View-dependent accuracy in body mass judgements of female bodies

Piers L. Cornelissen[a], Katri K. Cornelissen[a], Victoria Groves[a], Kristofor McCarty[a], Martin J. Tovée[b],*

[a]Department of Psychology, Faculty of Health & Life Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK

[b]School of Psychology, College of Social Science, University of Lincoln, Lincolnshire, LN6 7TS, UK

***Corresponding author. E-mail address: MTovee@lincoln.ac.uk (M.J. Tovée).**

Abstract

A fundamental issue in testing body image perception is how to present the test stimuli. Previous studies have almost exclusively used images of bodies viewed in front-view, but this potentially obscures key visual cues used to judge adiposity reducing the ability to make accurate judgements. A potential solution is to use a three-quarter view, which combines visual cues to body fat that can be observed in front and profile. To test this hypothesis, 20 female observers completed a 2-alternative forced choice paradigm to determine the smallest difference in body fat detectable in female bodies in front, three-quarter, and profile view. There was a significant advantage for three-quarter and profile relative to front-view. Discrimination accuracy is predicted by the saliency of stomach depth, suggesting that this is a key visual cue used to judge body mass. In future, bodies should ideally be presented in three-quarter to accurately assess body size discrimination.

Key words: BMI, body fat, body judgements, figural body scales.

Introduction

There has been a steady rise in obesity levels in the developed world with a concomitant pressure on public health resources (Ogden, Carroll, Kit, & Flegal, 2014; Swinburn et al., 2011). In tandem with this rise, there has also been an increase in the levels of negative body image, which may have contributed to the increasing prevalence of eating disorders and conditions such as muscle dysmorphia (Cash & Pruzinsky, 2002; Grabe, Ward, & Hyde, 2008; Pope, Phillips, & Olivardia, 2000; Swami et al., 2010). From both an epidemiological and clinical point of view, it is therefore important to develop psychometrically sound measurement scales for the self-assessment of body size/shape (Gardner & Brown, 2010; Thompson & Gray, 1995). Many different such measures have been constructed, but amongst the most commonly used include: (a) figural body scales that are composed of a series of images of either men or women varying in adiposity from emaciated to obese (Stunkard, Sorensen, & Schulsinger, 1983), (b) computerized tasks which either present many examples of such images in random order, one at a time, or which allow the stimulus to be smoothly animated between minimum and maximum body size endpoints (Gardner & Brown, 2010). Depending on the task, participants either estimate their own body size by choosing images closest to the size/shape they believe themselves to have or would like to have. Alternatively, participants make decisions about whether any particular stimulus is smaller/larger than the body size they believe themselves to have or would like to have (the difference between the two is a measure of body dissatisfaction) (Brodie, Bagley, & Slade, 1994; Gardner & Brown, 2011). In this paper we assert that judgements of this kind should properly be thought of as magnitude estimation tasks and should therefore follow Weber's law (1834). We then ask whether any of the three commonly used orientations for whole body stimuli (side, front, and three-quarter view) produce participant responses that conform to this expectation. Failure to do so may lead to systematic patterns of over- and/or under-estimation when people judge their body size.

Weber's Law

In whatever perceptual domain, be it sensory or proprioceptive, human magnitude estimation has been shown to follow Weber's law almost without exception. This is the phenomenon whereby the smallest difference between a pair of stimuli that can be reliably told apart (the just noticeable difference or JND) is a constant proportion of the stimulus magnitude. To illustrate, as a reference weight gets bigger, then a test weight which is to be compared to it needs to be heavier, by a constant proportion of the reference, in order that the test is correctly identified as being heavier than the reference (i.e., the Weber fraction $K = \Delta I / I$, where I = reference stimulus magnitude and K = constant). Weber's law only holds for physical properties that have magnitude. This is the mathematical property which determines whether an object is larger or smaller than other objects of the same kind, and is represented numerically by values that start at zero and must thereafter be positive. While rare exceptions do exist, for example for pure tone and noise intensity discrimination at high intensities in the auditory domain (Jesteadt, Wier, & Green, 1977), Weber's law should nevertheless be considered ubiquitous for human magnitude perception.

In the case of body mass index (BMI), we should expect that a plot of the JND for BMI (y-axis) as a function of reference BMI (x-axis) should be a straight line with a positive slope, and the Weber fraction, *K,* should be constant across the reference BMI range. In principle therefore, a useful way to design a figural scale for body size estimation would be based on JNDs for BMI. Starting from the smallest body size that one might want participants to judge, the next largest figure on the scale might be 2 JNDs larger, the next 2 JNDs larger still, and so on to the end point for the scale. Indeed, the Dol Pain scale was designed exactly in this way (Adair, Stevens, & Marks, 1968) and is still in use today.

A useful way to think about JNDs is in terms of the precision of magnitude judgements. Precision is said to be high when the JND is small. Precision is related to the statistical concept of variability (standard deviation, quartile deviation, or range), and to the concept of reliability or random error ("noise"). Since according to Weber's law, JND increases linearly with reference stimulus magnitude, this means that the precision with which judgements can be made falls correspondingly – hence leading to the need for bigger differences between stimulus pairs with increasing reference magnitude. However, a second implication is that the ideal stimuli for a figural scale should also give rise to the smallest possible JNDs at each reference magnitude. Given the example above of a straight-line plot of JND for BMI as a function of reference BMI, then the ideal figural scale would not only have a constant Weber fraction, *K*, but also an intercept for the relationship which is as close to zero as possible. This would lead to more precise body size estimates, lower variability across participants, and improved psychometric properties of the task. In the case of identifying individuals at risk from obesity in epidemiological samples, reducing the JNDs for the figural scales (e.g., as reported by Dratva et al., 2016) would lead to improved sensitivity and specificity.

Test validity

An important attribute of any psychometric test is that of content validity: "… if the items of a test can be shown to reflect all aspects of the subject being tested, then it is per se valid, given that the instructions are clear. This is not simply face validity, which is related to the appearance of the test items …" (Kline, 2015). With figural body scales and their computerized equivalents, an important consideration regarding content validity is the orientation of the body in the scale. The reason this is important is because, even though perceptual estimates of BMI should follow Weber's law, because BMI has magnitude, if the stimuli representing changes in BMI lack content validity, then we may nevertheless fail to observe Weber's law behaviour. Bodies in published figural scales have almost exclusively been presented in front-view (Gardner, Jappe, & Gardner, 2009; Harris, Bradlyn, Coffman, Gunel, & Cottrell, 2008; Li, Hu, Ma, Wu, & Ma, 2005; Peterson, Ellenberg, & Crossan, 2003; Swami, Salem, Furnham, & Tovée, 2008). However, to our knowledge, there have been no systematic studies to confirm whether the front view is indeed optimal – and here we would define optimal as producing participant responses which follow Weber's law. Indeed, there are reasons for believing that the front view may obscure visual cues normally used by an observer to judge body mass, thereby reducing content validity. For example, stomach depth, which has been suggested to be an important cue to body mass judgements (Cornelissen, Hancock, Kiviniemi, George, & Tovée, 2009; Rilling, Kaufman, Smith, Patel, & Worthman, 2009; Smith, Cornelissen, & Tovée, 2007; Tovée, Maisey, Emery, & Cornelissen, 1999) may be harder to judge in front-view than in profile. The use of front-view may also make it difficult to accurately estimate body fat in populations of African descent where the pattern of fat deposition differs from European populations with more fat deposited on the thighs and buttocks which are not visible in front-view (Cohen et al., 2015a; Cohen et al., 2015b; Marlowe, Apicella, & Reed, 2005).

The current study

Here we sought to determine which of three stimulus orientations: frontal, three-quarter or side view, is most suitable for use in body size estimation tasks. So, it is an investigation of basic stimulus properties. To do this, we used a 2-alternative forced choice (2-AFC) paradigm to determine the smallest difference in body fat that could be detected at the three different orientations (i.e., the JND for BMI). Our criteria for suitability were: (a) that participant responses obeyed Weber's law empirically because that is what we should expect them to do theoretically, (b) that participant responses maximize precision by minimizing JNDs across the reference range. We emphasize that the current study is an investigation of participants' basic ability to discriminate differences in body size between pairs of images. This is a judgement about others, made from a third-person point of view, which does not require participants to refer to their own body image in any way. Therefore, we should not expect these psychophysical estimates to be influenced by participants' body satisfaction or their attitudes to body shape, weight or eating, or indeed their own BMI.

Methods

Participants

We used a repeated measures design with two within-participants factors: CGI model orientation (3 levels: three-quarter, front, and side views) and reference BMI (4 levels: 15, 20, 27, & 36). We recruited 5 female participants to pilot this experiment. None of the participants who took part in this pilot study also took part in the main study. To estimate the sample size required for the main study from the pilot data, we used GLIMMPSE (General Linear Multivariate Model Power & Sample Size; Kreidler et al., 2013). We calculated conservative multivariate tests (by scaling the calculated covariance matrix by a factor of 2) of the interaction between main effects. This showed that a sample of 12 participants would be sufficient to quantify the main effects and interactions when modelling JND as a function of stimulus BMI and stimulus orientation, at a nominated alpha level of .01 and a power of .90. To offset attrition in participant numbers and/or unexpected sources of variability, we recruited 20 female participants (age *M* = 25.40 years, *SD* = 8.40) for this study from staff and students at Northumbria University in the UK. The participants had a mean BMI of 22.7 and a *SD* of 4.0. The BMI values of the participants range from 15.40 to 31.20 (3 are underweight, 11 are in the normal range, 5 are overweight and 1 is obese). We asked all potential participants whether they had a current diagnosis or history of an eating disorder and excluded those individuals from this study.

Stimuli

We wanted to identify the smallest change in BMI that observers could detect (the JND), at four separate points along the BMI continuum, corresponding to the World Health Organization's classification for underweight, normal, overweight, and obese. Accordingly, we chose reference BMIs for each of these four groups: 15, 20, 27, & 36 respectively. To create stimulus images which correctly represent how an individual body shape changes as a function of changing BMI, we used computer-generated imagery (CGI) methods to create graded 3D images of a standard model where: (a) the identity of the person in the image is clearly maintained over a wide BMI range and across the three body orientations (i.e., three-quarter view, front view, and side view); (b) the body shape changes at different BMI levels are extremely realistic and (c) the 3D rendered stimulus images are high definition and photorealistic (for further technical details see Supplementary Materials linked online to this article and Cornelissen, Bester, Cairns, Tovée, & Cornelissen, 2015; Cornelissen, Gledhill, Cornelissen, & Tovée, 2016). In addition, we made precise estimates of the BMI of the 3D model in our stimulus images. To achieve this, we used the Health Survey for England (2008, 2012) datasets to create calibration curves between waist and hip circumferences and height derived from ~3500 women in the UK, aged between 18 and 45. Because our CGI model exists in an appropriately scaled 3D world, having set the height of our models (1.6m) we can measure their waist and hip circumferences, and compare these with our Health Survey for England calibration curves in order to compute their BMI (Cornelissen, Bester, Cairns, Tovée, & Cornelissen, 2015).

Psychometric testing

Prior research has shown that an observer's attitudes to their body shape, weight, and eating habits, as well as their self-confidence, can together modulate estimates of their own body size (Cornelissen, Bester, Cairns, Tovée, & Cornelissen, 2015; Cornelissen, Johns, & Tovée, 2013). Therefore, we gathered these psychometric variables in order to characterize our participants and to be able to model potential effects of this kind in our statistical analyses, even though we did not expect to observe any: our participants were merely being asked to tell the difference between pairs of stimuli, and were not required to relate what they saw on screen to their beliefs/attitudes about their own body, as discussed in the Introduction. To assess participants' attitudes to body shape, weight, and eating we used the 16-item Body Shape Questionnaire (BSQ, range 0-96; Evans & Dolan, 1993) which indexes the degree of preoccupation and negative attitude toward body weight and body shape. In addition, we used the Eating Disorders Examination Questionnaire (EDE-Q, range 0-6), which is a self-report version of the Eating Disorder Examination (EDE) structured interview (Fairburn & Beglin, 1994). This is commonly used as a screening questionnaire for eating disordered behaviour and has been normed for young women and undergraduates

(Luce, Crowther, & Pole, 2008; Mond, Hay, Rodgers, & Owen, 2006). The questionnaire contains four subscales reflecting the severity of aspects of the psychopathology of eating disorders: (a) the Restraint (EDE-restraint) subscale investigates the restrictive nature of eating behaviour; (b) the Eating Concern (EDE-eating concerns) subscale measures preoccupation with food and social eating; (c) the Shape Concern (EDE-shape concerns) subscale investigates dissatisfaction with body shape and (d) the Weight Concern (EDE-weight concerns) subscale assesses dissatisfaction with body weight. The EDE-Q also measures overall disordered eating behaviour. Furthermore, it provides frequency data on key behavioural features of eating disorders. We also used the Beck Depression Inventory (BDI) (range 0-63; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) that measures participants' level of depression and the Rosenberg Self-Esteem Scale (RSE) (range 0-30; Rosenberg, 1965) that measures self-esteem.

Procedure

Having completed our set of questionnaires, the participants then completed the psychophysical task. To measure their JNDs at each of the three stimulus orientations (three-quarter, front, and side views), we used a 2-alternative forced choice (2-AFC) discrimination paradigm, based on the method of constant stimuli. The images were presented on a 19" flat panel LCD screen (1280w x 1024h pixel native resolution, 32-bit colour depth). On every trial, participants were presented a pair of images, side by side, and were asked to respond by button press which of the pair (left or right) represented a larger body. We presented 12 blocks of stimuli, each block corresponding to one of the 4 points along the BMI continuum and one of the three orientations. Within each block, we presented pairs of images at each of 13 levels of BMI difference between the left and the right images. One image was always the reference image, for a given BMI range, and it appeared at random on the left or right side with equal probability across trials. Comparisons were only ever made between images of the same orientation, and not between orientations. The set of differences in BMI between the image pairs was 0.0 to 3.0 BMI units in 0.25 BMI steps. The stimulus image pairs were therefore drawn from the 4 BMI ranges: 15-18; 20-23; 27-30; 36-39. Every image pairing, which represented a given BMI difference, was presented 20 times to each observer in order that we could calculate the probability that participants could detect that BMI difference, at that particular stimulus orientation. Each participant therefore carried out 3120 trials.

We randomized the order in which stimuli within a given block were presented, as well as the order of presentation of the BMI ranges and orientations themselves. In order to minimize effects of fatigue, participants were permitted to pause the psychophysical task at any point. Typically, they carried out the complete experiment over the course of two to three days. For each participant, we used probit analysis to fit psychometric functions which plot the percentage of correct 'this is the larger image' responses as a function of the difference in BMI between the image pairs. From this analysis, we extracted the BMI difference corresponding to the point of subjective equality (i.e., the PSE, where participants are responding at 50% correct) and the 75% correct response rate. The difference between these two values is the JND (Gescheider, 1997). For twenty-five out of a total of 240 fits, fiducial limits (i.e., the equivalent of confidence intervals in probit analysis) could not be estimated reliably, and were therefore discarded from the final analysis. JNDs were compared across participants, as a function of BMI and stimulus orientation, to test for Weber's law behaviour as well as any differences in sensitivity due to stimulus orientation.

Results

Univariate statistics

The responses to the questionnaires across the sample showed good internal reliability. For BSQ, EDEQ, RSE, and BDI, Cronbach's alpha was: .95, .94, .94, and .93 respectively. Table 1 shows the means and standard deviations for the psychometric performance for all 20 female participants. The mean BSQ score shown in Table 1 is consistent with mild concern with body shape (Evans & Dolan, 1993). The mean BDI and RSE scores are consistent with their minimal and normal ranges respectively, and the EDE-Q subscales are all within the normal range for women within this age group (Mond, Hay, Rodgers, & Owen, 2006).

Multivariate statistics: which stimulus orientations produce linear responses?

Figure 1 shows the mean JND across participants plotted as a function of the reference BMI for the 4 BMI ranges, separately for the 3 stimulus orientations. Consistent with Cornelissen et al. (2016), Fig. 1 shows very clearly on inspection, that participants viewing stimuli presented at the three-quarter and side view orientations produced the most linear pattern of responses. Indeed, the Weber fractions (i.e., $\Delta I / I = K$, where I = stimulus magnitude and K = constant) for these stimulus orientations at each of the reference BMIs were consistent with each other. For the three-quarter and side views they were: 0.082, 0.080, 0.077, & 0.082 and 0.082, 0.084, 0.071, & 0.075 respectively. The greatest departure from a linear pattern of responses was observed with participants judging stimuli in front view. For these judgements, the JNDs for the normal (BMI = 20) and overweight ranges (BMI = 27) were increased and showed elevated Weber fractions: 0.094, 0.124, 0.105, & 0.078. We used PROC MIXED in SAS v9.4 to run three separate repeated measures models, one for each stimulus orientation, to test statistically for non-linearity in the relationship between JND and reference BMI. Each model was optimized by ensuring that: (a) the change in -2 log-likelihood between the empty and full models was statistically significant, (b) second order polynomial terms were only retained if they produced a significant reduction in -2 log-likelihood and were statistically significant at $p<.05$.

The relationship between JND and reference BMI showed significant variance in intercepts across participants for the front and side views: Var(u_{0j}) = 0.036, Z = 2.05, p = .02 and Var(u_{0j}) = 0.038, Z = 1.91, p = .03, respectively. The models for the three-quarter and side views were linear, showing significant main effects for reference BMI only. For the three-quarter view, β = 0.024, $t(1, 51)$ = 5.58, $p<.0001$; 95%CI[0.015 – 0.033]. For the side view, β = 0.021, $t(1, 52)$ = 5.58, $p<.0001$; 95%CI[0.013 – 0.028]. However, the model for the front view was non-linear, and included a significant second order term for reference BMI. For the front view: BMI, β = 0.12, $t(1, 50)$ = 3.87, p=.0003; 95%CI[0.056 - 0.18] and BMI^2, β = -0.0019, $t(1, 50)$ = -3.27, p=.0019; 95%CI[-0.0031 - -0.00074].

Multivariate statistics: which orientations show differences at each reference BMI?

Aside from determining whether participants' response patterns were linear or not, we also wanted to know whether there were any statistically significant differences between the JNDs for each orientation, at each reference BMI. We used PROC MIXED in SAS v9.4 to build a mixed model to quantify the relationship between JND, reference BMI, and orientation. We included individual intercept variation for each subject by specifying an 'unstructured' variance–covariance structure for this random effect in the model. We computed all pairwise post-hoc comparisons (corrected for multiple comparisons) between the stimulus orientations, separately for each reference BMI. The Type III (i.e., not model order dependent) test of the fixed effects of reference BMI and stimulus orientation were statistically significant: $F(3, 185) = 29.67$, $p < .0001$, and $F(2, 185) = 4.15$, $p = .02$, respectively. Post-hoc pairwise comparisons, corrected for multiple comparisons, were statistically significant between the front and three-quarter, and the front and side views, at reference BMI 20: $t(1, 185) = 2.17$, $p = .03$, $d = 0.49$, 95%CI[0.018 – 0.37] ; $t(1, 185) = 2.23$, $p = .03$, $d = 0.50$, 95%CI[0.022 – 0.37] and reference BMI 27: $t(1, 185) = 1.93$, $p = .05$, $d = 0.43$, 95%CI[0.0035 – 0.34]; $t(1, 185) = 2.93$, $p = .004$, $d = 0.65$, 95%CI[0.082 – 0.42] respectively. We then checked whether this model could be improved by including age, participant BMI, BSQ, BDI, RSE, and EDE-global as covariates. To do this, we added each covariate separately to the model above, ran the new model with the added covariate, and checked whether this improved model fit compared to the model without a covariate. (We looked both for significant changes in -2 Log-likelihood between models, as well as whether the beta weight for the covariate was statistically significant). As expected, none of the 6 covariates had any statistically significant influence on JND or overall model fit.

This analysis shows that, statistically speaking, the pattern of responses derived from stimuli presented at all three orientations (i.e., three-quarter, front, and side views) were equivalent to each other for the underweight and obese images. Moreover, the side and three-quarter view responses were also equivalent to each other for the normal and overweight images. However, the JNDs for front view images for the normal and overweight images were significantly higher than those for the corresponding side and three-quarter views. This suggests that judgements with the front view are considerably less precise over this range, particularly in view of the fact that the Weber fractions for the front view were the least consistent of all. With respect to the side and three-quarter views, both showed linear response patterns and we could find no significant differences in the pairwise comparisons, suggesting equivalent levels of precision. Nevertheless, the three-quarter view showed more consistent Weber fractions over the range of reference BMIs, and may therefore be considered optimal.

Stimulus features that drive the JND

When female participants make judgements about female body size, they spend most of their time looking up and down the body, fixating between the top of the thighs and just below the costal margin (i.e., the lower edge of the chest formed by the bottom edge of the rib cage) (Cornelissen, Hancock, Kiviniemi, George, & Tovée, 2009). Moreover, in this region of the female human body, there is a linear relationship between BMI and both waist and hip circumferences (Cornelissen, Tovée, & Bateson, 2009). In other words, the most salient change in body shape that reflects changes in BMI is the horizontal separation of the left and right abdominal profiles. Added to this, there are also a set of predictable, localized, non-linear shape changes (see Figure 4, Crossley, Cornelissen, & Tovée, 2012). This suggests that there might be a very straightforward account of the Weber's law behaviour for detecting BMI that we observed. Specifically, since BMI is linearly related to the horizontal separation of the left and right abdominal profiles, then, for a unit increase in BMI, the proportional change in abdominal width(s) should be a negative, decelerating function of BMI. To illustrate, the average waist circumferences of UK women aged between 18 and 40, for the BMIs 15, 16, 34, and 35 are: 60.67, 62.71, 99.58, and 101.63cm as defined by the Health Survey for England (2008, 2012). Therefore, for a unit change in BMI from 15 to 16, the percentage increase in waist circumference is 3.27% compared to the corresponding change between BMIs 34 and 35, which is only 2.02%. In other words, as the percentage change in abdominal widths reduces with increasing BMI, we might expect perceptual JNDs for detecting the smallest difference in BMI to increase correspondingly, in a simple linear fashion. To test this prediction, we measured abdominal slice widths in our stimuli in 6 equally spaced slices from the subcostal region to the top of the thighs, at the reference BMIs of 15, 20, 27, and 36 as well as for the image corresponding to the respective JNDs, separately for the three stimulus orientations (See Figure 2a). Figure 2b shows plots of these data as a function of slice location. It is immediately clear that the difference in slice widths between the reference image and the corresponding image at the JND increases systematically with BMI, across all slices, and is therefore broadly consistent with Weber's law behaviour. Table 2 shows the mean difference, averaged across slice locations.

Table 2 also shows that the differences in mean slice width at reference BMIs 20 and 27 are larger for the front view, compared to both the side and three-quarter views in Table 2, consistent with the elevated JNDs that we observed (See Fig. 1). We hypothesized that this might be caused by differential widening with increasing BMI of the anterior-posterior dimension of the abdomen, in the sagittal plane[1], as compared to the lateral, left to right width in the coronal plane. To test this, as shown in Fig. 2c, we plotted the waist widths of the 50 women who agreed to be photographed in both front and side views in a previous study (Tovée & Cornelissen, 2001). Ordinary least squares regression showed regression coefficients for BMI of 0.180 and 0.143 respectively for the side and front view. In other words, the regression of waist width on BMI for the side views was 25.8% steeper than that for the front views, suggesting a more rapid increase in width with increasing BMI.

[1] The sagittal plane is an anatomical plane parallel to the sagittal suture which divides the body into left and right. The coronal plane is any vertical plane passing through the heart that divides the body into dorsal and ventral (back and front, or posterior and anterior) portions.

Moreover, we used PROC MIXED in SAS v 9.4 to compute a mixed model of these waist widths with BMI ($F(1, 46) = 792.56$, $p < .0001$), view ($F(1, 46) = 143.45$, $p < .0001$) and the interaction between BMI and view ($F(1, 46) = 18.06$, $p < .0001$) as main effects. The fact that the interaction term was statistically significant confirms that the waist widths of women increase faster in the sagittal plane (visible in three-quarter and side views, but not front view) than the coronal plane (visible in all three views) with increasing BMI (see Fig. 2d), and this effect may therefore have contributed to the elevated JNDs for the front view in the current study.

Discussion

We argue that because body size (indexed by BMI) has magnitude, we should expect that: (a) when human observers compare the size of pairs of bodies (i.e., a reference and a test) they should show just noticeable differences that scale linearly with increasing reference BMI and (b) that observers' JNDs should correspond to a constant proportion of the reference stimulus BMI. In short, we should expect human performance in body size judgement to conform to Weber's law. We also argued that this expectation can only be met if stimuli are configured to represent BMI dependent body shape change accurately, and in a way that is perceptually available to observers; i.e., the stimuli must have content validity. We tested which of three CGI body stimulus orientations: side, front, and three-quarter view, met these expectations and in so doing, would be suitable for building tasks that allow observers to estimate their own body size. The results were unambiguous. The three-quarter and side view stimuli produced responses that had the closest fit to Weber's law, with both a linear increase in JND and, particularly for the three-quarter view, a constant Weber fraction. In addition, the mean JNDs for the three-quarter and side views at each of the reference BMIs (corresponding to underweight, normal, overweight, and obese) could not be discriminated statistically. Therefore, to all intents and purposes, performance with the three-quarter and side view stimuli could be considered equivalent. The front view stimuli produced mean JNDs with the largest standard deviations at each reference BMI. While there were no statistically significant differences between these means at any of the three orientations for underweight and obese images, the JNDs for normal and overweight front view images were significantly increased compared to both the three-quarter and profile views. This loss of precision for normal and overweight images produced a substantial and significant non-linearity in the plot of JND as a function of BMI. Therefore, the front view images departed substantially from expected Weber's law behaviour.

Based on these results for the CGI stimuli used in this study, we would therefore choose either side or three-quarter view stimuli to build a body size estimation task, and not front view stimuli. Clearly, this investigation of basic stimulus properties would need to be repeated to compare JNDs at the same three orientations for line drawn stimuli of the kind originally developed by Stunkard et al. (1983) and also for photographic stimuli of real people, to identify which mode of stimulus presentation produces Weber's law behaviour. With respect to the photographic images, Cornelissen et al. (2016) report JNDs for front view stimuli in a 2-AFC discrimination task which used photographs of 6 different people at each reference BMI (representing a range of 0 to 2.5 BMI units in steps of 0.5). While the regression of JND against reference BMI was linear, nevertheless the Weber fraction, $\Delta I / I$, was far from constant over the reference BMI range, and therefore Weber's law was not adhered to.

What causes the differences in precision between stimulus orientations?

At least part of the reason why precision is so impaired for normal and overweight images in front view may have to do with a visual occlusion effect. As illustrated in Fig. 2c & 2d, the anterior to posterior width in the central abdomen (sagittal plane) increases more rapidly than the corresponding width in the lateral (coronal) plane, and this could represent a more salient cue to BMI difference in principle. However, unlike the side and three-quarter views, the front view automatically occludes this beneficial information because the changes are occurring directly along the line of sight and may well not be correlated with easily detectable changes in cues that allow observers to infer depth from shading. Therefore, in the absence of any other visual cues to compensate for this information loss, precision in body size estimation in the normal and overweight ranges for front view is impaired. The fact that the underweight and obese judgements do not suffer an equivalent loss of precision (although all front view responses are associated with the highest standard deviations for JND) may be because alternative and equally powerful cues are available to observers in front view for these body sizes – we should again note that BMI dependent body shape change has strong non-linear components (Crossley, Cornelissen, & Tovée, 2012), so it is perfectly plausible that complementary sources of information may be available at different stimulus orientations and body sizes.

While the preceding discussion illuminates why the front view may be sub-optimal, thereby reducing content validity, there are other reasons why the three-quarter view may indeed be optimal, and maximize content validity. Recognition and discrimination studies in object perception have suggested an improved performance when stimuli are presented in three-quarter view. This orientation is referred to as the canonical view. It is hypothesised that these recognition and discrimination judgements occur by comparing a novel view of an object against their stored prototypes (Edelman & Duvdevani-Bar, 1997; Palmer, Rosch, & Chase, 1981; Ullman, 1996). Viewpoints similar to, or the same as, the internal representation or representations allow participants to show improved performance. Previous studies have suggested that we make body judgements by comparison to a stored prototype or template, and this suggests that there may also be a similar canonical advantage for body judgements (Cornelissen, Bester, Cairns, Tovée, & Cornelissen, 2015; Cornelissen, Johns, & Tovée, 2013; Winkler & Rhodes, 2005).

Why do these basic stimulus properties matter?

Our data clearly show that the front view fails to produce Weber's law behaviour when participants are trying to tell apart pairs of images that differ in BMI. Specifically, our results show a loss of precision for these judgements in the normal and over-weight image ranges, but not the underweight or obese ranges. The *implication* of this finding is that if participants, who believe themselves to have a BMI in the normal-to-overweight range, used the same stimuli to judge their own body image, then the loss of precision (due to the front view stimuli) could lead to substantially greater variance in participants' responses than would be the case with the three-quarter or side view stimuli. The consequences of this are unknown currently, and would need to be investigated in a future study. However, we suggest at least two possible outcomes. In the first case, let us imagine that these stimuli, each of which is calibrated for BMI, are being used in an epidemiological study of obesity rates (cf. Dratva et al., 2016). Participants are being asked to identify which stimulus image is closest to the body size they think they have. Consider the average response across a set of, say, 100 overweight men whose average *actual* BMI is 27. Suppose that the mean

BMI of the images chosen to represent these men's body size is also 27 irrespective of whether they viewed the three-quarter, side or front view stimuli. If the standard deviation for both the three-quarter and side view responses is 3, then ~16% of the men would have given false positive responses consistent with being obese (i.e., BMI > 30). From our data in the current study, the JND at BMI 27 is ~25% greater for the front than the three-quarter or side views. Therefore, the standard deviation of the men's responses to the front view stimuli might be increased to ~3.75, leading to a false positive rate for obesity of ~21%. In short, loss of precision as a result of using the front view images could lead to elevated false positive rates in this group of individuals. The second scenario we imagine requires not only a loss of precision, leading to greater uncertainty in body shape/size estimation, but also a second factor which biases the average of a set of responses towards a new higher (or lower) location in the face of the increased uncertainty. Cornelissen et al. (2015) propose such a scenario for anorexia nervosa. In this case reduced sensitivity for body size judgements at higher BMIs (i.e., elevated JNDs) together with a pathological insistence for making correct responses, could in principal lead to body-size over-estimation.

This study addresses the visual estimation of the whole body, and does not consider judgements of individual body parts. A simple body scale such as we have discussed here cannot easily index weight change specific to individual body parts, which may be better addressed using interactive programmes which allow the adiposity of individual body parts to be independently varied (e.g., Crossley, Cornelissen, & Tovée, 2012; Tovée, Benson, Emery, Mason, & Cohen-Tovee, 2003). The best viewing angle to judge these changes would have to be assessed in additional, separate studies. Another limitation of using figure rating scales in isolation is that the results do not indicate level of importance of physical appearance, and do not provide indications of which body parts an individual may be most dissatisfied with as they are reporting overall dissatisfaction with their current appearance. For a fuller assessment, the use of body scales might therefore be combined with the use of behavioural or qualitative measures.

In conclusion, our results suggest that viewing orientation has a significant impact on the smallest difference in BMI that participants can detect when discriminating between pairs of images. This result may have important implications for the design of tasks used to measure body image. Future studies may need to consider the use of a three-quarter view for stimulus orientation, which captures both front- and profile view cues and represents a more ecologically valid, naturalistic view than a simple profile.

Data Statement

The raw data is available to download from https://goo.gl/cyv6b0

References

Adair, E. R., Stevens, J. C., & Marks, L. E. (1968). Thermally induced pain, the Dol Scale, and the psychophysical power law. *American Journal of Psychology*, *81*(2), 147-164.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571. Doi: 10.1001/archpsyc.1961.01710120031004

Brodie, D. A., Bagley, K., & Slade, P. D. (1994). Body-image perception in pre- and post-adolescent females. *Perceptual Motor Skills, 78,* 147–154.

Cash, T. F., & Pruzinsky, T. (Eds.) (2002). *Body image: A handbook of theory, research, and clinical practice*. New York: Guilford Press. Doi: 10.1037/10516-161

Cohen, E., Bernard, J. Y., Ponty, A., Ndao, A., Amougou, N., Saïd-Mohamed, R., & Pasquet, P. (2015a). Development and Validation of the Body Size Scale for Assessing Body Weight Perception in African Populations. *PLoS One*. Doi: 10.1371/journal.pone.0138983

Cohen, E., Ndao, A., Boëtsch, G., Gueye, L., Pasquet, P., Holdsworth, M., & Courtiol, A. (2015b). The relevance of the side-view in body image scales for public health: an example from two African populations. *BMC Public Health, 15,* 1169. Doi: 10.1186/s12889-015-2511-x

Cornelissen, K. K., Bester, A., Cairns, P., Tovée, M. J., & Cornelissen, P. L. (2015). The influence of personal BMI on body size estimations and sensitivity to body size change in anorexia spectrum disorders. *Body Image, 13*, 75-85. Doi: doi.org/10.1016/j.bodyim.2015.01.001

Cornelissen, K. K., Gledhill, L., Cornelissen, P. L., & Tovée, M. J. (2016). Visual biases in judging body weight. *British Journal of Health Psychology, 21,* 555-569. Doi: 10.1111/bjhp.12185

Cornelissen, P. L., Hancock, P. J. B., Kiviniemi, V., George, H. R., & Tovée, M. J. (2009). Patterns of eye movements when male and female observers judge female attractiveness, body fat and waist-to-hip ratio. *Evolution and Human Behaviour, 30,* 417-428. Doi: 10.1186/s12889-015-2511-x

Cornelissen, P. L., Johns, A., & Tovée, M. J. (2013). Body size over-estimation in women with anorexia nervosa is not qualitatively different from female controls. *Body Image*, *10,* 103-111. Doi: 10.1016/j.bodyim.2012.09.003

Cornelissen, P. L., Tovée, M. J., & Bateson, M. (2009). Patterns of subcutaneous fat deposition and the relationship between body mass index and waist-to-hip ratio: Implications for models of physical attractiveness. *Journal of Theoretical Biology, 256,* 343-350. Doi: 10.1016/j.jtbi.2008.09.041

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*. Doi: 10.1371/journal.pone.0050601

Dratva, J., Bertelsen, R., Janson, C., Johannessen, A., Benediktsdóttir, B., Bråbäck, L., . . . Gomez Real, F. (2016). Validation of self-reported figural drawing scales against anthropometric measurements in adults. *Public Health Nutrition, 19*(11), 1944-1951. Doi: 10.1017/S136898001600015X

Edelman, S., & Duvdevani-Bar, S. (1997). Similarity-based viewspace interpolation and the categorization of 3D objects. In *Proc. Similarity and Categorization Workshop* (pp. 75–81). Edinburgh, U.K.: Dept. of AI, University of Edinburgh.

Evans, C., & Dolan, B. (1993). Body Shape Questionnaire: derivation of shortened “alternate forms”. *International Journal of Eating Disorders*, *13,* 315-321. Doi: 10.1002/1098-108X(199304)13:3%3C315::AID-EAT2260130310%3E3.0.CO;2-3

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16,* 363-370. Doi: 10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-#

Gardner, R. M., & Brown, D. L. (2010). Body image assessment: A review of figural drawing scales. *Personality and Individual Differences, 48,* 107–111. Doi: 10.1016/j.paid.2009.08.017

Gardner, R. M., & Brown, D. L. (2011). Method of presentation and sex differences when using a revised figural drawing scale to measure body size estimation and dissatisfaction. *Perceptual and Motor Skills, 113,* 739-750. Doi: 10.2466/07.17.27.PMS.113.6.739-750

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,* 113–122. Doi: 10.1002/jclp.20526

Gescheider, G. A. (1997). *Psychophysics: the fundamentals (3rd ed.)*. New Jersey: Lawrence Erlbaum.

Grabe, S., Ward, L. M., & Hyde, J. S. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. *Psychological Bulletin, 134,* 460-476. Doi: 10.1037/0033-2909.134.3.460

Harris, C. V., Bradlyn, A. S., Coffman, J., Gunel, E., & Cottrell, L. (2008) BMI-based body size guides for women and men: development and validation of a novel pictorial method to assess weight-related concepts. *International Journal of Obesity*, *32,* 336–42. Doi: 10.1038/sj.ijo.0803704

Hasler, N., Stoll, C., Sunkel, M., Rosenhahn, B., & Seidel, H-P. (2009). A Statistical Model of Human Pose and Body Shape. *Computer Graphics Forum*, *28*(2), 337-346. Doi: 10.1111/j.1467-8659.2009.01373.x

Jesteadt, W., Wier, C. C., & Green, D. M. (1977). Intensity discrimination as a function of frequency and sensation level. *The Journal of the Acoustical Society of America*, *61*(1), 169-177. Doi: 10.1121/1.381278

Kline, P. (2015). *A handbook of test construction: Introduction to psychometrics*. New York, NY: Routledge.

Kreidler, S. M., Muller, K. E., Grunwald, G. K., Ringham, B. M., Coker-Dukowitz, Z. T., Sakhadeo, U. R., . . . Glueck, D. H. (2013). GLIMMPSE: Online Power Computation for Linear Models with and without a Baseline Covariate. *Journal of Statistical Software* *54*(10), i10. Doi: 10.18637/jss.v054.i10

Li, Y., Hu, X., Ma, W., Wu, J., & Ma, G. (2005). Body image perceptions among Chinese children and adolescents. *Body Image*, *2,* 91–103. Doi: 10.1016/j.bodyim.2005.04.001

Luce, K. H., Crowther, J. H., & Pole, M. (2008). Eating Disorder Examination Questionnaire (EDE-Q): norms for undergraduate women. *International Journal of Eating Disorders*, *41,* 273-276. Doi: 10.1002/eat.20504

Marlowe, F., Apicella, C., & Reed, D. (2005). Men's preferences for women's profile waist-to-hip ratio in two societies. *Evolution and Human Behaviour, 26,* 458–468. Doi: 10.1016/j.evolhumbehav.2005.07.005

Mond, J. M., Hay, P., Rodgers, B., & Owen, C. (2006). Eating Disorder Examination Questionnaire (EDE-Q): norms for young adult women. *Behaviour Research and Therapy, 44,* 53-62. Doi: 10.1016/j.brat.2004.12.003

National Centre for Social Research, University College London. Department of Epidemiology and Public Health. (2013*). Health Survey for England, 2008*. [data collection]. *4th Edition*. UK Data Service. SN: 6397. Doi: 10.5255/UKDA-SN-6397-2

National Centre for Social Research, University College London. Department of Epidemiology and Public Health. (2014). *Health Survey for England*, 2012. [data collection]. UK Data Service. SN: 7480. Doi: 10.5255/UKDA-SN-7480-1

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2014). Prevalence of childhood and adult obesity in the United States, 2011-2012. *JAMA, 311*(8)*,* 606-814. Doi: 10.1001/jama.2014.732

Palmer, S. E., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and Performance IX* (pp. 135-151). Hillsdale, NJ: Erlbaum.

Peterson, M., Ellenberg, D., & Crossan, S. (2003). Body-image perceptions: Reliability of a BMI-based silhouette matching test. *American Journal of Health Behaviour, 27,* 355–363. Doi: 10.5993/AJHB.27.4.7

Pope, H., Phillips, K., & Olivardia, R. (2000). *The Adonis complex*. New York, NY: Free Press.

Rilling, J. K., Kaufman, T. L., Smith, E. O., Patel, R., & Worthman, C. M. (2009). Abdominal depth and waist circumference as influential determinants of human female attractiveness. *Evolution and Human Behaviour, 30,* 21–23. Doi: doi:10.1016/j.evolhumbehav.2008.08.007

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Smith, K. L., Cornelissen, P. L., & Tovée, M. J. (2007) Colour 3D bodies and judgements of human female attractiveness. *Evolution and Human Behaviour*, *28*, 48-54. Doi: doi:10.1016/j.evolhumbehav.2006.05.007

Stunkard, A., Sorensen, T., & Schulsinger, F. (1983) Use of the Danish Adoption Register for the study of obesity and thinness. *Research Publications - Association for Research in Nervous & Mental Disease, 60*, 115–120.

Swami, V., Frederick, D. A., Aavik, T., Alcalay, L., Allik, J., Anderson, D., . . . Zivcic-Becirevic, I. (2010). The attractive female body weight and female body dissatisfaction in 26 countries across 10 world regions: results of the international body project I. *Personality and Social Psychology Bulletin, 36,* 309-325. Doi: 10.1177/0146167209359702

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

Swinburn, B. A., Sacks, G., Hall, K. D., McPherson, K., Finegood, D. T., Moodie, M. L., & Gortmaker, S. L. (2011). The global obesity pandemic: shaped by global drivers and local environments. *Lancet, 378,* 804–814. Doi: 10.1016/S0140-6736(11)60813-1

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

Tovée, M. J., Benson, P. J., Emery, J. L., Mason, S. M., & Cohen-Tovee, E. M. (2003) Measurement of body size and shape perception in eating-disordered and control observers using body-shape software. *British Journal of Psychology*, *94*, 501-516. Doi: 10.1348/000712603322503060

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. Doi: 10.1348/000712601162257

Tovée, M. J., Maisey, D. S., Emery, J. L., & Cornelissen, P. L. (1999). Visual cues to female physical attractiveness. *Proceedings of the Royal Society, London B Bio, 266,* 211-218. Doi: 10.1098/rspb.1999.0624

Ullman, S. (1996). *High level vision*. Cambridge, MA: MIT Press.

Weber, E. H. (1834). *De pulsu, resorptione, auditu et tactu: Annotationes anatomicae et physiolologicae*. Leipzig: Koehler.

Winkler, C., & Rhodes, G. (2005). Perceptual adaptation affects attractiveness of female bodies. *British Journal of Psychology, 96,* 141-154. Doi: 10.1348/000712605X36343.

Table 1: Demographic and questionnaire data from 20 participants.

Variable	*M (SD)*
Age (years)	25.40 (4.72)
BMI	22.66 (4.00)
BSQ	47.15 (18.0)
EDE-global	2.04 (1.21)
EDE-restraint	1.87 (1.21)
EDE-eating concerns	1.14 (1.07)
EDE-shape concerns	2.38 (1.62)
EDE-weight concerns	2.67 (1.63)
BDI	8.40 (8.37)
RSE	21.10 (6.39)

Note: BMI = Body mass index; BSQ = 16-item Body Shape Questionnaire; EDE-global = Eating Disorder Examination Questionnaire global score; EDE-restraint = Eating Disorder Examination Questionnaire eating restraint subscale; EDE-eating concerns = Eating Disorder Examination Questionnaire eating concern subscale; EDE-shape concerns = Eating Disorder Examination Questionnaire body shape concern subscale; EDE- weight concerns = Eating Disorder Examination Questionnaire weight concern subscale; BDI = Beck Depression Inventory; RSE = Rosenberg Self-Esteem Scale.

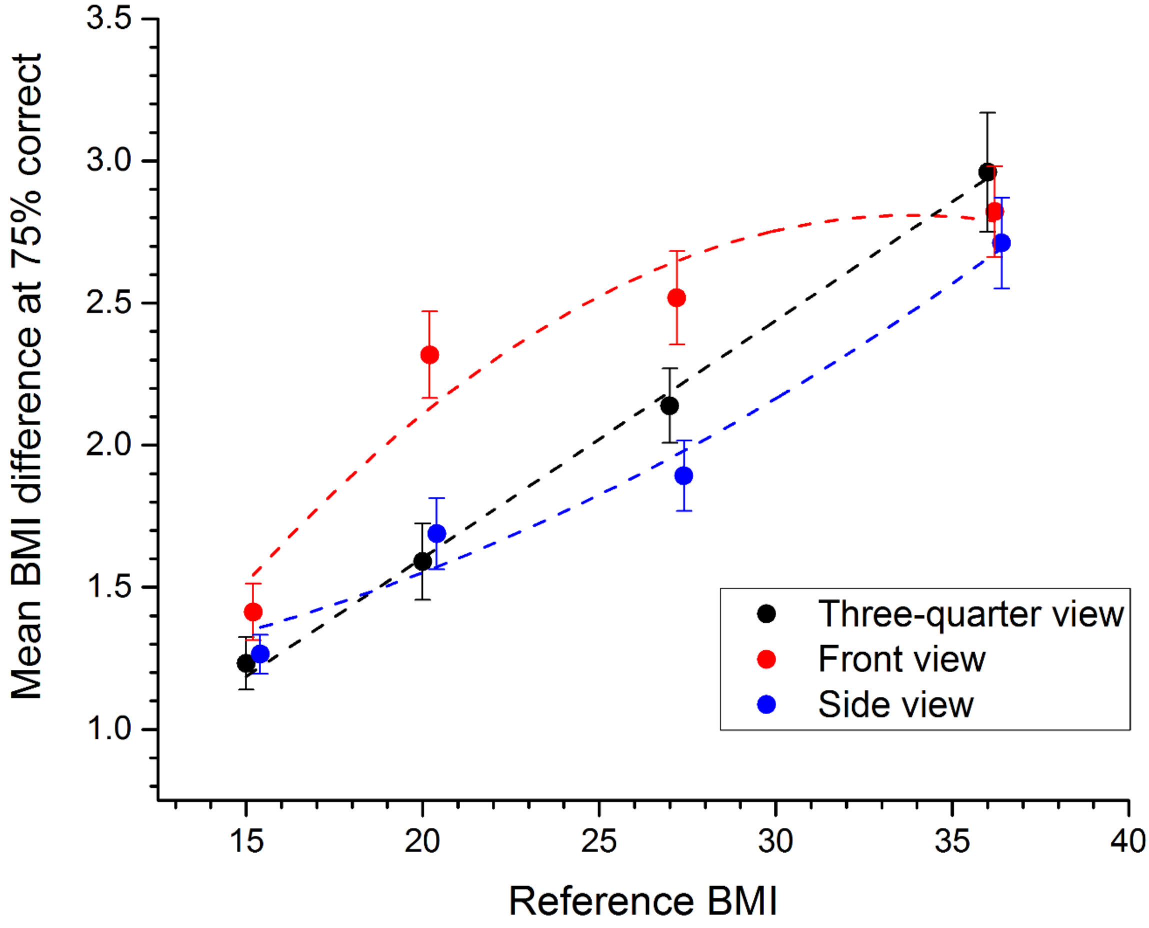
Table 2: Mean differences in slice width between reference BMI stimulus and stimulus at the JND.

Reference BMI	Three-Quarter View (pixels)	Front View (pixels)	Side View (pixels)
	M (SD)	*M (SD)*	*M (SD)*
15	4.58 (2.39)	3.02 (2.14)	6.54 (0.77)
20	6.94 (1.16)	8.10 (1.47)	6.83 (1.02)
27	9.58 (1.21)	11.42 (1.50)	8.97 (2.37)
36	13.25 (2.49)	11.69 (1.39)	12.57 (3.70)

Figure Legends

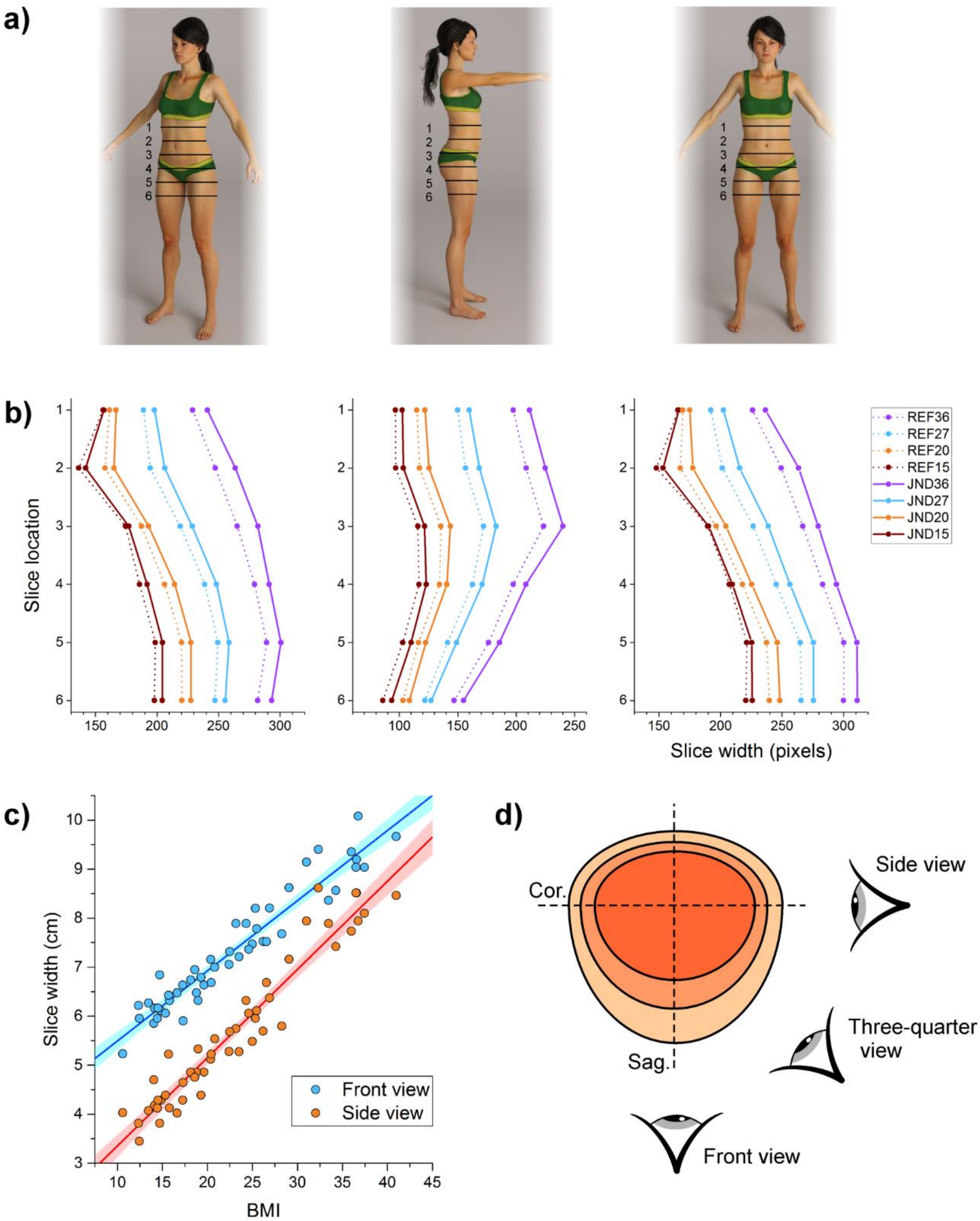
Figure 1: This shows a plot of mean JND (i.e., the difference in BMI between the reference and stimulus bodies at 75% correct response rate) as a function of the reference BMI value. Black circles represent the three-quarter view, red squares the front view and blue triangles the side view. The error bars represent standard errors of the respective means, corrected for repeated measures. Points at each reference BMI are offset horizontally so that error bars are visible. The dotted lines represent second order polynomial regression fits to the data for the front and side views, and a linear fit to the three-quarter view. See text for details.

Figure 2 A: The locations of the slice widths measured from the stimuli at each of the three orientations. **B:** Three plots showing the relationship between slice width as a function of slice location for the reference images (dotted lines) and the stimuli at the JND (solid lines). C: Plots of waist width seen from front (blue) and side (orange) views from 50 photographs of women in Tovée & Cornelissen, 2001. The blue and red lines represent the OLS regression lines through the respective data together with their 95% confidence intervals (cyan and pink shading). D: Illustration of abdominal cross-section with progressively increasing BMI. It shows how width increases in the sagittal (Sag.) plane more quickly than in the coronal (Cor.) plane, and how this is harder to see in front view than either the side of three-quarter view.



Mean BMI difference at 75% correct
3.5
3.0
2.5
2.0
1.5
1.0
15
20
25
30
35
40
Reference BMI
Three-quarter view
Front view
Side view

654


a)
b)
Slice location
Slice width (pixels)
REF36
REF27
REF20
REF15
JND36
JND27
JND20
JND15
c)
Slice width (cm)
BMI
Front view
Side view
d)
Cor.
Sag.
Side view
Three-quarter view
Front view

Supplementary Materials

In this study we used the same computer-generated imagery (CGI) methods as the film and games industries to create 3D images representing a full spread of BMI. This strategy therefore amounts to an updated version of a figural rating scale, like the Stunkard scale (Stunkard, Sorensen, & Schulsinger, 1983), with the advantage of a continuous variation in BMI, as well as highly realistic 3D imagery.

All the CGI stimuli were created in the Daz Studio v4.8 modelling environment. This program allows subtle manipulation of the body shape and posture of a fully rigged digital model. We used the Victoria 6 character model, which is based on the Genesis 2 female base model, in Daz Studio. From the neck down, there are 320 body shape controls, 16 of which influence whole body attributes such as adiposity. From the neck up there are 209 controls for head shape. For this study, we modified the Victoria 6 character model to capture the average body shape of a 25 year old UK Caucasian female, and this provided our baseline model whose adiposity we could then vary systematically. To do this, we extracted the appropriate averages from the Health Survey for England (2008, 2012) datasets to select the model's height, leg length, bust circumference, under-bust circumference, waist circumference and hip circumference. In addition, we ensured that these baseline models had an average 25-year old female's torso-to-leg ratio and waist-to-hip ratio.

The first question was whether participants judged the Victoria 6 baseline model to be a plausible representation of female body shape. To address this question, we applied the adiposity morphs to render a set of three images intended to capture the underweight, normal weight and overweight classifications defined by the World Health Organization

(WHO). We then asked 30 participants who were recruited from amongst friends and colleagues to provide qualitative feedback about these images. In addition, we carried out two further comparisons. First, the 3D volumes of the CGI modelled bodies were compared to a 3D statistical model of the relationship between BMI and shape changes in 114 scanned bodies (Hasler, Stoll, Sunkel, Rosenhahn, & Seidel, 2009). Secondly, we compared our models qualitatively to digital photographs of 220 women in a standard pose who vary in BMI from 11 (emaciated) to 45 (obese) (Tovée, Maisey, Emery, & Cornelissen, 1999). Based upon all the feedback we received, we further modified our baseline model by reducing chest size and shape to represent a more naturalistic breast shape, made the lips thinner, the eyes smaller and cheeks (buccae) flatter.