The New School Psychology Bulletin
2016, Vol. 13, No. 2


Print 2016 ISSN: 1931-793X; Online ISSN: 1931-7948

# Development of a Body Dissatisfaction Scale Assessment Tool

Gabriella J. Mutale
Nottingham Trent University

Andrew Dunn
Nottingham Trent University

James Stiller
Nottingham Trent University

Rebecca Larkin
Nottingham Trent University

The present research details the development of a new pictorial scale (Body Dissatisfaction Scale) to be used in the measurement of body dissatisfaction. The scale comprises nine female and nine male images of computer generated bodies that increase successively in body weight. Using a sample of 190 students (female = 130, male = 60) results showed that the new scale exhibits good validity, with participants being able to correctly identify body weight differences between all bodies in the scale. Evidence for construct validity was demonstrated by significant correlations between ratings of perceived actual body size using the current scale and participants' BMI. Body dissatisfaction measured using the current scale was also negatively correlated with a measure of body appreciation. Test-retest reliability remained stable over a 5 week period. The scale improves on previous pictorial scales by offering both male and female versions while offering improved realism and consistency between images.

*Keywords*: body image, body dissatisfaction, body weight, scale development

As the occurrence of eating disorders in young women has grown in Western society in recent years (Smink, van Hoeken, & Hoek, 2012), research into body image dissatisfaction has increased, with a large amount of research highlighting a link between body image disturbances and eating disorders (Cattarin & Thompson, 1994; Garner, 2002). Indeed, a large number of females and males in Western society are dissatisfied with some aspect of their bodies (Tiggemann, 2011). Consequently, the measurement of body dissatisfaction is an important aspect of research concerned with body image dissatisfaction and disturbances.

Grogan (2008) defines body dissatisfaction as "a person's negative thoughts about his or her own body" (p. 4). This includes judgements about size, shape, and muscle tone and generally involves a discrepancy between one's own body type and an ideal body type. Although there are several techniques for measuring a discrepancy between an individual's own body weight and their ideal body weight, figure rating scales are most commonly used (e.g., Gardner, Jappe, & Gardner, 2009; Stunkard, Sorensen, & Schulsinger, 1983; Thompson & Gray, 1995). This type of scale typically consists of a set of drawn stimuli that vary in body weight from underweight to overweight. To measure body dissatisfaction, participants are normally asked to choose a figure they think (a) best represents their perceived actual body shape and (b) best represents their ideal body shape. Body dissatisfaction is then defined as the discrepancy between a participant's actual and ideal body.

There has been extensive use of these types of figure rating scales (e.g., Gardner et al., 2009; Stunkard et al., 1983; Thompson & Gray, 1995) to represent different body weights when researching body dissatisfaction. This has proved useful in enabling researchers to keep the properties of stimuli consistent while changing waist to hip ratios, body mass, and size. However, there are various problems

*Corresponding authors:*
Gabriella Mutale, gabriella.mutale2005@my.ntu.ac.uk
Andrew Dunn, andrew.dunn@ntu.ac.uk
James Stiller, james.stiller@ntu.ac.uk
Rebecca Larkin, rebecca.larkin@ntu.ac.uk

with these types of stimuli. For example, line-drawn figures have been said to lack realism and ecological validity and can therefore often appear unrealistic (Tassinary & Hansen, 1998). Thompson and Gray (1995) noted that previous line drawn scales have figures with disproportionate arms and legs, differential thickness between right and left arms, or a lack of separation between the arms and bodies in obese drawings. This has led to further criticism of research into physical attraction (e.g., Singh, 1993; Swami, Furnham, & Joshi, 2008) that has previously relied on line-drawn stimuli (Bateson, Cornelissen, & Tovée, 2007; Gardner, Friedman, & Jackson, 1998). Often, line-drawn figures (e.g., Singh, 1993) that vary in waist to hip ratio (WHR) are modified by altering the width of the torso, but this also changes apparent Body Mass Index (BMI), meaning it is not possible to say whether attractiveness ratings are made on the basis of WHR, BMI or both. In addition, these figures are usually drawn by a professional artist (e.g., Thompson & Gray, 1995). Therefore, the increase in body weight between figures is subject to the artist's estimation of increase in body weight, rather than a metrically precise increment. This results in inconsistent size differences between successive figures (Gardner et al., 1998).

To address some of the previous problems with the measurement of figural stimuli, Gardner et al. (2009) developed a line-drawn figure rating scale where the bodies were based on real body dimensions. However, although these bodies can be said to change in size, the way real bodies do, using line-drawn stimuli still lacks realism (Tassinary & Hansen, 1998). Therefore, to address these issues with line-drawn stimuli, Swami, Salem, Furnham, and Tovée (2008a) developed The Photographic Figure Rating Scale (PFRS). This scale comprises 10 photographic images of real women varying in BMI from extremely thin to obese and has been used to assess body dissatisfaction.

Although the PFRS offers improvement on line-drawn scales by using images of real bodies, this can cause problems when trying to maintain consistency. For example, it has been noted that the bodies used in the PFRS vary in leg length (Swami et al., 2008a). This is problematic since leg to body ratio has been shown to affect attractiveness preferences in female bodies (Swami, Einon, & Furnham, 2006) and, therefore, when being used in the measurement of body dissatisfaction, some bodies may appear more attractive not because of their size but because of a more appealing leg to body ratio. In addition, the bodies used in the PFRS had their faces obscured in order to maintain consistency across images. It could be argued that this also reduces the ecological validity of the images as bodies in real life are rarely seen without faces. Most importantly, there is currently no male version of the PFRS.

Accordingly, our goal was to develop and report the psychometric evaluation of a new pictorial measure of body image that has two comparable versions which can be used with both females and males. The Body Dissatisfaction Scale (BDS) comprises images of computer generated bodies. Using computer generated bodies improves on the drawn stimuli used in previous figure rating scales (e.g., Stunkard et al., 1983; Thompson & Gray, 1995). Computer generated images are more realistic and life-like while also allowing for accurate control of size and shape differences between successive bodies. Using computer generated stimuli allows variation in body proportions, such as leg to body ratio, to be controlled for. This cannot be controlled for when using photographic images. It also gives greater control over variation in physical characteristics like skin and facial features (identity). Therefore, as skin and facial features can be controlled, there is no need to obscure faces in the BDS, and presenting a body with the face will likely increase the ecological validity of the scale.

Although using computer generated stimuli can never be completely ecologically valid, it can reduce some of the problems found when using photographic images and is especially useful in an experimental setting where greater control is needed. Indeed, the use of computer generated stimuli is being used increasingly in body perception research (e.g., Crossley, Cornelissen, & Tovée, 2012; Tovée, Edmonds, & Vuong, 2012). However, there is currently no pictorial scale using computer generated stimuli that has been psychometrically evaluated for measuring body dissatisfaction. This study is the first to develop and test a scale using this new type of computer generated stimuli.

The aim of our research was to examine the validity and test-retest reliability of both the female and male versions of the BDS. Similar to previous research (e.g., Swami et al., 2008a; Thompson & Gray, 1995), the scale was initially tested for validity by assessing the extent to which participants could successfully distinguish the size differences between the nine bodies. It was expected that, for the scale to be valid, participants would be able to do this successfully. In line with previous research (Swami et al., 2008a), construct validity was assessed by examining the correlation between the BDS and a measure of positive body image. It was hypothesized that these variables would be negatively correlated. In addition, the correlation between participants' perceived actual body size ratings and their Body Mass Index (BMI) was also used to asses construct validity. It was predicted that there would be a positive correlation between participants' BMI and their perceived actual body. Finally, test-retest reliability was examined by the correlation of body dissatisfaction scores from a first and second testing, and this was expected to be positively correlated to show reliability.

## Method

### Participants

Participants were 130 females and 60 males. They were all Psychology students from Nottingham Trent University. Participants were given research credits in exchange for taking part in the research.

Of the female participants, 76.9% were White European, 10.8% were of Asian descent, 7.7% had mixed ethnicity, and 4.6% were of African Caribbean descent. Of the male participants, 83.3% were White European, 8.3% were of Asian descent, 5.1% were of African Caribbean descent, and 3.3% were of mixed ethnicity.

### Materials

***Body Dissatisfaction Scale (BDS).*** The scale was developed by creating an average sized body using DAZ Studio 4 software (www.daz3d.com). Using this software the average sized body was systematically altered to create 4 bodies which were successively thinner than the original body and 4 bodies which were larger than the original body. Thus, nine bodies were created in total which range from extremely thin to obese. This was done for both the female (see Appendix A1) and male bodies (see Appendix A2), creating two versions of the scale.

Following Crossley et al. (2012), the bodies, without clothing, were then exported out of DAZ studio into 3ds Max (www. autodesk.com). 3ds Max is able to estimate the height of the body in real world measurements (cms). In addition, 3ds Max calculates the volume of the body. Once volume is known the weight of the body can be estimated by multiplying the density of either the average young female body (1.04 $g/cm^3$) or the average young male body (1.06 $g/cm^3$). This enables the BMI of each body to be calculated by dividing the weight (kg) by the height (m) squared (see Appendix B for measurements, weights, and BMIs). The BMIs of the nine bodies in both the female and male scale range from underweight to obese: bodies 1-3 are underweight, bodies 4-6 are in the normal range and bodies 7-9 are overweight/obese.

All the bodies are depicted at a 25 degree angle to enable more visual information about the 3D shape of the body to be made available. This gives a more realistic portrayal of the human body shape that would not be achieved by using simple front view bodies. Gardner et al. (2009) recommended omitting facial and body features in pictorial stimuli that reflect obvious Caucasian ethnicity. However, since removing or obscuring the facial features would reduce the ecological validity of the images, the bodies are instead presented in grey scale and without hair to minimize the effects of race or ethnicity. It is therefore thought that when presenting the bodies in this way there would be little effects of perceived ethnicity. All bodies are depicted wearing black shorts and a T-shirt, so the size and shape of the body is clearly visible.

The BDS scale can be used to measure body dissatisfaction by numbering the bodies from 1 to 9 (in ascending order of size); each body is scored as one body unit. Participants are asked to choose the body they would most like to look like (ideal) and the body they thought was closest to their perceived actual body shape (actual). The discrepancy between the participant's selected actual and ideal body was the participant's body dissatisfaction score. For example, if a participant chose body number 5 as their actual body and body number 2 as their ideal body, their body dissatisfaction score would be 3. A higher score means a greater discrepancy between perceived ideal and actual body chosen, meaning greater dissatisfaction. This difference score does not reflect the direction of the body dissatisfaction and participants may have chosen an ideal body that was thinner or heavier than their chosen actual body. The highest body dissatisfaction score a participant could receive would be 8 if they selected body 1 and body 9. If a participant has no body dissatisfaction, i.e., they selected the same number body for both their perceived ideal and actual body, they would have a score of zero.

***Body Appreciation Scale (BAS; Avalos, Tylka, & Wood-Barcalow, 2005).*** The BAS comprises 13 items which are intended to measure body appreciation. Items are statements for which participants are asked to select a response. Some example items from the BAS include "I feel that my body has at least some good qualities" and "My self-worth is independent of my body shape or weight." Items are scored on a 5 point scale (1 = *never*, 2 = *seldom*, 3 = *sometimes*, 4 = *often*, 5 = *always*) and are averaged to obtain an overall body appreciation score. A higher score indicates higher body appreciation. The BAS has been shown to have good internal consistency ($\alpha = .94$). Construct validity was demonstrated by a positive correlation with a tendency to evaluate one's appearance favorably ($r = .68$), and negative correlations with body preoccupation ($r = -.79$) and eating disorder symptomatology ($r = -.60$). In addition, test-retest reliability was found to be good ($r = .90$) over a 3 week period.

***Initial validation task.*** Following Swami et al.'s (2008a) and Thompson and Gray's (1995) validation of the PFRS and the CDFRS, the BDS was tested for validity by assessing the extent to which participants could successfully identify the body weight differences between the nine bodies. In Swami et al. (2008a) and Thompson and Gray (1995), participants were asked to order the images from thinnest to heaviest and the percentage of correctly positioned bodies was used to establish validity. In the current research, participants were given a two alternative forced choice (2AFC) task in which participants were presented with pairs of bodies. Participants had to decide which body they thought was the thinnest body out of the pair. All possible pair combinations were presented in a random order. The percentage of combinations where the participants correctly identified the thinnest body was used to assess the ability of participants to distinguish the body weight differences between all bodies and therefore validity of the BDS.

In order for the scale to be valid it is essential that participants are able to see that all the bodies vary in body weight. It is not sufficient to tell participants that bodies are presented from underweight to obese before the scale is used. In addition, this means that in further research the bodies do not need to always have to be presented from underweight to obese and could be presented in a variety of ways if it is established that participants can distinguish weight differences between all the bodies.

***Test-retest reliability.*** Test-retest reliability was examined by the correlation of body dissatisfaction scores from the initial testing and then again after five weeks. There is no evidence to establish that one particular time period is best for reliability testing in pictorial body dissatisfaction measurement scales. The length of time does affect the reliability of a measure to the extent that the first testing may influence the second testing. Previous studies have used one week (Thompson & Gray 1995) and three weeks (Swami et al., 2008a). The current research chose five weeks as it seemed a suitable amount of time for the first testing to no longer have an effect on the second testing.

***Demographics.*** Participants were also asked their age, sex, ethnic origin, and self-reported height and weight. Height and weight were used to calculate each participant's BMI.

## Procedure

Participants were presented with the bodies as a scale on paper, with images going from left to right, from underweight to obese so as to measure their perceived body dissatisfaction. Bodies were numbered from 1 to 9 (in ascending order of size) and each body is scored as one body unit. Participants were asked to choose the body they would most like to look like (ideal) and the body they thought was closest to their perceived actual body shape (actual). This was a categorical judgment and participants could only choose one body each time. Female participants were presented with the female version of the scale and male participants were administered the male version.

Participants were then given the 2AFC task to complete. Participants completed this task in SuperLab 4.5 (www.superlab.com) on a 44.3 x 25.4 cm screen. Participants had to decide which body they thought was the thinnest body out of each pair.

Five weeks after the initial test, participants were invited to use the new rating scale to measure their body image again. Of the original sample, 64 female participants and 20 male participants returned to complete the scale for a second time.

## Results

### Female Version

***Descriptive statistics.*** The descriptive statistics for the female participants are presented in Table 1.

***Initial validation.*** Participants were assessed on their performance on the 2AFC validation task. Same body pairings were removed from the analysis, meaning 72 pair combinations for each participant were analyzed. Due to a computer error, two of the pair combinations were displayed incorrectly for 30 of the participants and therefore the responses for these combinations were removed from the analysis.

The results showed that on average participants were able to correctly identify the thinnest body 97.63% ($SD = 2.44$) of the time. A chi-square analysis showed that amount of times participants correctly identified the thinnest body was significantly higher than chance $\chi^2(1, N = 130) = 8440.82$, $p < .001$. This validation task was conducted as to attain whether participants were able identify the thinnest body out of each pair and therefore detect body weight differences between the bodies. The high percentage correct found here suggests that participants are accurately able to detect the body weight differences between the nine bodies.

***Construct validity.*** Validity was assessed by the correlation between body dissatisfaction scores on the BDS and body appreciation scores on the BAS. Results showed a significant negative correlation,

Table 1
*Female Participants Mean Scores on all Measures*

| | M | SD |
|---|---|---|
| Perceived actual body score | 5.58 | 1.31 |
| Perceived ideal body score | 4.08 | 1.0 |
| Body dissatisfaction score | 1.61 | .90 |
| BAS score | 3.44 | .62 |
| BMI | 23.1 | 3.50 |

Table 2
*Male Participants Mean Scores on all Measures*

| | M | SD |
|---|---|---|
| Perceived actual body score | 5.07 | 1.54 |
| Perceived ideal body score | 4.90 | .82 |
| Body dissatisfaction score | 1.49 | .82 |
| BAS score | 3.54 | .64 |
| BMI | 24.12 | 3.83 |

$r(128) = -.60$, $p < .001$, providing evidence of construct validity. To provide further validation, the correlation between participants' ratings of their perceived actual body size and their BMI was assessed. This was also found to be significant, $r(128) = .77$, $p < .001$, suggesting that the scale can be used to accurately assess perception of one's own body size.

***Test-retest reliability.*** The correlations between scores for perceived actual body, ideal body and body dissatisfaction from the first testing and five weeks after were analyzed for test-retest reliability. These were all found to be significant: perceived actual body, $r(62) = 0.81$, $p < .001$; ideal body, $r(62) = 0.89$, $p < .001$; and body dissatisfaction, $r(62) = 0.82$, $p < .001$. Scores obtained at the first and second testing were highly correlated, suggesting that the scale is a reliable measure of body dissatisfaction.

### Male Version

***Descriptive statistics.*** The descriptive statistics for the male participant sample are presented in Table 2.

***Initial validation.*** Participants were assessed on their performance in the 2AFC validation task. In total, 72 pair combinations were analyzed for each participant to identify if participants had correctly selected the thinnest body out of each pair. Due to a computer error, one of the pair combinations was displayed incorrectly for seven of the participants and therefore the responses for these combinations were removed from the analysis.

Results showed that participants were able to correctly select the thinnest body 98.04% ($SD = 1.82$) of the time when the bodies were presented in pairs. A chi-square analysis showed that amount of times participants correctly identified the thinnest body was significantly higher than chance $\chi^2(1, N = 60) = 3980.70, p < .001$. Being able to correctly identify the thinnest body out of each pair of bodies suggests that body weight differences between the nine male bodies are easily identifiable.

***Construct validity.*** A significant negative correlation was found between body dissatisfaction on the BDS and body appreciation scores, $r(57) = -.46, p < .001$, providing evidence of construct validity. The correlation between participants' ratings of their perceived actual body size and their BMI was also found to be significant, $r(57) = 0.83, p < .001$, providing further validation.

***Test-retest reliability.*** To examine test-retest reliability, original scores for perceived actual body, ideal body, and body dissatisfaction were correlated with scores from five weeks after the initial testing. All three correlations were found to be significant: perceived actual body, $r(18) = 0.96$ $p < .001$; ideal body, $r(18) = 0.88$ $p < .001$; and body dissatisfaction, $r(18) = 0.97$ $p < .001$.

## Discussion

The results of the present research suggest that both the male and female version of the BDS exhibit good construct validity and test-retest reliability over a five-week period. It would appear that participants can easily detect the subtle differences in size between the nine bodies on the scale. The significant correlations with the BAS suggest that both the female and male version of the BDS have good construct validity. Perceived actual body size was highly correlated with participants BMI for both male and female versions, which indicates that the BDS is a useful tool in assessing perception of one's own body size and provides further evidence of construct validity.

The findings support the use of the BDS in body image measurement for females and males. The current scale offers improvement on scales which have used line-drawn stimuli (e.g., Stunkard et al., 1983; Thompson & Gray, 2005) by offering greater realism and providing more life-like figures. By using computer generated figures, it also avoids the problems associated with using images of real people (e.g., PFRS), such as biases in judgements associated with certain racial groups. It also controls for unwanted variation in body part ratios, inherent in real human beings both across and within the sexes.

Another advantage of using the BDS is that it is time efficient and easy to administer either in a digital or paper format. A future version of the scale where the images are rotated and presented in three dimensional formats could be developed using the same software. The use of more life-like computer generated stimuli, like the bodies used in the BDS, is also particularly useful as it allows for the presentation of stimuli in more realistic settings, which could not be achieved with line-drawn stimuli. Therefore, the current stimuli is not limited to the use of images simply being presented in isolation. Future research could see the bodies, for example, being presented against various computer generated backgrounds or settings to allow for a more realistic presentation.

Although the current research provides a male version of the scale that is directly comparable with the female version, it has been questioned if bodyweight is a strong predictor of physical attractiveness in males. Some research has suggested that upper body muscularity is a more important indicator of attractiveness (Maisey, Vale, Cornelissen, & Tovée, 1999; Swami & Tovée, 2005; Swami & Tovée, 2008). Therefore, a male version of the scale which varies in the muscularity of the body shape may be useful. Notwithstanding muscularity, BMI does appear to be a significant predictor of male physical attractiveness (Maisey et al., 1999). Males are concerned with their body weight (Pope, Phillips, & Olivardia, 2000), with research suggesting around half wanting to lose weight and half wanting to gain muscle (Drewnowski & Yee, 1987). Therefore, it would seem that body weight is still an important component of perceived body dissatisfaction in males.

In addition, a male version of the scale which varied in muscularity would not be directly comparable with the female version. The stimuli in the BDS, like the PFRS, can also be used in an experimental setting to measure attractiveness and health preferences for body weight (e.g., Swami & Tovée, 2006). In this way the stimuli can be used to ask participants which body they find most attractive and healthy. In this type of experiment, it would be necessary to have both a male and female version of the scale that both varied in BMI to ensure the results were directly comparable with each other. Although males and females may not be equally concerned about their body weight, when measuring body dissatisfaction in an experimental setting it is extremely useful to have a measure that is identical for males and females. Therefore, although a version of the male scale that varied in muscularity would be useful, the current version of the male scale is necessary.

One limitation of the current research is that the sample was larger for female participants than for male participants, potentially suggesting that the results are more reliable for the female sample. Therefore, findings of the male version should be treated with more caution than the female version. However, the male version of the scale has good face validity and appears to be measuring body dissatisfaction. A-priori power analysis is useful in achieving a reliable sample size on which to judge statistical effects. However, selecting an appropriate sample size and calculating power is complex (see Baguley, 2012; Hoenig & Heisey, 2001). Post-hoc power analyses are particularly problematic because such analyses tend to involve transforming the *p*-values of the effects being explored. This is problematic because it is paradoxical, as it would involve using significant or non-significant *p*-values to confirm that a given effect is (accordingly) significant or non-significant (see Hoenig & Heisey, 2001; Thomas, 1997). Therefore, it was decided not to run and report post-hoc power analyses. The apparent strength of the correlations and the near ceiling performance on the repeated measures 2AFC body detection task are used instead as indicators of sufficient power for both the female and male version. Consequently, this does not mean that the male version should not be used, especially since, as mentioned, there is currently a lack of scales that have both female and male versions.

A further limitation is the use of an opportunity sample, meaning the participants were all students, which is not representative of the population as a whole. As this scale was primarily being developed to be used in a set of experiments with students, it is certainly valid to be used in the context for which it is intended. Further research could aim to validate the scale with a more diverse sample. In addition, it could be argued that using self-reported weight and height to calculate BMI is not as accurate as using the actual weight and height of participants. However, previous research has suggested that self-report measures of weight and height are highly correlated with participants' actual weight and height (Spencer, Appleby, Davey, & Key, 2001), and it is time consuming to collect the actual weight and height of participants. Therefore, using self-report measures is a more time efficient method. Finally, although the stimuli used are more life-like in appearance than previous line drawn stimuli (e.g., Thompson & Gray, 1995), computer generated stimuli may never be as ecologically valid as using real life images. Nonetheless, given the problems with maintaining control and consistency over photographic stimuli, computer generated bodies provide a useful alternative, particularly for researchers wanting to measure body dissatisfaction in an experimental setting.

Future research should aim to develop and validate a version of the BDS so that it can be successfully used with children. Body image measurement with young children can be particularly difficult. As a consequence, researchers often use pictorial scales as other measures can be too complex, placing increased cognitive demands on younger children. However, pictorial scales in this area are limited to line-drawings (e.g., Collins, 1991) which have similar methodological problems to line drawn stimuli used with adults. A version of the BDS that could be used with children would be useful for the assessment of body dissatisfaction in young children. Future research should also aim to further validate the scale and demonstrate its reliability.

In conclusion, the current research suggests that both the male and female version of the BDS show good construct validity, and test-retest reliability is stable over a five-week period. The current scale has improved realism on previous line drawn scales while avoiding the consistency issues associated with the use of photographic stimuli. In addition, it is convenient and easy to administer, consequently making it a useful tool in the measurement of body dissatisfaction.

### References

Avalos, L., Tylka, T. L., & Wood-Barcalow, N. (2005). The body appreciation scale: Development and psychometric evaluation. *Body Image, 2*(3), 285-297. doi:10.1016/j.bodyim.2005.06.002

Baguley, T. (2012). *Serious statistics: A guide to advanced statistics for the behavioural sciences*. Basingstoke, UK: Palgrave MacMillan.

Bateson, M., Cornelissen, P. L., & Tovée, M. J. (2007). Methodological issues in studies of female attractiveness. In V. Swami & A. Furnham (Eds.), *Body beautiful: Evolutionary and sociocultural perspective* (pp. 46-63). Basingstoke, UK: Palgrave MacMillan.

Cattarin, J. A., & Thompson, J. K. (1994). A three-year longitudinal study of body image, eating disturbance, and general psychological functioning in adolescent females. *Eating Disorders: The Journal of Treatment & Prevention, 2*(2), 114-125. doi:10.1080/10640269408249107

Collins, M. E. (1991). Body figure perceptions and preferences among preadolescent children. *International Journal of Eating Disorders, 10*(2), 199-208. doi:10.1002/1098-108X(199103)10:2<199::AID-EAT2260100209> 3.0.CO;2-D

Crossley, K. L., Cornelissen, P. L., & Tovée, M. J. (2012). What is an attractive body? Using an interactive 3D program to create the ideal body for you and your partner. *PLoS ONE, 7*(11), e50601. doi:10.1371/journal.pone.0050601

Drewnowski, A., & Yee, D. K. (1987). Men and body image: Are males satisfied with their body weight? *Psychosomatic Medicine, 49*(6), 626-634. doi:10.1097/00006842-198711000-00008

Gardner, R. M., Friedman, B. N., & Jackson, N. A. (1998). Methodological concerns when using silhouettes to measure body image. *Perceptual and Motor Skills, 86*(2), 387-395. doi:10.2466/pms.1998.86.2.387

Gardner, R. M., Jappe, L. M., & Gardner, L. (2009). Development and validation of a new figural drawing scale for body-image assessment: The BIAS-BD. *Journal of Clinical Psychology, 65*(1), 114-122. doi:10.1002/jclp.20526

Garner, D. (2002). Body image and anorexia nervosa. In T. F. Cash & T. Pruzinsky (Eds.), *Body image: A handbook of theory, research, and clinical practice* (pp. 295-303). New York, NY: Guilford Press.

Garner, D. M., Olmsted, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders, 2*, 15-34. doi: 10.1002/1098-108x(198321)2:2<15::aid-eat2260020203>3.0.co;2-6

Grogan, S. (2008). Culture and body image. In S. Grogan (Ed.), *Body image: Understanding body dissatisfaction in men, women and children* (pp. 9-40). Hove, UK: Routledge.

Hoenig, J. M., & Heisey, D, M. (2001). The Abuse of Power. *The American Statistician, 55*, 19-24. doi:10.1198/000313001300339897

Kinsey, A. C., Pomeroy, W. B., & Martin, C. E. (1998). *Sexual behavior in the human male*. Bloomington, IN: Indiana University Press.

Maisey, D. S., Vale, E., Cornelissen, P. L., & Tovée, M. J. (1999). Characteristics of male attractiveness for women. *Lancet, 353*(9163), 1500-1500. doi: 10.1016/S0140-6736(99)00438-9

Pope, H. G., Phillips, K. A., & Olivardia, R. (2000). *The adonis complex: The secret crisis of male body obsession*. New York, NY: Free Press.

Singh, D. (1993). Adaptive significance of female physical attractiveness: Role of waist-to-hip ratio. *Journal of Personality and Social Psychology, 65*(2), 293-307. doi:10.1037/0022-3514.65.2.293

Smink, F., van Hoeken, D., & Hoek, H. W. (2012). Epidemiology of eating disorders: Incidence, prevalence and mortality rates. *Current Psychiatry Reports, 14*(4), 406-414. doi:10.1007/s11920-012-0282-y

Spencer, E. A., Appleby, P. N., Davey, G. K., & Key, T. J. (2001). Validity of self-reported height and weight in 4808 EPIC-oxford participants. *Public Health Nutrition, 5*, 561-565. doi:10.1079/PHN2001322

Stunkard, A. J., Sorensen, T., & Schulsinger, F.

(1983). Use of the Danish adoption register for the study of obesity and thinness. In S. S. Kety, L. P. Rowland, R. L. Sidman, & S. W. Matthysse (Eds.), *The genetics of neurological and psychiatric disorder* (pp. 115-120). New York, NY: Raven.

Swami, V., Einon, D., & Furnham, A. (2006). The leg-to-body ratio as a human aesthetic criterion. *Body Image, 3*(4), 317-323. doi:10.1016/j.bodyim.2006.08.003

Swami, V., Furnham, A., & Joshi, K. (2008). The influence of skin tone, hair length, and hair colour on ratings of women's physical attractiveness, health and fertility. *Scandinavian Journal of Psychology, 49*(5), 429-437. doi:10.1111/j.1467-9450.2008.00651.x

Swami, V., Salem, N., Furnham, A., & Tovée, M. J. (2008a). Initial examination of the validity and reliability of the female photographic figure rating scale for body image assessment. *Personality and Individual Differences, 44*(8), 1752-1761. doi:10.1016/j.paid.2008.02.002

Swami, V., Salem, N., Furnham, A., & Tovée, M. J. (2008b). The influence of feminist ascription on judgements of women's physical attractiveness. *Body Image, 5*(2), 224-229. doi:10.1016/j.bodyim.2007.10.003

Swami, V., & Tovée, M. J. (2005). Male physical attractiveness in Britain and Malaysia: A cross-cultural study. *Body Image, 2*(4), 383-393. doi:10.1016/j.bodyim.2005.08.001

Swami, V., & Tovée, M. J. (2006). Does hunger influence judgments of female physical attractiveness? *British Journal of Psychology, 97*(3), 353-363. doi:10.1348/000712605X80713

Swami, V., & Tovée, M. J. (2008). The muscular male: A comparison of the physical attractiveness preferences of gay and heterosexual men. *International Journal of Men's Health, 7*(1), 59-71. doi:10.3149/jmh.0701.59

Tassinary, L. G., & Hansen, K. A. (1998). A critical test of the waist-to-hip-ratio hypothesis of female physical attractiveness. *Psychological Science, 9*(2), 150-155. doi:10.1111/1467-9280.00029

Thomas, L. (1997). Retrospective power analysis. *Conservation Biology, 11*, 276–280. doi:10.1046/j.1523-1739.1997.96102.x

Thompson, J. K., & Altabe, M. N. (1991). Psychometric qualities of the figure rating scale. *International Journal of Eating Disorders, 10*(5), 615-619. doi:10.1002/1098-108X(199109)10:5<615::AID-EAT2260100514>3.0.CO;2-K

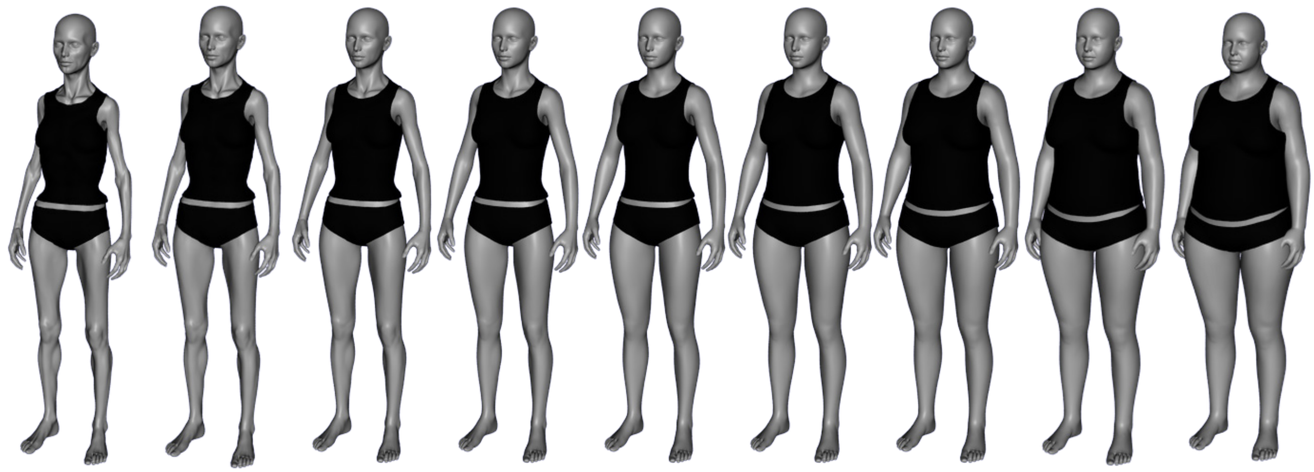
Thompson, M. A., & Gray, J. J. (1995). Development and validation of a new body-image assessment scale. *Journal of Personality Assessment, 64*(2), 258-269. doi:10.1207/s15327752jpa6402_6

Tiggemann, M. (2011). Sociocultural perspectives of human appearance and body image. In T. F. Cash & L. Smolak (Eds.), *Body image: A handbook of science, practice and prevention* (2nd ed., pp. 12-19). New York, NY: Guilford Press.

Tovée, M. J., Edmonds, L., & Vuong, Q. C. (2012). Categorical perception of human female physical attractiveness and health. *Evolution and Human Behavior, 33*(2), 85-93. doi:10.1016/j.evolhumbehav.2011.05.008

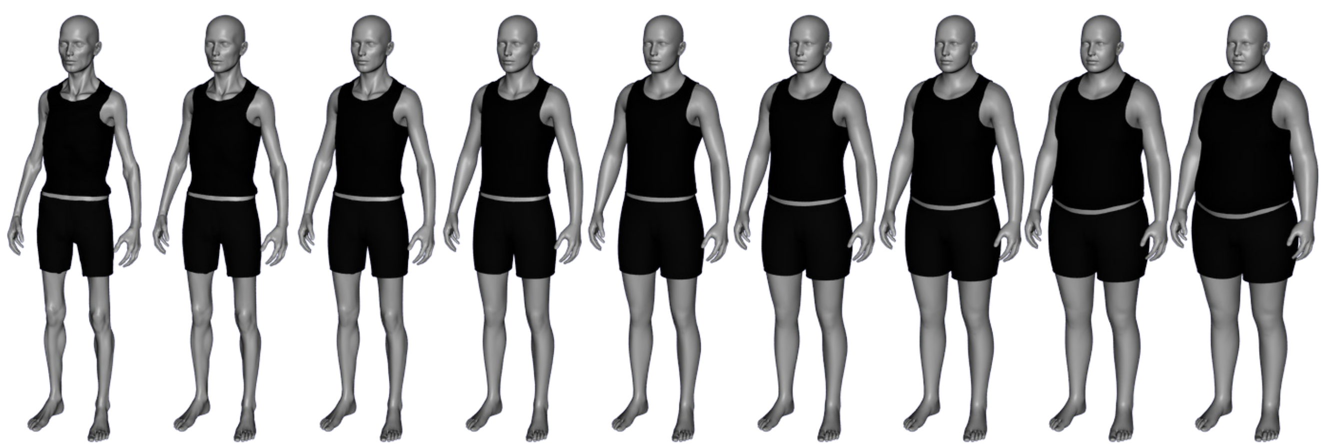
**Appendix A**

1. Female body stimuli



**Body 1** **Body 2** **Body 3** **Body 4** **Body 5** **Body 6** **Body 7** **Body 8** **Body 9**

2. Male body stimuli



**Body 1** **Body 2** **Body 3** **Body 4** **Body 5** **Body 6** **Body 7** **Body 8** **Body 9**

## Appendix B

Female body measurements

| | Height (cm) | Volume | Density ($g/cm^2$) | Weight (g) | Weight (kg) | Height (meters) | Body Mass Index |
|---|---|---|---|---|---|---|---|
| 1 | 172.62 | 39433.94 | 1.04 | 41011.30 | 41.01 | 1.73 | 13.76 |
| 2 | 172.62 | 43104.10 | 1.04 | 44828.26 | 44.83 | 1.73 | 15.04 |
| 3 | 172.62 | 47421.31 | 1.04 | 49318.16 | 49.32 | 1.73 | 16.55 |
| 4 | 172.61 | 52287.90 | 1.04 | 54379.42 | 54.38 | 1.73 | 18.25 |
| 5 | 172.61 | 57364.37 | 1.04 | 59658.94 | 59.66 | 1.73 | 20.02 |
| 6 | 172.62 | 66488.58 | 1.04 | 69148.12 | 69.15 | 1.73 | 23.21 |
| 7 | 172.63 | 76290.74 | 1.04 | 79342.37 | 79.34 | 1.73 | 26.62 |
| 8 | 172.64 | 86402.47 | 1.04 | 89858.57 | 89.96 | 1.73 | 30.15 |
| 9 | 172.65 | 98237.68 | 1.04 | 102167.2 | 102.17 | 1.73 | 34.27 |

Male body measurements

| | Height (cm) | Volume | Density ($g/cm^2$) | Weight (g) | Weight (kg) | Height (meters) | Body Mass Index |
|---|---|---|---|---|---|---|---|
| 1 | 172.62 | 40862.48 | 1.06 | 43314.23 | 43.31 | 1.73 | 14.54 |
| 2 | 172.62 | 44786.76 | 1.06 | 47473.97 | 47.47 | 1.73 | 15.93 |
| 3 | 172.62 | 49155.39 | 1.06 | 52104.71 | 52.10 | 1.73 | 17.49 |
| 4 | 172.61 | 53826.42 | 1.06 | 57056.01 | 57.06 | 1.73 | 19.15 |
| 5 | 172.61 | 58956.77 | 1.06 | 62494.18 | 62.49 | 1.73 | 20.97 |
| 6 | 172.62 | 68373.29 | 1.06 | 72475.69 | 72.48 | 1.73 | 24.32 |
| 7 | 172.62 | 78023.91 | 1.06 | 82705.34 | 82.71 | 1.73 | 27.75 |
| 8 | 172.64 | 88868.72 | 1.06 | 94200.84 | 94.2 | 1.73 | 31.61 |
| 9 | 172.65 | 101008.70 | 1.06 | 107069.20 | 107.07 | 1.73 | 35.92 |