fleming, Greer, Packard, & Sattar, 2001; Wass et al., 1997; Zaadstra et al., 1993). Other factors less directly related to fertility, such as diabetes, Kwashiorkor, and cretinism, can also result in increased WHR with decreased fertility as a correlate.

Many studies have found that both men and women find a female WHR of 0.7 most attractive (Furnham, Lavancy, & McClelland, 2001; Furnham, Tan, & McManus, 1997; Henss, 1995, 2000; Singh, 1993a, 1993b, 1994a, 1994b, 1994c, 1994d, 1995; Singh & Luis, 1994; Singh & Young, 1995), but these findings have not gone unchallenged. Some have argued that WHR is insignificant in relation to weight, as measured by body mass index (BMI) (Tovée & Cornelissen, 1999, 2001; Tovée, Maisey, Emery, & Cornelissen, 1999; Tovée, Mason, Emery, McCluskey, & Cohen-Tovée, 1997; Tovée, Reinhart, Emery, & Cornelissen, 1998; Yu & Shepard, 1998). BMI also predicts health and fertility (Tovée et al., 1999, and references therein). We will address this issue in the Discussion section.

Wetsman and Marlowe (1999) find that the hunter–gatherer Hadza prefer higher WHR than 0.7, and suggest that the Hadza base their attractiveness ratings on weight because starvation risk in this society makes fat stores a better predictor of fecundity than WHR. Singh and Luis (1994), however, found wide cross-cultural agreement on a preference for WHR of 0.7. Although they did include nonwestern populations, none of their data come from hunter–gatherer societies.

Manning, Trivers, Singh, and Thornhill (1999) propose a reason independent of weight for a preference for higher WHR. They found that women with higher WHRs had given birth to more sons than daughters, and propose that in cultures where boy children are preferred, the optimal WHR would be driven slightly higher than 0.7. However, a prospective study of WHR and sex ratio of offspring did not confirm the results of Manning et al. (Tovée, Brown, & Jacobs, 2001).

Tassinary and Hansen (T&H) (1998) challenge the validity of the stimuli used to test the hypothesis in most WHR studies, on the grounds that Singh's stimuli confound WHR with weight. They developed their own set of drawings, supposedly free from such confounds, but as we will discuss later, they did not succeed in unconfounding WHR and weight. T&H conclude that attractiveness actually increases with increases in WHR, contrary to most prior studies, but this conclusion depends entirely on the way they chose to plot their data: when plotting attractiveness as a function of WHR for a set of stimuli that vary both by waist and hip size, one can connect either the data points with the same hip size or those with the same waist size, and by choosing the latter, T&H also chose their conclusion. Fig. 1 shows T&H's data plotted both ways. The dotted lines connect the data points by waist size, as T&H plotted them, and imply that attractiveness increases with higher WHR, but the solid lines show the data plotted by hip size, and indicate that attractiveness *decreases* with higher WHR, as all others have found. We will discuss later which way of plotting the data is more appropriate.

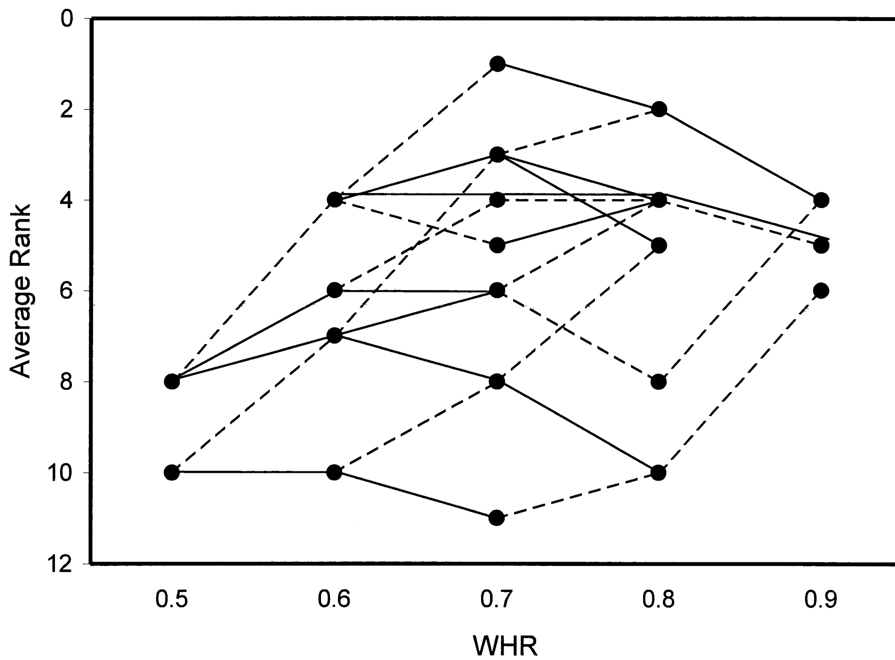

Fig. 1. Attractiveness as a function of WHR (data of T&H, 1998). Solid lines connect figures with the same hip size. Dotted lines connect figures with the same waist size. See text.

Data we present in this paper, however, show the conventional relationship regardless of which way they are plotted.

2. Method

2.1. Stimuli

Our study differs in several ways from previous experiments. First, we included a wider range of WHRs. Previous research has often used 0.7 as the lowest WHR, and it has also been found to be the most preferred, leaving open the possibility that even lower ratios might be even more preferred and that males simply prefer the female with the lowest WHR (Jones, 1997; T&H, 1998). T&H used a range of WHRs from 0.5 to 1.0. We used 1.2 as the upper limit, based on the theoretical limits determined by a systematic evaluation of anthropometric measurements (Gordon et al., 1989).

Most previous investigators have used the same upper body size for all stimuli. In order to determine if the proportion between chest and hip is important, we also varied chest size. From anthropometric surveys (Gordon et al., 1989), we set the extreme values associated with waist, hip, and chest measurements for use in our stimulus figures as in Table 1. By combining the largest hip with the smallest waist, the smallest possible expected WHR

Table 1
Dimensions of figures and average ratings by figure

Figure	WHR	Waist	Hip	Chest	Mean weight estimate	Mean attractiveness rating	Adjusted attractiveness rating
1	0.5	S	L	S	161.9 (±18.9)	3.79	−1.16
2	0.6	S	M	S	134.1 (±18.4)	2.04	−0.17
3	0.6	M	L	S	159.9 (±15.1)	4.65	−0.91
4	0.7	S	S	S	118.7 (±12.3)	3.90	−0.13
5	0.7	M	M	S	132.5 (±12.4)	2.09	−0.18
6	0.7	L	L	S	159.4 (±22.8)	4.09	−0.88
7	0.9	M	S	S	121.7 (±12.8)	2.38	−0.51
8	0.9	L	M	S	148.9 (±15.1)	2.28	−0.93
9	1.2	L	S	S	158.2 (±22.2)	1.69	−0.74
10	0.5	S	L	M	166.0 (±21.7)	2.13	−0.64
11	0.6	S	M	M	139.0 (±13.1)	5.03	1.17
12	0.6	M	L	M	162.0 (±17.7)	2.89	0.05
13	0.7	S	S	M	130.9 (±11.3)	5.39	1.11
14	0.7	M	M	M	142.3 (±11.3)	5.13	1.40
15	0.7	L	L	M	164.7 (±19.3)	2.88	0.06
16	0.9	M	S	M	131.0 (±10.5)	5.27	1.04
17	0.9	L	M	M	156.5 (±15.0)	2.93	−0.14
18	1.2	L	S	M	161.5 (±20.2)	2.54	−0.34
19	0.5	S	L	L	167.4 (±19.9)	2.21	−0.48
20	0.6	S	M	L	151.4 (±15.1)	3.92	0.61
21	0.6	M	L	L	174.0 (±20.3)	2.99	0.62
22	0.7	S	S	L	144.1 (±14.4)	3.91	0.26
23	0.7	M	M	L	151.9 (±15.1)	4.44	1.04
24	0.7	L	L	L	174.3 (±21.3)	2.78	0.42
25	0.9	M	S	L	144.0 (±14.1)	3.73	0.09
26	0.9	L	M	L	164.4 (±18.4)	2.54	−0.24
27	1.2	L	S	L	168.9 (±22.6)	2.46	−0.22

would be approximately 0.5. The largest waist measurement combined with the smallest hip resulted in a WHR of approximately 1.2. These combinations include some that are more than two standard deviations from the mean, but are nevertheless encountered, especially in individuals with developmental anomalies such as cretinism and conditions such as Kwashiorkor.

The resulting stimuli include some that appear unusual, and certainly unattractive. To leave them out for that reason, however, would be to preclude determining the attractiveness of the full range of anatomically possible body shapes, and amounts to begging the question.

2.2. Participants

Ninety-five undergraduate Introductory Psychology students (38 women and 57 men) participated for course credit. All were over the age of 18, and were treated in accordance with the ethical principles of the American Psychological Association (1992). Data from two men were dropped for failure to follow instructions.

2.3. Stimulus materials

A photograph of one woman was obtained by permission from Henss (1998), and was manipulated using Adobe PhotoShop software. The original photograph depicts a young woman in 3/4 view wearing a form-fitting garment. By altering waist, hip, and chest widths, we modified the photograph to produce 27 different figures, with the proportions listed in Table 1. These figures included three sizes each of hip, waist, and chest (small, medium, and large). Combining the three values of each variable resulted in stimuli that varied slightly from the desired ratios. Minor adjustments to waists were required in three cases in order to ensure exact ratios. Stimuli used by previous investigators allowed the WHR to vary somewhat from the nominal ratio. Relative hip and chest dimensions were comparable; that is, "small" chest was the same width as "small" hip, etc. This procedure afforded the opportunity to consider ratios of chest to hip and chest to waist, in addition to WHR. Breast size was proportionate to the chest width.

2.4. Design

Combining the 3 waists × 3 hips × 3 chests gave 27 figures, resulting in WHRs of 0.5 ($n=3$), 0.6 ($n=6$), 0.7 ($n=9$), 0.9 ($n=6$), and 1.2 ($n=3$). Each hip and chest size was associated with three WHRs. Thus, the experiment was organized as a 3 (waist) × 3 (hip) × 3 (chest) × 2 (sex) factorial design.

This experimental design differed from most previous designs in two additional ways. We asked participants to rate the figures on a seven-point scale, rather than use rank ordering (Singh, 1993a, 1994a, 1994b, 1994c, 1994d; Singh & Luis, 1995; T&H, 1998), and subjects viewed the stimulus pictures individually rather than concurrently. These two considerations are related: 27 stimuli far exceed the capacity of working memory.

2.5. Procedure

Participants were scheduled in groups of 10 or fewer. Each subject judged the stimuli in a different random order, with the restriction that each figure appeared first an equal number of times. They first assessed attractiveness on a seven-point scale (extremely unattractive to extremely attractive), recording their ratings in response booklets, one page per rating. After completing these ratings, they were instructed to estimate the weight in pounds for the same set of figures (in the same random order), and were told to assume that the woman was average height (5 ft 4 in.), and that the average weight for a woman of that height was 140 lb. Again, the estimates were made in booklets accompanying the binders, with one page per rating. Upon completing their ratings, participants were debriefed and were thanked for their participation.

3. Results

3.1. Weight estimates

Average weight estimates are given in Table 1. Estimates varied from a low of 118.7 to a high of 174.3 lb, corresponding to BMIs of 20 and 30, respectively.

3.2. Waist–hip ratio

Fig. 2 shows attractiveness ratings as a function of WHR. A WHR of 0.7 was most preferred [$F(4,87) = 106.37$, $P < .001$]. Both sexes rated 0.7 most attractive, but men rated the figures as more attractive than did women [$F(1,90) = 10.50$, $P = .002$), and also responded more to variation in WHR [Sex × WHR interaction: $F(1,4) = 4.72$, $P = .001$].

Following a suggestion by Bronstad and Singh (D. Singh, personal communication, April 1999), we regressed attractiveness ratings on weight estimates for each subject in order to examine the effect of WHR independent of weight. The residuals, hereafter referred to as "adjusted attractiveness," were then used for all subsequent analyses.

Linear regression showed that estimated weight was responsible for 66% of the variance in attractiveness ratings of the figures. Even when the effect of weight was controlled, 0.7 remained the preferred WHR [$F(4,87) = 54.12$, $P < .001$], with attractiveness decreasing with distance from 0.7. All pairwise comparisons of adjusted values for WHR were significantly

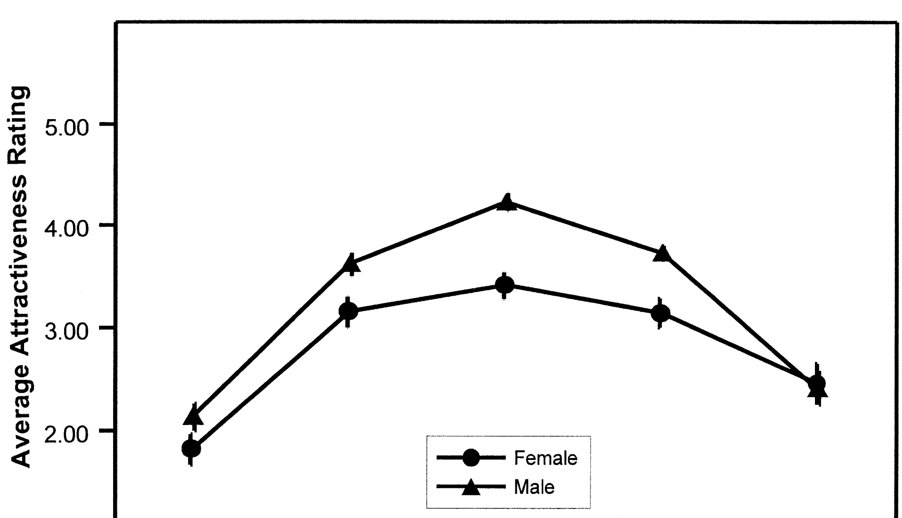

Fig. 2. Attractiveness as a function of WHR and sex of rater. Error bars show ±1 standard error of the mean.

different from one another at the $P = .001$ level or better, with the exception of the figures with WHR of 0.6 and 0.9, which did not differ.

Fig. 3 shows all data as a function of both hip size and chest size using the adjusted attractiveness values. The least favored WHRs are 0.5 and 1.2; all the figures with these extreme WHRs were among the lowest in attractiveness.

3.3. Chest, hip, and waist size

The main effect of chest size was highly significant [$F(2,89) = 107.551$, $P < .001$]. Medium chest size was most preferred, followed by large and then small. Preferences for all chest sizes differed from each other at the .003 level or better.

The main effect of hip size was highly significant [$F(2,89) = 38.65$, $P < .001$]. Medium hip size was most preferred, followed by small and then large. Preferences for all hip sizes differed from each other at the .001 level or better. Hip size differences had the greatest effect on weight estimates. The effect of range of hip size (small to large) on weight estimate was 23.40 lb; for chest size, 16.14 lb; waist size, 15.90 lb.

The main effect of waist size was highly significant [$F(2,89) = 34.07$, $P < .001$]. Medium waist size was most preferred, followed by small and then large. All waists differed significantly from each other at $P < .001$.

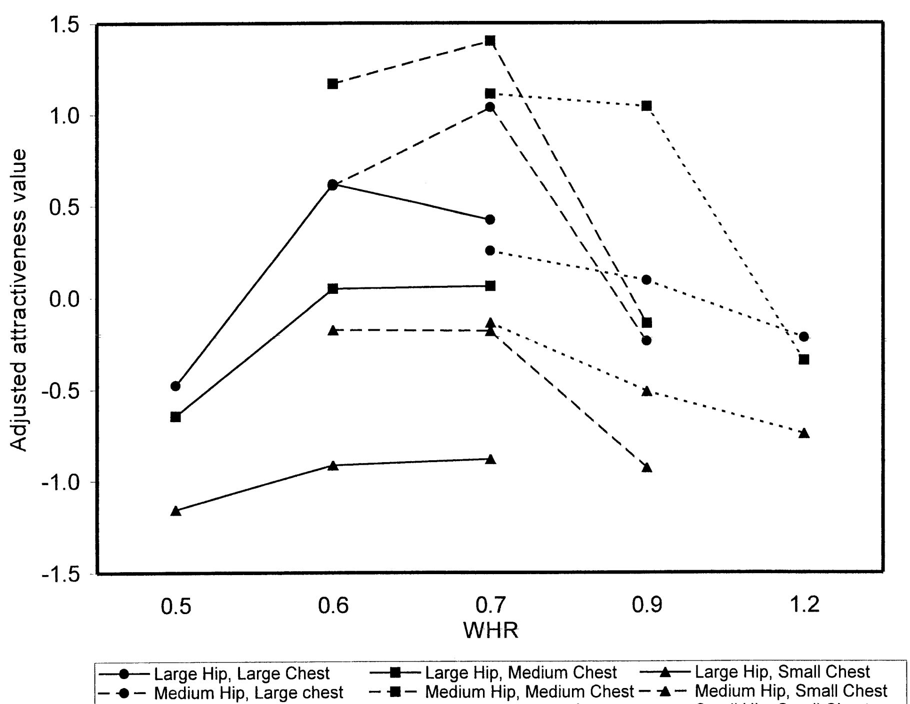

Fig. 3. Attractiveness values adjusted for the effect of estimated weight as a function of WHR.

3.4. Effect sizes

We analyzed the relative effects of the independent variables using η^2. We found that chest size had the largest effect (.547), followed by WHR (.324), hip size (.303), and waist (.277). Because the analysis was done on the weight-adjusted data, there is no effect size for weight.

4. Discussion

The present data confirm the results of Singh and others who found that a WHR of 0.7 is most preferred. This is true regardless of whether the data are plotted by connecting points with the same hip size as in Fig. 3, or the same waist size, as T&H did. Fig. 4 shows our data plotted with waist size as the parameter, following T&H, and the general pattern still favors 0.7: for six of the nine waist and chest size combinations, 0.7 was the most preferred WHR. Only one of the waist sizes shows a monotonic trend favoring the highest WHR (and 0.7 is the highest in the series).

Which way of plotting the data is most appropriate? Either is defensible, but we maintain that the WHR effect will be most clearly revealed when the variable responsible for the most variance in the judgments is used as the parameter. Both T&H and the present study found the effect of hip size to be larger than the waist effect, considerably so in the data of T&H.

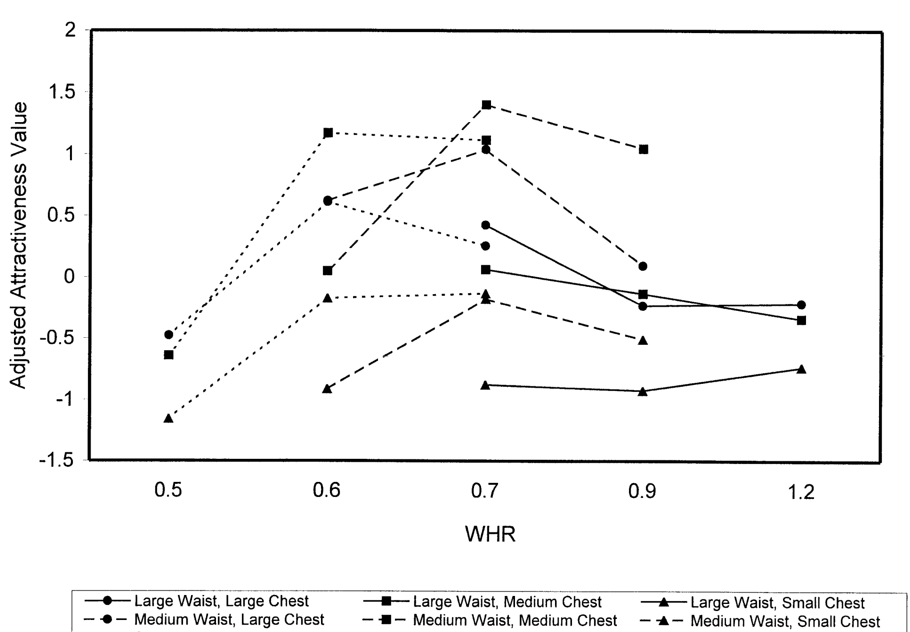

Fig. 4. Same data as Fig. 3 plotted by connecting figures with the same waist size, following T&H (1998).

It is obvious that the proportion of variance accounted for in a set of data depends crucially on the ranges of the variables investigated. Tovée and Cornelissen (1999, 2001), for example, used figures varying in weight from emaciated to obese. (Assuming hypothetical women of a constant height, weight would have varied fourfold in their figures, variability that is presumably outside the range that obtained in the EEA.) By contrast, WHR varied relatively little (0.68–0.98). A less extreme range of BMI would have reduced the variance accounted for by weight. It would have been a simple matter to show that some other feature accounted for more variance, by disfiguring the faces or making the bodies asymmetrical in some of the stimuli, for example. For this reason, we do not consider the percent of variance accounted for by BMI relative to WHR to be of particular interest.

T&H point out correctly that Singh's figures confound WHR and weight. They claim that these variables are manipulated orthogonally in their own figures. However, this is impossible to do (Henss, 2000) because one cannot vary waist or hip in a factorial design without affecting weight. We approached the problem of confounding in two ways. First, we obtained judgments of weight for all stimuli, and regressed preference on estimated weight; this removed the effect of weight from the data. Second, we plotted our data in both of the alternative ways discussed above, and found the same effect of WHR.

T&H (1998) and Tovée et al. (1997, 1998, 1999, 2001) maintain that weight is more important than WHR in judgments of attractiveness. We do not dispute this. The present study did find that weight accounted for 66% of the total variance. The important question, however, is not whether weight accounts for more variance than WHR, but whether WHR predicts preference when the effect of weight is removed.

The arguments above in favor of WHR as an important predictor of attractiveness independent of weight are purely empirical. We should not lose sight of the fact that the concept of WHR is supported by a wealth of theoretical evidence concerning health and fertility (Pirwany et al., 2001; Wass et al., 1997; Zaadstra et al., 1993). The two measures reflect two separate predictors of fertility. WHR signals youth and health; weight signals resources to support a pregnancy as well as general health. A recent study by Pawlowski and Dunbar (unpublished, as cited by Barrett, Dunbar, & Lycett, 2002) found that "WHR was the better predictor of neonatal weight . . . in women weighing over 54 kg, but BMI was the better predictor in lighter women" (p. 109).

T&H presented their figures to the subjects grouped according to weight category (light, moderate, and heavy). In an effort to maintain comparability between the data from groups of figures, they presented one figure from each of the other weight categories in each group. This attempt is methodologically flawed. It has been known since the work of Helson (1964) that judgments are influenced markedly by the range of stimuli presented. In this case, the judgment of the stimulus intended to provide comparability would be influenced by the nine stimuli from the category of interest.

In summary, the effect of WHR in these data is clear. We found the expected non-monotonic relationship by using the full range of possible WHRs. Additionally, we show this relationship when the effect of weight on attractiveness judgments is removed. Singh's claim that WHR predicts attractiveness in females is supported by this study.

Acknowledgments

This report is based in a Bachelor of Philosophy thesis at the University Honor's College, University of Pittsburgh, by SAS under the direction of DMcB. We thank UHC for generous support, and Alec Sarkas for preparing the photographs. The data were previously reported at the annual meeting of the Human Behavior and Evolution Society, Salt Lake City, UT, 1999.

References

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, *47*, 1597–1611.

Barrett, L., Dunbar, R., & Lycett, J. (2002). *Human evolutionary psychology.* Princeton: Princeton University Press.

Buss, D. M. (1989). Sex differences in human mate preferences: evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*, 1–49.

Furnham, A., Lavancy, M., & McClelland, A. (2001). Waist to hip ratio and facial attractiveness: a pilot study. *Personality and Individual Differences*, *30*, 491–502.

Furnham, A., Tan, T., & McManus, C. (1997). WHR and preferences for body shape: a replication and extension. *Personality and Individual Differences*, *22*, 539–549.

Gordon, C. C., Churchill, T., Clauser, C. E., Bradtmilller, B., McConville, J. T., Tebbbets, I., & Walker, R. A. (1989). *1988 Anthropometric survey of U.S. Army personnel: methods and summary statistics: final report*. Natick, MA: U.S. Army Natick Research, Development and Engineering Center (NATICK/TR-89/044, AD A225 094).

Helson, H. (1964). *Adaptation-level theory: an experimental and systematic approach to behavior*. New York: Harper & Row.

Henss, R. (1995). WHR and attractiveness. Replication and extension. *Personality and Individual Differences*, *19*, 479–498.

Henss, R. (1998). *Waist-to-hip ratio and attractiveness of the human figure.* Retrieved November 1998, from University of the Saarland: http://www.cops.uni-saarland.de/ronald/PICTURES/WHRARCHV/WHRJPG/FIG4_1.JPG.

Henss, R. (2000). WHR and female attractiveness. Evidence from photographic stimuli and methodological considerations. *Personality and Individual Differences*, *28*, 501–513.

Jones, D. (1997). An evolutionary perspective on physical attractiveness. *Evolutionary Anthropology*, *5*, 97–109.

Ley, C. J., Lees, B., & Stevenson, J. C. (1992). Sex- and menopause-associated change in body fat distribution. *American Journal of Clinical Nutrition*, *55*, 950–954.

Manning, J. T., Trivers, R. L., Singh, D., & Thornhill, R. (1999). The mystery of female beauty. *Nature*, *399*, 214–215.

Pirwany, I. R., Fleming, R., Greer, C. J., Packard, C. J., & Sattar, N. (2001). Lipids and lipoprotein subfractions in women with PCOS: relationship to metabolic and endocrine parameters. *Clinical Endocrinology*, *54*, 447–453.

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

Singh, D. (1993b). Body shape and women's attractiveness: the critical role of waist–hip ratio. *Human Nature*, *4*, 297–321.

Singh, D. (1994a). Is thin really beautiful and good? Relationship between WHR and female attractiveness. *Personality and Individual Differences*, *16*, 123–132.

Singh, D. (1994b). Ideal female body shape: role of body weight and WHR. *International Journal of Eating Disorders*, *16*, 283–288.

Singh, D. (1994c). WHR and judgments of attractiveness and healthiness by male and female physicians. *International Journal of Obesity*, *18*, 731–737.

Singh, D. (1994d). Body fat distribution and perception of desirable female body shape by young black men and women. *International Journal of Eating Disorders*, *16*, 289–294.

Singh, D. (1995). Female health, attractiveness and desirability for relationships: role of breast asymmetry and WHR. *Ethology and Sociobiology*, *16*, 465–481.

Singh, D., & Luis, S. (1994). Ethnic and gender consensus for the effect of WHR on judgment of women's attractiveness. *Human Nature*, *6*, 51–65.

Singh, D., & Young, R. (1995). Body weight, WHR, breasts and hips: role in judgments of female attractiveness and desirability for relationships. *Ethology and Sociobiology*, *16*, 483–507.

Tassinary, L. G., & Hansen, K. A. (1998). A critical test of the WHR hypothesis of female physical attractiveness. *Psychological Science*, *9*, 150–155.

Tovée, M. J., Brown, J. E., & Jacobs, D. (2001). Maternal waist-to-hip ratio does not predict child gender. *Proceedings of the Royal Society of London*, *268*, 1007–1010.

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

Tovée, M. J., & Cornelissen, P. L. (2001). Female and male perceptions of female physical attractiveness in front-view and profile. *British Journal of Psychology*, *92*, 391–402.

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*, *266*, 211–218.

Tovée, M. J., Mason, S. M., Emery, J. L., McCluskey, S. E., & Cohen-Tovée, E. M. (1997). Supermodels: stick insects or hourglasses? *Lancet*, *9089*, 1474–1475.

Tovée, M. J., Reinhardt, S., Emery, J. L., & Cornelissen, P. L. (1998). Optimum body-mass index and maximum sexual attractiveness. *Lancet*, *352*, 548.

Wass, P., Waldenstrom, U., Rossner, S., & Hellberg, D. (1997). An android body fat distribution in females impairs the pregnancy rate of in-vitro fertilization–embryo transfer. *Human Reproduction*, *12*, 2057–2060.

Wetsman, A., & Marlowe, F. (1999). How universal are preferences for female waist-to-hip ratios? Evidence from the Hadza of Tanzania. *Evolution and Human Behavior*, *20*, 219–228.

Yu, D. W., & Shepard, G. H. (1998). Is beauty in the eye of the beholder? *Nature*, *396*, 321–322.

Zaadstra, B. M., Seidell, J. C., Van Noord, P. A., te Velde, E. R., Habbema, J. D., Vrieswijk, B., & Karbaat, J. (1993). Fat and female fecundity: prospective study of body fat distribution on conception rates. *British Medical Journal*, *306*, 484–487.