Research Report

A CRITICAL TEST OF THE WAIST-TO-HIP-RATIO HYPOTHESIS OF FEMALE PHYSICAL ATTRACTIVENESS

Louis G. Tassinary[1] and Kristi A. Hansen[2]

[1]Environmental Psychophysiology Laboratory and [2]Department of Architecture, College of Architecture, Texas A&M University

Abstract—*Recent research has led to increasingly sophisticated conjectures as to the roles that genetic heritage, prior experience, and environmental context play in the production and maintenance of complex behaviors. The field of evolutionary psychology was born of such conjectures (Stanley, 1895) and now serves as a niche for a growing number of researchers (Buss, 1995; Kenrick, 1994). One of the more provocative lines of experimental research to emerge from this alembic derives from the linkage of evolutionary theories of human mate selection with definitions of physical attractiveness based on somatic characteristics that simultaneously signal attractiveness and predict reproductive potential (Buss, 1989). The waist-to-hip ratio (WHR) has been purportedly demonstrated to be a robust example of just such an invariant perceptual cue (e.g., Henss, 1995; Singh, 1993a). Here, we report that judgments of attractiveness and fecundity can be either unrelated or related, positively or negatively, to the WHR depending on waist size, hip size, and weight, and are thus inconsistent with the evolutionary argument that human physical attractiveness is fundamentally a sign of mate value.*

Recent studies have demonstrated that simple variations in female body shape and size are potent determinants of perceived attractiveness, age, reproductive potential, health status, and personality (Henss, 1995; Singh, 1993a, 1993b, 1994a, 1994b, 1994c; Singh & Luis, 1995; Singh & Young, 1995), and these results have captured the attention of scientists as well as the lay public (Cowley, 1996; Stanish, 1996). These experiments varied the weight and waist-to-hip ratio (WHR) of two-dimensional representations of a female figure and consistently found that both men and women find a figure of moderate weight with a WHR of approximately 0.7 to be the most attractive and fecund. When combined with the evidence that the WHR is a predictor of actual fecundity (e.g., Zaadstra et al., 1993), these results have been interpreted to provide support for a utilitarian model of human mate selection in which attractiveness is a marker for mate quality and have been extolled as necessary for the development of biologically credible explanations of evaluation (Tesser & Martin, 1996). Videlicet, the specification of a "social" stimulus (Arnoult, 1963) has advanced the science of social psychology by allowing a quantitative function to be defined over objectively describable stimuli and observable responses.

Unfortunately, despite the allure of the argument, nearly all previous studies have employed the same set of stimuli, which confounded weight with hip size, as well as WHR with relative waist size. In addition to these limitations, no studies have systematically examined WHRs of less than 0.7. WHRs smaller than 0.7 are of theoretical interest because of the potential to test runaway models of human mate selection in which attractiveness is a marker for arbitrary but noncapricious aesthetic preferences (Kirkpatrick & Ryan, 1991). To overcome these limitations, we constructed a new set of simple stimuli in which we systematically and independently varied weight, waist size, and hip size (cf. Brandon, 1994; Platt, 1964).

Address correspondence to Louis G. Tassinary, Environmental Psychophysiology Laboratory, College of Architecture, Texas A&M University, College Station, TX 77843-3137; e-mail: lou@archone.tamu.edu.

METHOD

Research Participants

One hundred and thirty-six undergraduates (83 women, 53 men) received credit toward fulfilling a course requirement in return for their voluntary participation and were treated in accordance with the "Ethical Principles of Psychologists and Code of Conduct" (American Psychological Association, 1992). The self-reported median age, height, weight, and body mass index[1] were 18 years, 65 in., 125.0 lb, and 17.25 for the women and 18 years, 72 in., 165 lb, and 18.75 for the men, respectively.

Stimulus Materials

All stimuli consisted of black line drawings composited on 8.5- by 11-in. white paper. The stimulus set taken directly from previously published research (hereafter referred to as *confounded*) consisted of 12 line drawings of a female figure that varied in weight (light, moderate, heavy) and WHR (0.7, 0.8, 0.9, 1.0) (see Fig. 1). The new stimulus set (hereafter referred to as *independent*) comprised 27 line drawings of a female figure that varied in weight (light, moderate, heavy), waist size (small, medium, large), and hip size (small, medium, large) (see Fig. 2). The independent set was generated based on an appropriately scaled normal-weight female, 64 in. in height, and all modifications were consistent with anthropometric norms (Tilley, 1993; Young et al., 1983). The linear dimensions of the waists and hips of these figures were slightly modified to represent more accurately the anatomical fact that a greater percentage of the hip circumference is due to variation in the saggital plane. Two different random orders for each set were used to determine the location of each figure on a page. For the confounded set, the figures were placed within a rectangular area as three rows with four figures per row. For the independent set, the figures were placed within a pentagonal area with each figure the same distance from each of its neighbors.

Design and Procedure

The experiment consisted of two phases. All participants were asked to rank-order individually four sets of line drawings, once with respect to attractiveness and once with respect to capability of bearing

1. Quetelet's Index (kg/m^2) served as the body mass index. Self-reported weight and height were given in pounds and inches and converted to kilograms and meters prior to calculation of this index.

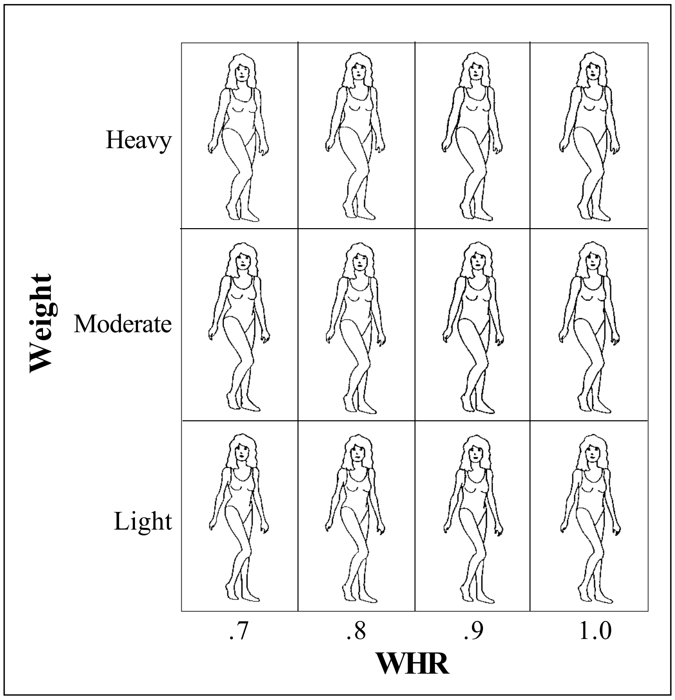

Fig. 1. Confounded stimulus figures originally designed to represent three body-weight categories (rows) and four waist-to-hip ratios (columns). Modified and redrawn from Singh (1993a).

children. Order of set presentation and task order were counterbalanced across eight groups, with group size ranging from 12 to 23 participants. The confounded stimuli used in prior research were presented simultaneously and ranked from 1 to 12. The independent stimuli were presented in three successive subsets of 11 figures, with 3 figures representing 0.7 WHRs (one at each weight category, all of medium waist and hip size) repeated in each subset, and ranked from 1 to 11.

All analyses reported are based on the entire group of 136 participants. We initially included the sex and body mass index of each subject as independent variables but subsequently dropped them because their influence was consistently small and insignificant.[2]

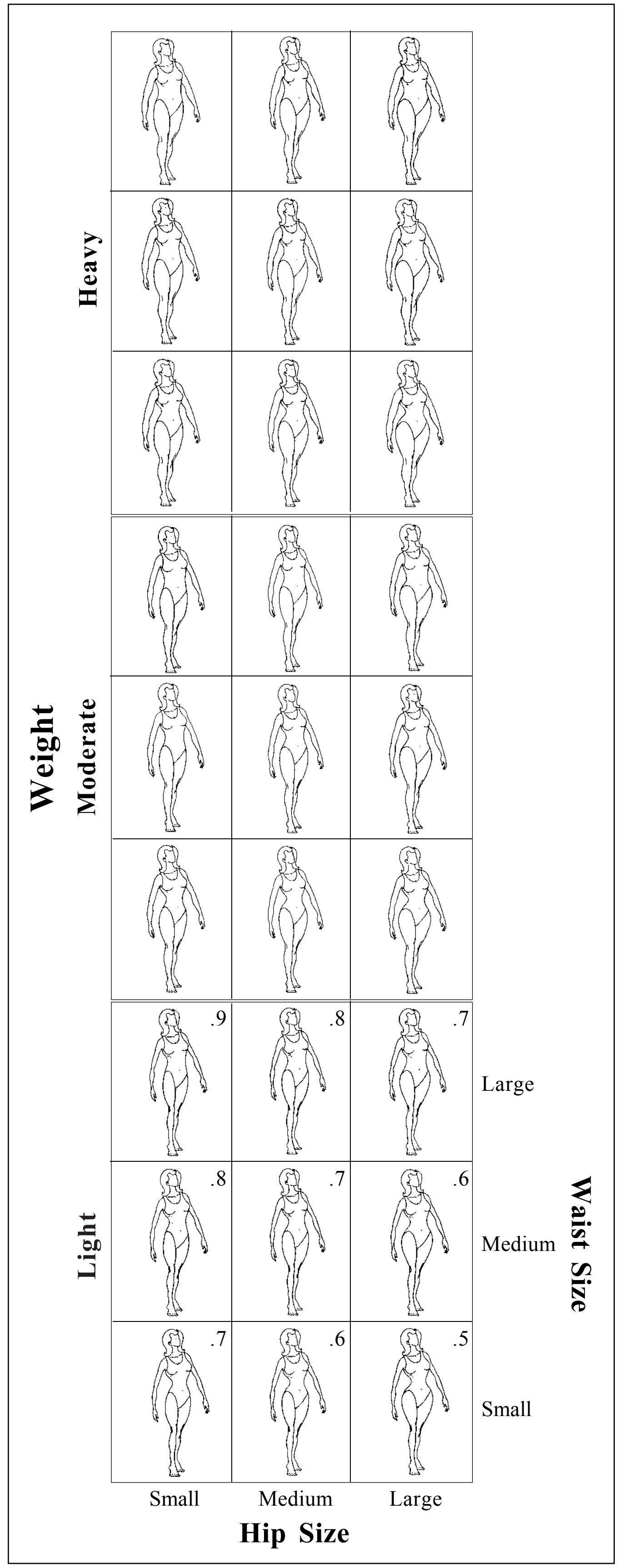

Fig. 2. New independent stimulus figures representing three body-weight categories (squares), three relative waist sizes (rows), and three relative hip sizes (columns). The small numbers in the upper right-hand corners of the cells in the lower square are the waist-to-hip ratios for the figures. These ratios are identical in the corresponding cells of the middle and upper squares.

2. The lone statistically significant sex difference occurred with the confounded stimulus set and emerged as a significant three-way interaction (Sex × Weight × WHR), $F(6, 804) = 2.292$, $p < .05$. A visual inspection of the 24 means revealed that the only dis-ordinal effect occurred for attractiveness judgments of the moderate- and light-weight figures with a WHR of 0.8. In this case, the female participants, on average, ranked the moderate-weight figure as less attractive than the light-weight figure (4.4 vs. 3.7), whereas the male participants, on average, ranked the moderate-weight figure as more attractive than the light-weight figure (3.2 vs. 4.2).

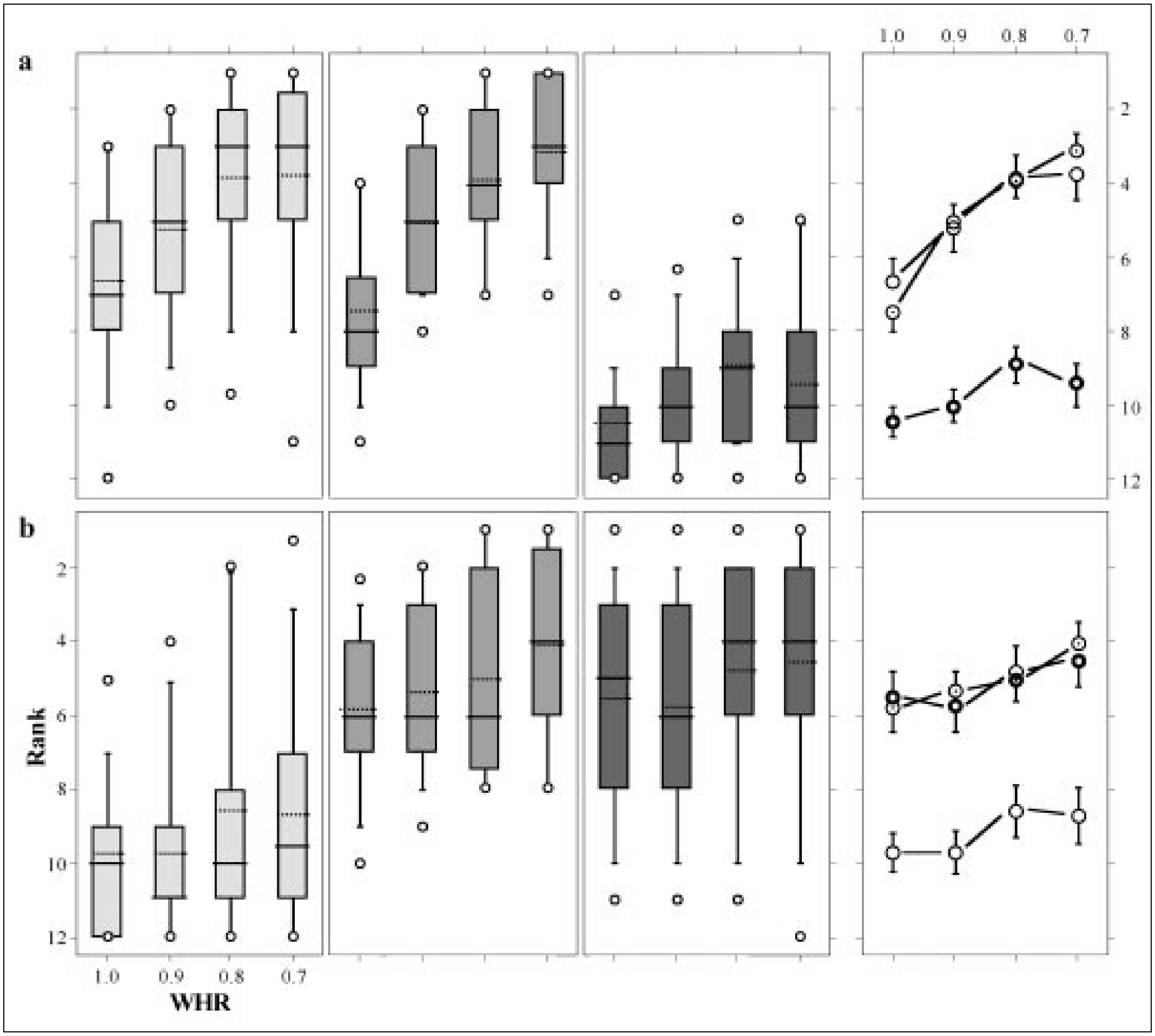

Fig. 3. Attractiveness (a) and fecundity (b) rankings of the confounded stimulus set for all subjects. The first three plots in each panel display box plots of the distribution of ranks for each figure. In both panels, the three weight categories of light, moderate, and heavy are represented by the colors light, medium, and dark gray, respectively. Each box indicates the interquartile range, the solid and dotted lines inside the box indicate the median and arithmetic mean, the capped bars indicate the 10th and 90th percentiles, and the circles indicate the data in the upper and lower vigintiles. The last plot in each panel plots the average rank as a function of both weight category and waist-to-hip ratio (WHR), and the three weight categories of light, moderate, and heavy are represented by open, dotted, and partially filled circles, respectively. The error bars represent the 99% confidence intervals.

RESULTS AND DISCUSSION

The utilitarian interpretation of the WHR hypothesis predicts that conspecifics will judge human female figures with 0.7 WHRs to be highly attractive and fecund because this particular value is a marker for a distribution of subcutaneous fat that is causally related to the ability to bear children successfully (Singh, 1993a). Using the prior confounded stimuli, we successfully replicated previous findings. The female figures with the higher WHRs were judged generally to be both less attractive and less fecund in all weight categories, the effect of WHR was most pronounced for the figures in the moderate-weight category, and the figure with a WHR of 0.7 in the moderate weight category was judged to be the most attractive and fecund of the entire set of figures (see Fig. 3). As in previous studies, however, we also found the weight of the figure to be a much more potent factor than the WHR. Light- and moderate-weight figures were judged to be much more

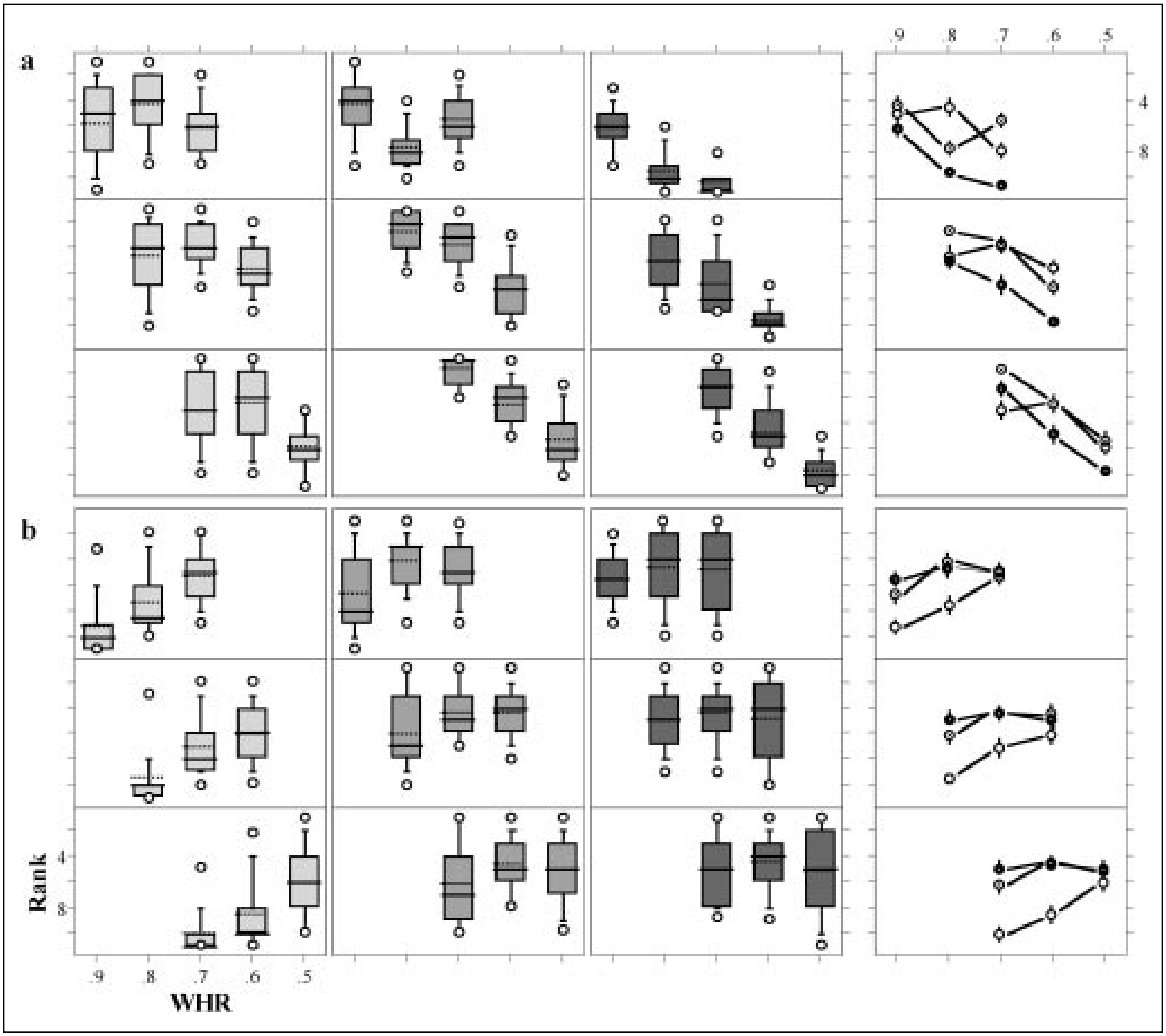

Fig. 4. Attractiveness (a) and fecundity (b) rankings of the independent stimulus set for all subjects. The first three plots in each panel display box plots of the distribution of ranks for each figure. In both panels, the three weight categories of light, moderate, and heavy are represented by the colors light, medium, and dark gray, respectively. Each window within each plot displays the data for a given relative waist size, from small (bottom) to large (top). Within each of these windows, relative hip size varies from small (left) to large (right). Each box indicates the interquartile range, the solid and dotted lines inside the box indicate the median and arithmetic mean, the capped bars indicate the 10th and 90th percentiles, and the circles indicate the data in the upper and lower vigintiles. The last plot in each panel plots the average rank as a function of weight category, relative waist size, and waist-to-hip ratio (WHR), and the three weight categories of light, moderate, and heavy are represented by open, dotted, and partially filled circles, respectively. The error bars represent the 99% confidence intervals.

attractive than the heavy figures, whereas moderate-weight and heavy figures were judged to be much more fecund than the light figures.

Could the apparent positive association between the WHR, judged attractiveness, and judged fecundity be an artifact of a limited stimulus set, or is the WHR the phenotype that links attractiveness and mate quality? The results for the new stimulus set provide evidence decidedly in favor of the former alternative. As with the previous stimuli, light- and moderate-weight figures were judged to be much more attractive than the heavy figures, whereas moderate-weight and heavy figures were again judged to be much more fecund than the low-weight figures (see Fig. 4). If the utilitarian interpretation of the WHR hypothesis is correct, a WHR of 0.7 should be ranked as most attractive and fecund relative to larger and smaller WHRs, regardless of waist or hip size. If the aesthetic-preference interpretation of the WHR hypothesis is correct, then a WHR of 0.5 should be ranked as the most attractive and fecund relative to larger WHRs, regardless of weight. The data, however, show little evidence for either effect. Rather, they show clear evidence for two strong main effects; that is, relative hip size and weight are both positively associated with ranked fecundity and negatively associated with attractiveness, regardless of WHR. In addition, the distributions of idiographic rank-order correlations between attractiveness and fecundity provided no evidence of consensus (see Fig. 5). For the confounded stimuli, the median Spearman rank correlation across participants was –.05, with a range of –.95 to +.93, and the confidence interval for the median correlation included 0.00. For the independent stimuli, the median Spearman rank correlation across participants was –.09, with a range of –.95 to +.88, and the confidence interval for the median correlation again included zero.

The pattern of the effect sizes for the various factors across the two experiments is also consistent with the artifactual interpretation of previous findings (see Fig. 6). The effect size for WHR is reduced by more than 50% for attractiveness rankings and more than 90% for fecundity rankings using the independent set rather than the confounded set. In addition, the large effect size for weight with the confounded set appears to be decomposable into weight and hip factors when using the independent set. The obtained pattern also suggests that relative hip size is a more potent influence on attractiveness than is relative weight (cf. Singh & Young, 1995, Study 2), whereas the opposite is true for fecundity.

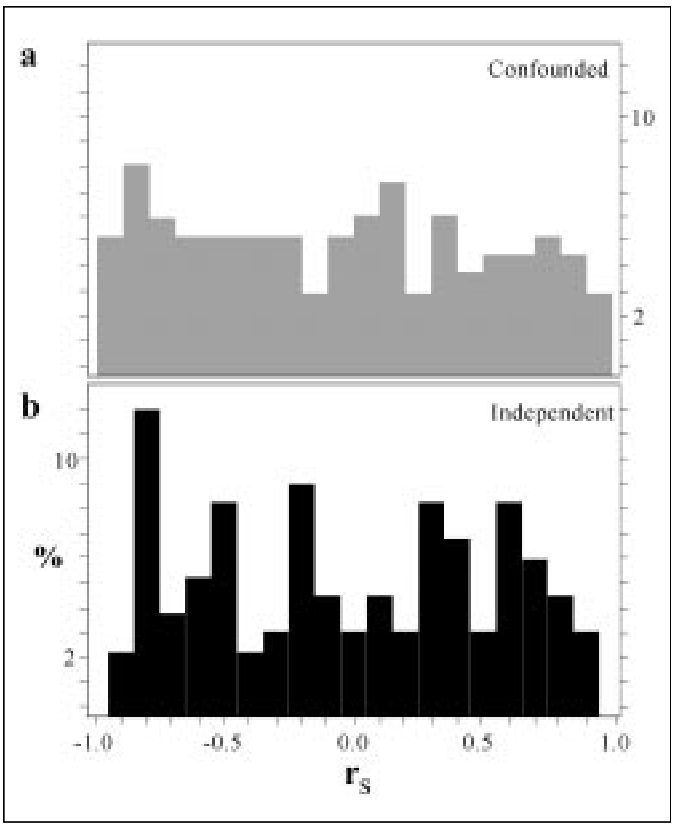

Fig. 5. Histograms of idiographic correlation coefficients between attractiveness and fecundity rankings. The *y*-axis represents the percentage of participants, and the *x*-axis the size of the Spearman rank correlation in steps of 0.1. Distributions based on the confounded stimulus set (a) and the independent stimulus set (b) are shown.

Despite the fact that we replicated previous findings using the confounded stimuli and constructed our independent stimuli based on explicit anthropometric guidelines, it is possible that our participants perceived our new stimuli as qualitatively different or simply found the extended ranking task to be onerous and unnatural. If so, this would certainly compromise the validity of our more rigorous test of the WHR hypothesis. Fortunately, three of the new independent stimuli are analogous to three of the old confounded stimuli, and their rankings can be compared directly. The mean rankings and confidence intervals for these three figures are plotted in Figure 7. As is visually apparent, the mean rankings for these figures occupy comparable regions of the phase space defined by relative attractiveness and fecundity.

These findings demonstrate that weight and hip size are important and independent co-determinants of both relative attractiveness and fecundity, and that the WHR is of marginal importance for predicting relative attractiveness. Further, they indicate that perceived attractiveness and fecundity can be either positively related, unrelated, or inversely related, at both the individual and the group level. This pattern of results is inconsistent with either the *good sense* or the *good taste* model of

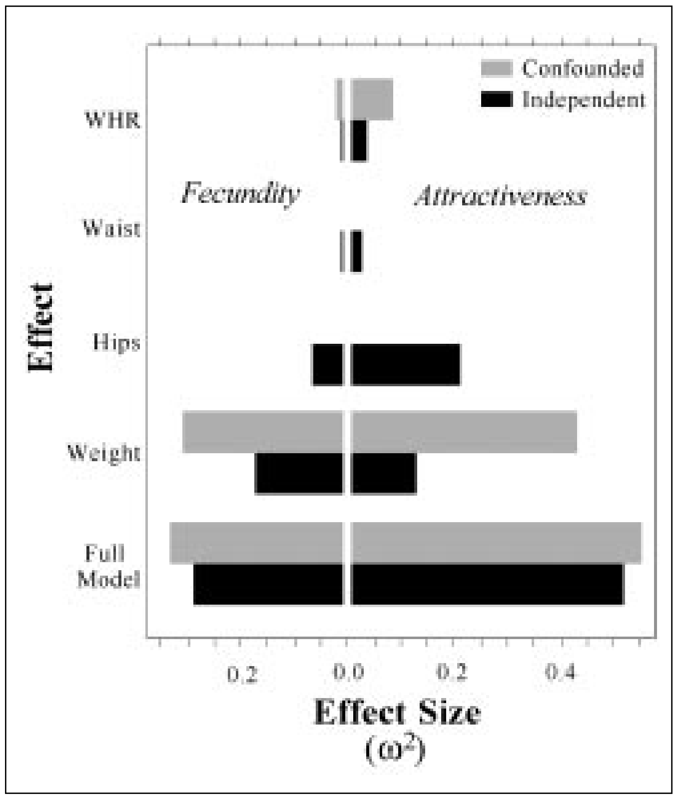

Fig. 6. Effect sizes for the major independent variables. These relative magnitudes were computed based on the complete factorial repeated measures analysis of variance using the Type 3 (partial) sums-of-squares model according to published procedures (Vaughan & Corballis, 1969). The magnitudes for the full model represent all main effects and higher order interactions combined. The magnitudes for the waist-to-hip-ratio (WHR) effect are based on a specific planned contrast designed to represent the WHR hypothesis. For the confounded stimulus set, for the WHRs of 0.7, 0.8, 0.9, and 1.0, the weights were .75, .25, –.25, and –.75, respectively. For the independent stimulus set, for the WHRs of 0.5, 0.6, 0.7, 0.8, and 0.9, the weights were –.4, –.1, 1.0, –.1, and –.4, respectively.

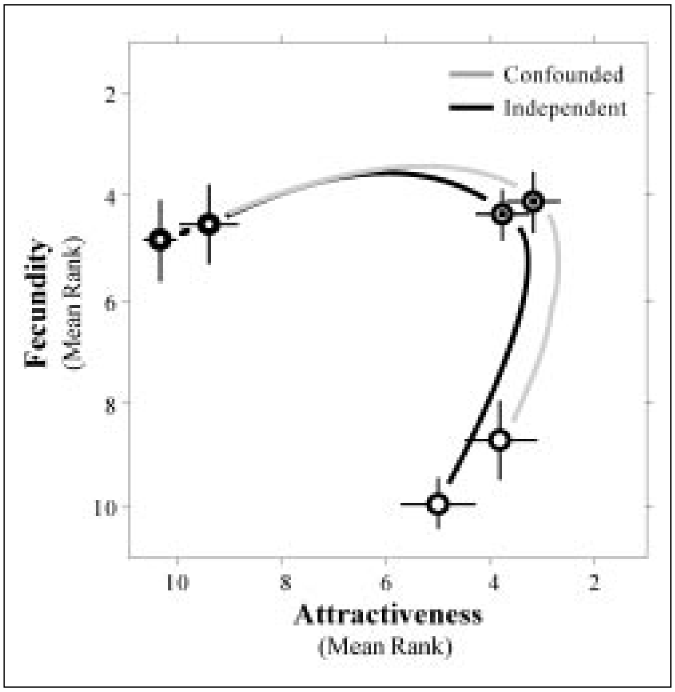

Fig. 7. Plot of three figures from the two stimulus sets whose proportions are directly comparable. All have waist-to-hip ratios of 0.7 and vary conjointly in weight, waist, and hip size. The data for the confounded stimulus set are based on the stimulus figures in the first column of Figure 1. The data for the independent stimulus set are based on the stimulus figures found in the lower left, middle, and upper right cells of the lower, middle, and upper squares in Figure 2. The weight categories of light, moderate, and heavy are represented by open, dotted, and partially filled circles, respectively. Error bars represent the 99% confidence intervals.

human female attractiveness (cf. Cronin, 1991; Perper, 1989), and thus constitutes a clear and unambiguous disconfirmation of the WHR hypothesis. At present, we do not have an explanation for these results that is as simple and elegant as this hypothesis. We suspect, however, that people use some very simple, probably learned, and yet-to-be-fully-articulated heuristics when evaluating figures such as these (Allgeier & Wiederman, 1994; Brown, Cash, & Noles, 1987; Salusso-Deonier, Markee, & Pederson, 1993). We also suspect the WHR, not unlike the golden section (Chandler, 1934, pp. 33–45) or the grand mean (Perrett, May, & Yoshikawa, 1994), will eventually be understood to be a dimensionless number with great intuitive appeal but with highly circumscribed explanatory or predictive efficacy with respect to aesthetic judgments.

Acknowledgments—This study was supported in part by an NSF Presidential Faculty Fellowship to the first author and formed part of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science to the second author. Portions of this research were presented at the annual Life-like Characters Conference (September 1995; Snow Bird, Utah) and the Texas A&M Graduate Student Poster Competition (1996). We thank D. Bontempo, S. King, K. Carrol, and B. Zgabay for general assistance, and W. Graziano, M. Ihrig, M. Shepley, E. Vanman, and W. Wood for criticism.

REFERENCES

Allgeier, E.R., & Wiederman, M.W. (1994). How useful is evolutionary psychology for understanding contemporary human sexual behavior? *Annual Review of Sex Research, 5,* 218–256.

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

Arnoult, M. (1963). The specification of a "social" stimulus. In S.B. Sells (Ed.), *Stimulus determinants of behavior* (pp. 16–30). New York: Ronald Press.

Brandon, R.N. (1994). Theory and experiment in evolutionary biology. *Synthese, 99,* 59–73.

Brown, T.A., Cash, T.F., & Noles, S.W. (1987). Perception of physical attractiveness among college students: Selected determinants and methodological matters. *The Journal of Social Psychology, 126,* 305–316.

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

Buss, D.M. (1995). Evolutionary psychology: A new paradigm for psychological science. *Psychological Inquiry, 6,* 1–30.

Chandler, A.R. (1934). *Beauty and human nature.* New York: D. Appleton-Century.

Cowley, G. (1996, June 3). The biology of beauty. *Newsweek,* 61–69.

Cronin, H. (1991). *The ant and the peacock: Altruism and sexual selection from Darwin to today.* New York: Cambridge University Press.

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

Kenrick, D.T. (1994). Evolutionary social psychology: From sexual selection to social cognition. *Advances in Experimental Social Psychology, 26,* 75–121.

Kirkpatrick, M., & Ryan, M.J. (1991). The evolution of mating preferences and the paradox of the lek. *Nature, 350,* 33–38.

Perper, T. (1989). Theories and observations on sexual selection and female choice in human beings. *Medical Anthropology, 11,* 409–454.

Perrett, D.I., May, K., & Yoshikawa, S. (1994). Attractive characteristics of female faces: Preference for non-average shape. *Nature, 368,* 239–242.

Platt, J.R. (1964). Strong inference. *Science, 146,* 347–353.

Salusso-Deonier, C.J., Markee, N.L., & Pederson, E.I. (1993). Gender attractiveness in the evaluation of physical attractiveness ideals for male and female body builds. *Perceptual and Motor Skills, 76,* 1155–1167.

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-to-hip ratio. *Human Nature, 4,* 297–321.

Singh, D. (1994a). 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. (1994b). Ideal female body shape: Role of body weight and waist-to-hip ratio. *International Journal of Eating Disorders, 16,* 283–288.

Singh, D. (1994c). Is thin really beautiful and good? Relationship between waist-to-hip ratio (WHR) and female attractiveness. *Personality and Individual Differences, 16,* 123–132.

Singh, D., & Luis, S. (1995). Ethnic and gender consensus for the effect of waist-to-hip ratio on judgment of women's attractiveness. *Human Nature, 6,* 51–65.

Singh, D., & Young, K. (1995). Body weight, waist-to-hip ratio, breasts, and hips: Role in judgments of female attractiveness and desirability for relationships. *Ethology and Sociobiology, 16,* 484–507.

Stanish, M. (1996, June 1). Hip deep in controversy: UT professor throws researchers a curve with study on waist-hip ratios. *Austin American-Statesman,* pp. C1, C12.

Stanley, H.M. (1895). *Studies in the evolutionary psychology of feeling.* London: Swan Sonnenshein.

Tesser, A., & Martin, L. (1996). The psychology of evaluation. In E.T. Higgins & A.W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 400–422). New York: Guilford Press.

Tilley, A.R. (1993). *The measure of man and woman: Human factors in design.* New York: Whitney Library of Design.

Vaughan, G.M., & Corballis, M.C. (1969). Beyond tests of significance: Estimating the strength of effects in selected ANOVA designs. *Psychological Bulletin, 72,* 204–213.

Young, J.W., Chandler, R.F., Snow, C.C., Robinette, K.M., Zehner, G.F., & Lofberg, M.S. (1983). *Anthropometric and mass distribution characteristics of the adult female.* Oklahoma City, OK: FAA Civil Aeromedical Institute. (NTIS No. FAA-AM-83-16)

Zaadstra, B.M., Seidell, J.C., van Noord, P.A.H., te Velde, E.R., Habbema, J.D.F., Vrieswijk, B., & Karbaat, J. (1993). Prospective study of effect of body fat distribution on conception rates. *British Medical Journal, 306,* 484–487.

(RECEIVED 6/25/97; REVISION ACCEPTED 10/7/97)