Title: **Women's self-estimates of body size are more accurate and precise when made with three-quarter view than front-view stimuli**

.

Authors: Piers L. Cornelissen[a], Lise Gulli Brokjøb[a], Jiří Gumančík[a], Katri K. Cornelissen[a],

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

CORRESPONDENCE TO:

Piers Cornelissen: piers.cornelissen@northumbria.ac.uk

Declarations of interest: none

Abstract

Recently, Cornelissen, Cornelissen, Groves, McCarty, & Tovée (2018) asked which image orientations (e.g. front-, side-, or three-quarter view) are most appropriate for tasks which are used for self-estimates of body size and shape. Based on psychophysical measurements, they showed that front view stimuli showed substantially poorer content validity compared to side- and three-quarter view stimuli. Here, we tested the real-world consequences of Cornelissen et al.'s (2018) findings. We carried out a body size self-estimation task in a sample of healthy adult women, once with front view stimuli, and once with three-quarter view stimuli. The order in which front- and three-quarter view tasks were carried out was randomized across participants. Compared to three-quarter view stimuli, we found that: a) the precision of participants' judgements was worse with front view stimuli, and b) that front view stimuli led to over-estimation of body size by ~1.7 BMI units. While these results need to be replicated, they do suggest that careful consideration needs to be given to stimulus orientation in future studies.

Keywords: BMI, body fat, body size judgement, viewpoint, figural body scales

1. Introduction

A core diagnostic feature of anorexia nervosa is a "disturbance in the way in which one's body weight or shape is experienced, undue influence of body weight or shape on self-evaluation, or denial of the seriousness of the current low body weight." (DSM-5; American Psychiatric Association, 2013). This distortion of a person's body image can be partitioned into two broad components which are statistically independent of each other (Cornelissen, Widrington, McCarty, Pollet, Tovée & Cornelissen, 2019): (a) an attitudinal component which corresponds to the feelings someone has about their body, and (b) a perceptual component which captures the size and shape they believe themselves to have. Cash and Deagle's (1997) meta-analysis showed convincingly that when eating disordered individuals are compared with controls, perceptual measures yield effect sizes of d = 0.61 to 0.64 and attitudinal measures from d = 1.10 to 1.13. Research since Cash and Deagle (1997) has continued to focus on perceptual and attitudinal components of body image, with a bias towards the latter, perhaps reflecting the asymmetry of these effect sizes. But for both domains, there is a need for valid, reliable measures that are essential for construct validity and internal validity (Murnen & Smolak, 2019). Here we focus on the validity of perceptual measures of body image.

1.1 Expectations of perceptual measurement

Historically, researchers and clinicians have tried many different ways to measure the body size/shape that someone believes they have. More recent techniques 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 (e.g., Stunkard, Sorensen, & Schulsinger, 1983), or (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 an image closest to the size/shape they believe themselves to have (or would like to have), or participants make decisions across a number of trials about whether any particular stimulus is smaller/larger than they believe themselves to be (or would like to be) (Brodie, Bagley, & Slade, 1994; Gardner & Brown, 2011). Irrespective of the particular methodological details, we contend that all size judgements of this kind should properly be thought of as magnitude estimations, which should therefore conform to Weber's law (1834). This is a historically important psychological law quantifying the perception of change in a given stimulus. Weber's law states that the smallest difference between a pair of stimuli that can be reliably told apart, i.e., the just noticeable difference or JND, is a constant proportion of the stimulus magnitude. (Technically, the proportion is called the Weber fraction $K = \Delta I / I$, where K = a constant, ΔI = the difference in magnitude between a pair of stimuli, and I = reference stimulus magnitude). This principle has been found to hold across a wide range of attributes, measurement procedures and sensory modalities (for reviews see: Billock & Tsou, 2011; Gescheider, 1997). Consequently it has been hailed by some as the starting point of modern scientific psychology (Boring, 1950; Krech & Crutchfield, 1958).

In a typical forced choice task, a participant is presented pairs of stimuli on each trial, and is asked to judge which of the pair is heavier, brighter, louder, sweeter, depending on the particular task demands. In these examples, both stimuli of the pair are external to the observer, and we will refer henceforth to JNDs obtained in this way as JND_{stim}. However, in the context of someone judging their own body size, the "pair of stimuli" comprises: (a) the participant's belief about their own body size, which is a mental representation, and (b) the image they have been presented for comparison, which is of course external to them. If we use body mass index (BMI) as an outcome measure for body size self-estimates, then the JND amounts to the smallest detectable difference in BMI between the body size that a participant believes she has and the body size of the stimulus presented. Henceforth we will refer to this form of a JND measurement as JND_{self}. But more than this, Weber's law also shows how this just noticeable difference in BMI should get systematically larger as people's beliefs about their bodies get heavier. For example, if a participant believes she has a body size in the underweight range, the JND_{self} might be ~1 BMI unit (cf. Cornelissen, Gledhill, Cornelissen, & Tovée, 2016a). But if the participant believes she has a body size in the obese range, the JND_{self} might increase to ~2.5 BMI units. Therefore, a plot of the JND_{self} for BMI (y-axis) as a function of actual BMI (x-axis) should be a straight line with a positive slope. In addition, the Weber fraction, K, should be constant across the actual BMI range. In two recent studies, we have shown that when healthy adult women carry out such a task, their responses do conform to Weber's law very well. In comparison, women with a history of anorexia nervosa show a much steeper plot with a Weber fraction that increases with BMI (Cornelissen, Bester, Cairns, Tovée, & Cornelissen, 2015; Cornelissen, McCarty, Cornelissen, & Tovée, 2017).

1.2 Stimulus content validity

It is one thing to assert that a body size estimation task should conform to Weber's law. But this can only happen if the stimuli themselves represent body size correctly, i.e., they must have 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). So, if the stimuli representing different BMIs lack content validity, then we may nevertheless fail to observe Weber's law behaviour, even though it is expected. The most obvious aspect of content validity for a body size estimation task is that the stimuli are correctly calibrated for BMI. To this end, we and others (see e.g, Mölbert, Thaler, Mohler, Streuber, Romero, et al., 2018) have used computer generated imagery (CGI) techniques which allow the creation of high resolution, photorealistic 3D models, which can be calibrated for BMI. Another important consideration regarding content validity is the orientation of the body in the stimulus images: e.g., whether they appear in front view, side view, or three-quarter view. 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). Yet 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. These concerns raise the question whether content validity could be affected by the orientation of bodies in such stimuli.

As an objective test of stimulus content validity with respect to stimulus orientation, Cornelissen et al. (2018) proposed that participants should carry out a traditional JND_{stim} task, where both stimuli in the pair are external to the observer. Accordingly, on each trial of the task, participants were presented pairs of images of the same CGI avatar on a computer screen. The only difference between the two versions of the same person on screen was their BMI, and participants had to choose whether the version on the left or the right of the screen was heavier. In other words, this was a judgement about another person's body, made from a third-person point of view, which did not require participants to refer to their own body image in any way. It can be equated with choosing which of a pair of lights is brighter, or which of a pair of loudspeakers is louder. Therefore, if performance in this objective situation led to Weber's law behaviour, then the stimuli could be said to have content validity: in the context of magnitude estimation, the stimuli gave rise to the expected pattern of behavioural responses. Using this approach, Cornelissen et al. (2018) measured the JND_{stim} for their CGI stimuli at 4 reference BMI levels (15, 20, 27 & 36). In separate sessions, they repeated this procedure for three different avatar orientations: i.e. on each trial both avatars appeared in front view, or side view, or three-quarter view. The criteria for conforming to Weber's law behaviour were: (a) that there should be a positive, straight line relationship between JND_{stim} and reference BMI, and (b) that the Weber fraction should be constant over the full range of reference BMI. Cornelissen et al. (2018) found that the side and three-quarter view avatars conformed to both criteria for Weber's law behaviour, with the three-quarter view showing the most consistent Weber fraction. In comparison, the front view conformed to neither criterion. Based on these results, the authors suggested that the front view should not be used for self-estimates of body size, because the front view stimuli lacked content validity.

1.3 What might be the impact of poor content validity?

Here, we ask what are the real-world consequences of using stimuli with poor content validity when healthy adult women estimate their own body size/shape? Specifically, we want to know what happens to the just noticeable difference, as well as women's estimates of their own body size, when the body stimuli they are judging are presented in three–quarter view (i.e., good content validity) and front view (i.e., poor content validity). An important point to emphasize in this situation is that any individual participant only produces a single estimate of their own JND_{self} and a single estimate of the body size they believe themselves to have. This is slightly different to the situation just described in Cornelissen et al. (2018), where each of the "pair of stimuli" is external to the viewer, and the same participant can make multiple estimates of JND_{stim} for stimuli at different points along the BMI spectrum. Now, the "pair of stimuli" constitute: (a) the participant's belief about their own body size, which is a mental representation, and (b) the image they have been presented for comparison, in either front or three-quarter view. There are two consequences to this difference. First, the decision about whether participants' responses conform to Weber's law is really a statement about how the group of participants behave overall, because each individual contributes only one data point on a plot of JND_{self} (y-axis) as a function of BMI (x-axis). Secondly, we need to be able to interpret what one person's JND_{self} means, under these circumstances. A useful way to think about JND in this situation is in terms of the precision of a participant's 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.

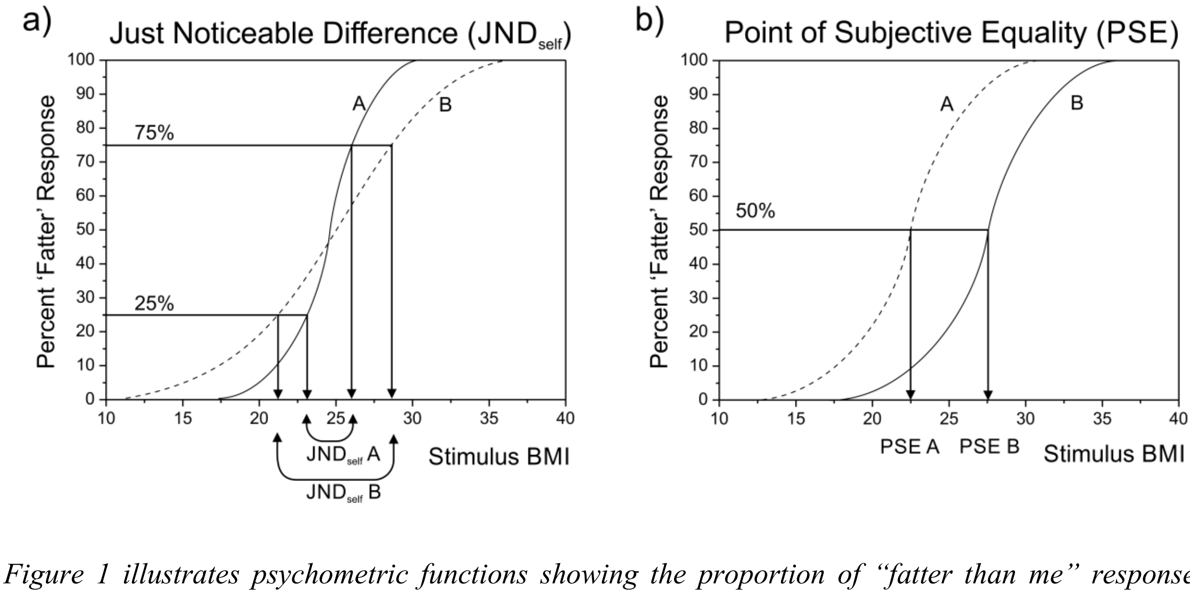

Figure 1 illustrates psychometric functions showing the proportion of "fatter than me" responses plotted as a function of stimulus BMI. (a) Shows how the JND_{self} is derived from the difference in stimulus BMI between the 25% and 75% "fatter than me" responses for two individuals A and B, where A has a steeper psychometric function. As a consequence, JND_{self} for A is smaller than JND_{self} for B. (b)

Shows how the PSE is estimated from the 50% "fatter than me" response rate for two individuals A and B, where A believes she has a small body size than B.

To be explicit, consider the experimental set-up for self-estimates of body size. We use a forced choice task where, on each trial, participants are presented a single CGI avatar selected at random from a wide range of BMI. They are asked to respond by key press whether "this image is fatter than me" or "this image is thinner than me". Figure 1a shows plots of the percentage of "fatter than me" responses as a function of the BMI of the presented stimulus. Typically, these so-called psychometric functions are sigmoidal in shape. Here, JND_{self} is defined as the difference in BMI between the 25% and 75% "fatter than me" responses (cf. Gescheider, 1997). Since the curve for participant A is steeper, this gives rise to a smaller JND_{self} than is the case for participant B, whose curve is shallower. In other words, over the same 25% to 75% range of responses, participant B shows greater variability in stimulus BMI, a higher JND_{self}, and therefore lower precision. As far as individual women's estimates of their own body size are concerned, this judgement is calculated from the same psychometric function, and is defined as the stimulus BMI where the "fatter than me" response rate is 50%. This is known as the point of subjective equality (PSE; Gescheider, 1997), because when presented with a stimulus that is the same size as the participant believes herself to be, she is as likely to respond "this image is fatter than me" as "this image is thinner than me". Figure 1b illustrates the situation for participants A and B who believe their body sizes match CGI avatars with BMIs of 22.5 and 27.5, respectively.

1.4 The current study

In a number of recent studies, we have asked healthy adult women to visually estimate their body size using CGI stimuli in a forced choice task together with the method of constant stimuli (Cornelissen et al., 2015, 2017, 2019; Irvine, McCarty, McKenzie, Pollet, Cornelissen, et al., 2019a). In all of these studies, the avatar stimuli were presented in three-quarter view, which we know have good content validity. We found that women with a low BMI over-estimated their size (indexed by the PSE) and those with a high BMI under-estimated. Those whose body size was closest to the population average for women (i.e., BMI ~27) were most accurate. This is revealed by a regression of estimated body size on actual body size that has a slope of less than one, and is a pattern which is consistent with a normal perceptual phenomenon called contraction bias (Poulton, 1989). In addition, the same studies have shown that visual body size estimation depends simultaneously on attitudinal factors indexed by performance on psychometric tasks measuring attitudes towards body shape, body size and eating habits. Specifically, when self-estimated body size (i.e., PSE) is plotted on the y-axis as a function of actual body size (x-axis), the vertical location of the regression line for this relationship is controlled by attitudinal factors: the same increase in psychological concerns about her body leads to the same increase in body size estimates, anywhere within the range of actual body size. Therefore, in the current study, we used a forced choice task as illustrated in Figure 1 to measure healthy adult women's JND_{self} and PSE using CGI model stimuli presented at three-quarter view and front view. In addition we also measured their attitudes to their own body shape, weight, and eating habits, as well as any tendency towards depressive symptomatology. Our first aim was to check that the findings from previous studies using three-quarter view avatars to measure women's self-estimates of body size were replicated. Specifically, that we see evidence of: (a) contraction bias, and (b) an independent contribution to over-estimation from attitudinal factors. Our second aim was to test which of the two modes of avatar presentation, i.e., three-quarter versus front view, gave rise to JNDs that conformed to Weber's law in the context of self-estimates of body size. Our third aim was to quantify any differences in PSE between the two orientations of stimulus presentation.

2. Methods

2.1. Participants

To our knowledge, this is the only published study of its kind to estimate JND_{self} and PSE for self-estimates of body size, comparing the results obtained from a standard avatar stimulus, presented in front-view and three-quarter view. Therefore, there are no direct precedents from which we can calculate an appropriate sample size. The closest published study which uses the same avatar CGI model for all participants as we do here, and which has also measured JND_{stim} at different orientations of the avatar is that by Cornelissen et al. (2018). However, as described in the Introduction, because this was a test of content validity participants were presented a thinner and a fatter version of the same avatar on screen and asked to judge which of the two was fatter. The JND_{stim} values for these judgements were calculated and compared between three-quarter and front view stimulus presentations. While this design is not identical to the current study in which we measure self-estimates of body size, we argue it is nevertheless similar enough to estimate a sample size for the current study. Therefore, we used the means and SDs for JND_{stim} from Cornelissen et al. (2018), as well as the correlations between responses at different BMI levels (15, 20, 27, & 36) and stimulus orientations (front and three-quarter view) to build a two-factor repeated measures model in GLIMMPSE v3.0 (General Linear Multivariate Model Power & Sample Size; Kreidler et al., 2013). From this, we could calculate sample sizes to estimate the main effects of BMI and avatar orientation at an alpha value = 0.05, and a power = 0.8. This returned integer sample sizes of 9 and 53 respectively. We recruited 65 females (as assigned at birth) from the students and staff at Northumbria University in the Northeast of England who gave their written consent to take part in the study. However, due to a combination of failure to complete all aspects of the study (11 participants), or poor data fits for the psychophysical tasks (2 participants), we excluded 14 of these individuals, leaving complete datasets from 51 participants.

2.2. Psychometric and anthropometric measures.

Previous research (Cornelissen et al., 2015, 2017, 2019; Irvine et al., 2019a) has demonstrated that an observer's attitudes about their own body contribute independently to body size self-estimation; increasing psychological concerns about body shape, weight and eating, together with low self-esteem and a tendency towards depressive symptoms all lead to larger body size estimates. To measure this attitudinal component of participants' body image, we administered the following questionnaires: (a) the Eating Disorders Examination Questionnaire (EDE-Q), which is a self-report version of the Eating Disorder Examination (EDE) interview (Fairburn & Beglin, 1994). The questionnaire contains four sub-scales; the Restraint subscale investigates the restrictive nature of eating, the Eating Concern subscale measures the preoccupation with food and social eating, the Shape Concern subscale measures dissatisfaction with body shape, and the Weight Concern subscale measures dissatisfaction with body weight. A global score of overall disordered eating behaviour is also calculated and frequency data on key behavioural features of eating disorders is provided. (b) The 16-item Body Shape Questionnaire (BSQ) was used to measure body shape preoccupations (Evans & Dolan, 1993). (c) The Beck Depression Inventory (BDI) was used to measure level of depression (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). Participants' Body Mass Index (BMI) was calculated from their weight and height measured with a set of calibrated scales and a stadiometer respectively.

2.3. Stimuli

The stimuli were CGI (computer-generated imagery) images of a standard female model whose BMI ranged from 12.5-55 (see Cornelissen, et al., 2017; Cornelissen, 2016b). The images were created with DAZ v4.8 and were calibrated for BMI, based on the waist and hip circumference data from the Health Survey for England (HSE, 2003, 2008). They were rendered in Luxrender. The advantages of this stimulus set are that the images: (a) are high definition and photorealistic, (b) maintain the identity of the female model across a wide BMI range, and (c) demonstrate extremely realistic changes in BMI-dependent body shape.

2.4. Psychophysical measurement

In this study we used a forced choice task with the method of constant stimuli (cf. Gardner, 1996) to measure two components of the participants' judgements of their own body size: (a) the point of subjective equality (PSE) and (b) the just noticeable difference (JND_{self}). The PSE is the participant's subjective estimate of their own body size. The JND_{self} is an estimate of how sensitive a participant is to the difference in body size between the stimulus presented on screen, and the body size the participant believes they have. It equates to the smallest difference in body size between the two that she can detect reliably (Gescheider, 1997).

Participants carried out the forced choice task twice, once for each stimulus orientation. The order of presentation of stimulus orientation (i.e. front view and three-quarter view) was randomized for each participant. On each instance of the forced choice task, participants were presented with a randomized sequence of images of the standard CGI female body model. Across the image set, BMI varied continuously from 12.5 to 44.5. On each trial of the task, one image was presented, and participants were required to decide whether the body depicted was larger than they were and to record this decision by button press. Stimuli were presented on a 19” flat panel LCD screen (1280w x 1024h pixel native resolution, 32-bit colour depth) for as long as it took participants to make a decision. At the standard viewing distance of ~60cm, the image frame containing the female body subtended ~26° vertically and ~8° degrees horizontally. Each participant first judged seven images covering the whole BMI range (from m 12.5 to 44.5 in equal BMI steps) presented in two separate blocks. Each stimulus image appeared 10 times in each block, and the order of presentation was randomized. Based on the responses from each block, the participants’ point of subjective equality or PSE (the BMI they believe themselves to be) was calculated automatically by fitting a cumulative normal distribution. These two values were then averaged to give an initial estimate of the participant’s PSE. On the basis of this initial estimate, the program presented a further set of 21 images (spread over a range of 5 BMI units centred on the participant’s initial PSE, at a spacing of 0.25 units per image) for the participants to judge. Each image was presented ten times in randomized order. This final set of judgements allowed us to transform the probabilities of larger than responses to z-scores, and use ordinary least squares regression to calculate the PSE as well as the JND_{self} (Gescheider, 1997).

2.4. Procedure

Consenting participants first had their height and weight measured and then they completed the psychometric questionnaires. They were informed that the purpose of the experiment was to assess the accuracy of their judgements about their own body size. Prior to carrying out the psychophysical measurements on the PC, participants were reminded that in the event that they felt fatigued or discomforted in any way, they were at liberty to pause the task for as long as necessary. The whole experiment took ~ 45 minutes for each participant to complete.

3. Results

3.1. Univariate Analysis

Table 1

Characteristics of the 51 female participants

	M	*SD*	Range: Actual	Range: Potential
Age (years)	22.65	6.40	19.00 – 47.0	
Body Mass Index (kg/m^2)	24.24	3.99	17.87 – 32.80	
EDE-Q Global	1.79	1.26	0.09 – 5.42	0 – 6
BSQ-16	46.22	19.80	18.00 – 88.00	16 – 96
BDI	12.75	8.43	0.00 – 36.00	0 – 36

EDE-Q = Eating Disorders Examination Questionnaire, BSQ = Body Shape Questionnaire, BDI = Beck Depression Inventory

The internal reliability of the psychometric measurements was good. Cronbach's alpha for the BSQ, EDEQ, and BDI was .97, .97, and .89 respectively. Descriptive statistics for the 51 female participants are presented in Table 1. Utilizing a cut-off of ≥ 4 as the marker of clinical significance for the EDEQ (Fairburn, Cooper, and O'Connor, 2008), 9.2% of the women in our sample scored in the clinically significant range on restraint, 6.2% on eating concern, 29.2% on shape concern, 24.6% on weight concern, and 10.8% on the global scale. According to cut-off criteria for the BSQ-16 (https://www.psyctc.org/psyctc/psyctc-org-home/instruments/body-shape-questionnaire-bsq), 40.0% of the women in our sample had "no concern" with their body shape, 15.4% "mild concern", 16.9% "moderate concern", and 27.7% "marked concern". Finally, using conventional cut-off scores for the BDI, 44.6% of women were "non-depressed", 21.5% were "mildly depressed", and 33.6% "moderately to severely depressed". Since the distribution of raw JND_{self} scores was not normal (Shapiro–Wilk's $W = .59$, $p < .0001$) we applied a logarithmic transformation to these data.

3.2. Multivariate Analysis

3.2.1. Replicating findings from three-quarter view presentations

In all our previous studies, we have used three-quarter view stimuli. Therefore, our first analysis was directed towards replicating previous findings with respect to the PSE and JND_{self} for self-estimates of body size using only the three-quarter view stimuli. Table 2 shows the correlation matrix between all variables. JND_{self} shows moderate correlations with participants actual BMI and their EDEQ scores.

PSE shows moderate to strong correlations between age, actual BMI, BSQ and EDEQ. Table 2 also shows moderate to strong correlations between all three psychometric tasks.

Table 2

Pearson correlations between participant characteristics, psychometric scores and psychophysical performance for three-quarter view stimuli

	Log_{10} JND_{self} 3Q	PSE 3Q	Age	BMI	BDI	BSQ-16
PSE 3Q	0.33*	-				
Age	0.036	0.45**	-			
BMI	0.34*	0.75***	0.22	-		
BDI	-0.098	0.12	0.068	-0.0014	-	
BSQ-16	0.25	0.54***	0.38**	0.40**	0.47**	-
EDE-Q	0.30*	0.54***	0.16	0.46**	0.39**	0.83***

NB JND = Just Noticeable Difference, PSE = Point of Subjective Equality, BMI = Body Mass Index, BDI = Beck Depression Inventory, BSQ = Body Shape Questionnaire, EDE-Q = Eating Disorders Examination Questionnaire, * = p < .05, ** = p < .01, *** = p < .001

We wanted to build separate multiple regression models to explain variance in JND_{self} and PSE, based on participants' age, actual BMI and psychometric performance. Given the substantial correlations between the psychometric tasks, we sought to include a selection procedure in the models which would avoid potential problems with multicolinearity. Since stepwise selection algorithms are known to lead to biases in parameter estimation (Grafen & Hails, 2002; Hurvich & Tsai 1990; Steyerberg et al. 1999), we used PROC GLMSELECT in SAS v9.4 (SAS Institute, North Carolina, USA) to run adaptive LASSO (least absolute shrinkage and selection operator) regression for variable selection (Efron, Hastie, Johnstone & Tibshirani, 2004; Osborne, Presnell & Turlach, 2000; Tibshirani, 1996). LASSO and stepwise regression differ in their criteria for retaining predictors in the final model, and LASSO has been shown to produce more stable results. The LASSO algorithm selects an optimal value for *t*, the tuning or shrinkage parameter which, in our case, minimized the Schwarz Bayesian information criterion (SBIC) for model fitting. We included chronological age, actual BMI, BDI, EDEQ, and BSQ as explanatory variables at the start of the selection procedure. By the end of selection, the optimal subsets of variables chosen to model JND_{self} and PSE had SBIC values of -175.3 and 117.6, respectively. We then used PROC REG in SAS (v9.4) to run ordinary least squares multiple regression models with these reduced sets of explanatory variables, derived from the LASSO process. The outcome from this analysis is shown in Table 3.

Table 3

Output from ordinary least squares multiple regression of $log_{10}JND_{self}$ and PSE obtained from three-quarter view stimulus presentations

Outcome variable	Explanatory variable	Parameter estimate	Standard error	t-value	p-value	Model R-square
Log_{10} JND_{self}	Intercept	-0.37	0.13	-2.79	.008	0.12
	Actual BMI	0.013	0.0054	2.51	.02	
PSE	Intercept	-0.42	2.89	-0.15	.9	0.68
	Actual BMI	0.75	0.12	6.18	<.0001	
	Age	0.23	0.068	3.36	.002	
	EDEQ	0.86	0.39	2.23	.03	

For JND_{self}, Table 3 shows a statistically significant, positive linear relationship with participants' actual BMI. Moreover, from the lowest (17.9) to the highest (35.9) actual BMI values, the Weber fraction was reasonably constant from 0.041 to 0.037. Both of these outcomes are consistent with Weber's law behaviour for this sample of 51 women. For PSE, Table 3 shows statistically significant, independent contributions from participants' actual BMI and their EDEQ scores. The slope of the regression line of estimated BMI on actual BMI was positive, i.e. participants estimated BMI increased as a function of their actual BMI. However, this relationship had a slope significantly less than 1 (F1,49 = 3.97, p =.05), consistent with the presence of perceptual contraction bias. Table 3 shows that the effect of EDEQ on body size self-estimates was also positive. Therefore, for any value of actual BMI, women increasingly over-estimated their own body size as their concerns about body shape, weight, and eating habits, increased. Both of these effects on PSE are illustrated in Figure 2a, which is a numerical simulation based on the regression parameters in Table 3.

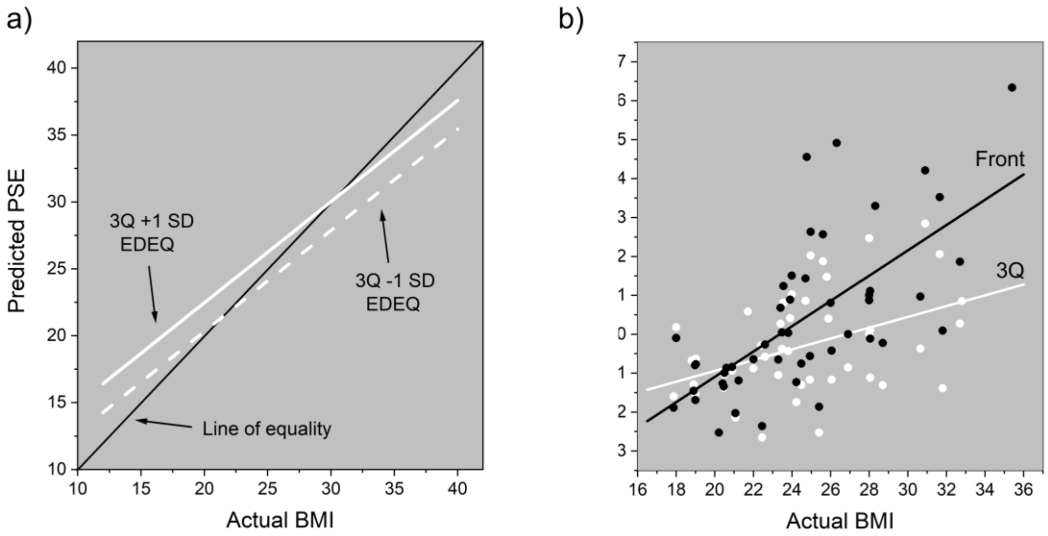

Figure 2. Scatterplots showing (a) PSE and (b) log_{10} JND_{self} predicted from the statistical models and plotted as a function of participants' actual BMI. The data in plot (a) is restricted to the three-quarter view stimuli. Predicted PSE is plotted as a function of participant BMI when calculated for mean EDEQ +1SD (solid white line) and mean EDEQ -1SD (dashed white line). The fact that the solid white line sits directly above the dashed white line, in parallel, means that for any actual BMI, participants over-estimate their body size by the same amount, as EDEQ increases from the mean -1 SD to the mean + 1SD. Veridical responses, where participants judge their body size perfectly accurately, is shown by solid black "the line of equality". The presence of contraction bias is indicated by the fact that both white regression lines have a slope less than one. Therefore, at low BMI, PSE values sit above the line of equality (i.e. participants over-estimate body size), and at high BMI, PSE values sit below the line of equality (i.e. participants under-estimate body size). In plot (b), the model predicted log_{10} JND_{self} values from responses to three-quarter view stimuli are shown as white dots with a white regression line. Those from front view stimuli are shown in black dots with a black regression line, which is significantly steeper than the white regression line.

3.2.2. Direct comparison of JND and PSE between three-quarter and front view avatars

To model the effects of avatar orientation on JND_{self} and PSE, we used PROC MIXED in SAS v9.4 (SAS Institute, North Carolina, USA) to run separate linear mixed effects models for each outcome measure. In each case the model included fixed effects for participants' actual BMI, stimulus orientation, and the interaction between actual BMI and stimulus orientation. We included a random effect to control the intercept for each participant. For dummy coding, front view (for stimulus orientation) was treated as the reference. To compute the denominator degrees of freedom, we specified the Satterthwaite method. The outcomes for these two models are shown in Table 4.

Table 4

Linear mixed effects model output for $log_{10}JND_{self}$ and PSE comparing responses obtained from three-quarter and front view avatar presentations

Outcome variable	Explanatory variable	Parameter estimate	Standard error	t-value (DF)	p-value	95% CI
$Log_{10}\,JND_{self}$	Intercept	-0.75	0.17	-4.41 (78.8)	<.0001	-1.09 – -0.41
	Actual BMI	0.033	0.0069	4.81 (79)	<.0001	.019 – 0.047
	3Q view	0.30	0.16	1.85 (43.8)	.07	-0.027 – 0.62
	Front view	.	.	.	.	.
	BMI × 3Q view	-0.015	0.0066	-2.34 (44.6)	.02	-0.029 – -0.0021
	BMI × Front view	.	.	.	.	.
PSE	Intercept	2.14	2.79	0.77 (57.4)	.4	-3.45 – 7.73
	Actual BMI	0.98	0.11	8.74 (57.1)	<.0001	0.75 – 1.20
	3Q view	-1.32	0.32	-4.15 (46)	.0001	-1.97 – -0.68
	Front view	.	.	.	.	.

For JND_{self}, Table 4 shows a statistically significant, positive linear relationship with participants' actual BMI. While the main effect of stimulus orientation was not statistically significant at $p < .05$, we found a significant interaction between orientation and participants' BMI, indicating a steeper slope for the front view stimuli than the three-quarter view. This is illustrated by the black regression line in Figure 2b for the front view stimuli, as compared to the white regression line for the three-quarter view stimuli. Moreover, for the front view stimuli, the Weber fraction more than doubled from 0.037 at the lowest actual BMI (17.9) to 0.084 at the highest (35.9) actual BMI. Therefore, participants' responses to the front view stimuli did not conform to Weber's law. For PSE, Table 4 shows statistically significant, independent contributions from participants' actual BMI and the orientation of the stimulus avatar. We did not find evidence for an interaction between participants' actual BMI and stimulus orientation. As can be seen from the parameter estimate for 3Q view in Table 4 (i.e., -1.32), the three-quarter view systematically gives rise to lower estimates of body size across the full BMI range. So, for a participant at any particular BMI, this model predicts that their self-estimate of body size will be, on average, 1.73 BMI units (*SE* = 0.98) higher when comparing their beliefs about their own body size to the front view avatar than the three-quarter view avatar. While this result shows clearly that body size estimates with the front view stimuli were systematically larger than those made with three-quarter view stimuli, it is important to distinguish between two possible interpretations. The first is that estimates made with front view stimuli are a closer fit to actual BMI, and the three-quarter view gives rise to under-estimated body size. The second possibility is that estimates made with three-quarter view stimuli are closest to actual BMI, and the front view stimuli produce over-estimates. To address this question, we calculated the difference between PSE and actual BMI, and calculated t-tests of location with respect to a difference of zero. The respective mean differences for three-quarter and front view, together with the t-tests were: $M = 0.31$, $SE = 0.48$, $t(50) = 0.65$, $p = .51$ and $M = 1.76$, $SE = 0.52$, $t(50) = 3.40$, $p = .001$. This suggests that front view presentations of avatar stimuli led to body size over-estimation by comparison to the three-quarter view presentations.

Finally, we addressed the question whether individual differences in women's attitudes about their bodies (indexed by the psychometric tasks: EDEQ, BSQ, BDI) may have contributed to the difference in self-estimates of body size (indexed by PSE) that we observed when comparing responses to front view versus three-quarter view stimuli. To do this, we calculated the difference in PSE between the two stimulus orientations for each participant, and calculated the Pearson correlations between these differences and psychometric task performance. The correlations with EDEQ, BSQ, and BDI were r = -0.12, -0.080, and -0.0071, respectively, none of which were statistically significant at $p < .05$.

4. Discussion

In this study, we asked what are the consequences of using stimuli with poor content validity, when healthy adult women estimate their own body size? Specifically, we wanted to know what happens to their just noticeable difference (JND_{self}), as well as their point of subjective equality (PSE), when the avatar stimulus they are comparing themselves to is presented in three–quarter view (i.e., good content validity) and front view (i.e., poor content validity). To address this, we asked participants to carry out a forced choice task in order for them to estimate their own body size. On each trial of the task, participants were presented a CGI image of a standard female model whose BMI varied across trials. Participants were asked to respond by button press whether they believed the image on each trial to be thinner or fatter than themselves. Each participant's responses were compiled into a psychometric function. From this, we calculated two measures: (a) their point of subjective equality (PSE) which corresponds to the body size they believe they have, and (b) their just noticeable difference (JND_{self}), which is an estimate of task precision and defines the smallest difference in body size between their own beliefs and the presented stimulus that they can just detect. Participants carried out this task twice, once with the CGI model facing forward, and once in a three-quarter view. Across the sample of women, we could then compare PSE and JND_{self} from the two stimulus orientations.

4.1 Replication of JND_{self} and PSE findings with three-quarter view avatars

Cornelissen et al. (2015, 2017) asked healthy adult women, who varied widely in actual BMI, to visually estimate their own body size using three-quarter view CGI avatars in the same forced choice task that we have used here. With respect to JND_{self}, both studies showed that women's responses with the three-quarter view stimuli conformed to Weber's law: individuals with larger actual BMIs showed a higher JND_{self}, and the Weber fraction remained constant across the range of BMI in these samples. With respect to PSE, Cornelissen et al. (2015, 2017, 2019) and Irvine et al. (2019a) have shown that women with a low BMI over-estimate their size and those with a high BMI under-estimate, a pattern which is consistent with a normal perceptual phenomenon called contraction bias (Poulton, 1989). Contraction bias arises when one uses a standard reference or template for a particular kind of object against which to estimate the size of other examples of that object. The estimate is most accurate when estimating the size of an object of a similar size to the reference but becomes increasingly inaccurate as the magnitude of the difference between the reference and the object increases. When this happens, the observer estimates that the object is closer in size to the reference than it actually is. As a result, an object smaller in size than the reference will be over-estimated and an object larger will be under-estimated.

In these same studies, Cornelissen et al. (2015, 2017, 2019) and Irvine et al. (2019a) found that women's estimates of their own body size, indexed by PSE, also depended on attitudinal factors indexed by psychometric tasks measuring attitudes towards body shape, body size and eating habits, as well as the tendency towards depression. Broadly speaking, for any actual BMI, as a woman's psychological concerns about her own body increase, so does the size she thinks she is. Put together, this means that a complete description of an individual's performance in a forced choice body size estimation task requires information about two independent sources of information: (a) actual body size, which is subject to a perceptual contraction bias, and (b) an attitudinal component, whereby increasing psychological concern leads to larger body size estimates.

As is clear from Figure 2b (white regression line) and Table 3, in the current study we have replicated this pattern of results for JND_{self}, which therefore conforms to Weber's law. We found a positive linear slope (β = 0.013) for the regression of Log_{10} JND_{self} on actual BMI, and the Weber fraction remained constant at ~ 0.04 across the sample range of BMI. Similarly, as is illustrated by Figure 2a, we replicated the contraction bias effect for PSE coupled with the independent contribution of EDEQ scores to the body size that women believed they have.

4.2 The consequence of poor stimulus content validity

Cornelissen et al. (2018) showed that, at least for the CGI avatars they tested, images of women in front view showed poor content validity. In the current study, the regression of JND_{self} on actual BMI, calculated from women's responses to such avatars presented in front-view, showed a positive slope that was significantly steeper than that for three-quarter view stimuli (see Figure 2b). This produced Weber fractions for the responses to front view stimuli that more than doubled over the full range of participants' actual BMI. Therefore, unlike the three-quarter view stimuli, we conclude that the poor content validity of the front view stimuli prevented women's body size estimates from conforming to Weber's law.

Perhaps the most striking outcome from the current study is that the front view stimuli also gave rise to self-estimates of body size that were ~1.7 BMI units greater than those derived from three-quarter view stimuli, and this was true across the full range of actual BMI from ~17 to ~36. Not only was this viewpoint dependent difference in body size estimates statistically significant, but it is likely to represent a clinically meaningful effect size (cf. Kazis, Anderson, & Meenan, 1989), in the sense that it is easily detectable perceptually. Cornelissen et al. (2018) showed that the smallest difference in BMI that can just be detected between pairs of CGI avatars, both presented simultaneously on screen (i.e., JND_{stim}), increased systematically from ~0.6 to ~1.0 BMI units over the stimulus BMI range 15-36. Therefore, the ~1.7 BMI unit difference in PSE estimates obtained from front and three-quarter view stimulus presentations are ~180% and ~70% larger, respectively, than these smallest detectable differences. In other words, the differences in body size estimates between three-quarter and front view are not subtle, but manifestly easy to see. Put together, our results with respect to both JND_{self} and PSE in the current study strongly suggest that using front view CGI stimuli for body size self-estimation tasks will lead to imprecise, over-estimates.

4.3. When might these viewpoint dependent differences really matter?

If this front view stimulus set were to be used in the future, to compare body size self-estimates between two or more groups, then this systematic error may not be critically important because one would most likely be seeking evidence of *relative* differences between the groups. Since the difference between body size estimates between the front and three-quarter view is a constant ~1.7 BMI units across the stimulus range, it is likely that this error would apply across the board and could therefore be ignored. However, this would *not* be the case, if one wanted to use the front view stimuli to obtain *absolute* estimates of body size. For example, there would be significant implications for epidemiological studies of obesity rates that use figural scales, like those reported by Dratva et al. (2016). In such studies, where it is not logistically feasible to measure the heights and weights of thousands of people in order to calculate BMI, participants are asked instead to identify which image on a scale is closest to the body size they think they have. Therefore, in the presence of a systematic ~1.7 BMI unit offset, individuals who have an actual BMI of ~28.3 would self-report a BMI of 30, hence obesity rates would be over-estimated using such a scale.

4.4. What causes these viewpoint dependent differences?

In a recent study using the bubbles masking technique, Irvine et al. (2019b) showed that information extracted primarily from the left and right torso edges is largely responsible for driving decisions about self-estimates of body size. The strong implication of this finding is that participants' visual systems are extracting information which relates to the width of the torso, i.e. the separation between the left and right torso edges.

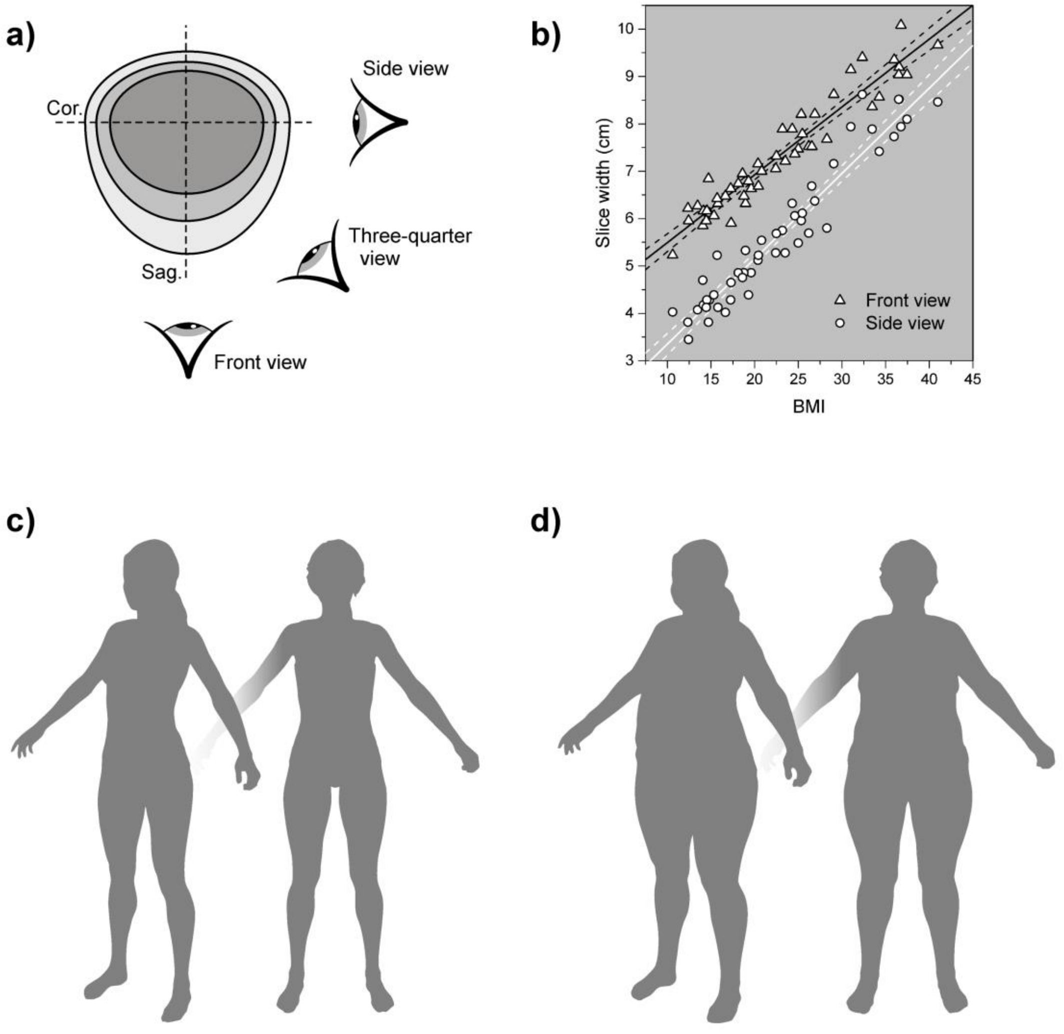

Figure 3 (a) Illustration of abdominal cross-section with progressively increasing BMI. It shows how width increases in the sagittal (Sag.) plane more quickly than in the coronal (Cor.) plane, and how this is harder to see in front view than either the side of three-quarter view. (b) Plots of waist width seen from front (triangles) and side (circles) views from 50 photographs of women in Tovée & Cornelissen, 2001. The black and white lines represent the ordinary least squares regression lines through the respective data together with their 95% confidence intervals. The images in (c) and (d) are silhouettes *of three-quarter (left) and front (right) view stimuli at BMIs of ~20 and ~36 respectively. The left edge of the three-quarter view captures the component of the anterior-to-posterior (sagittal plane) bulge with increasing BMI much more effectively than the front view.*

Figure 3a is a sketch of a set of cross-sections through the abdomen. It shows how the anterior to posterior width in the central abdomen (i.e., anatomically the sagittal plane) increases more rapidly than the corresponding width from the left to the right torso edge (i.e., anatomically the coronal plane). Moreover, viewing a body from the front gives access primarily to changes in the coronal plane (left to right torso edge), whereas a side view primarily accesses changes in the sagittal plane (anterior to posterior). The three-quarter view however gives access to both. Figure 3b shows plots of waist width seen from the front view (coronal plane; white triangles with black regression line) and the side view (sagittal plane; white circles with white regression line) as a function of BMI, extracted from 50 photographs of women in Tovée & Cornelissen (2001). This confirms that the rate of change of waist width is greater (i.e. there is a steeper regression slope) for the side view than the front view as BMI increases. Put together, this evidence suggests that the three-quarter view may well carry more information about adiposity, and changes to adiposity, than the front view. Consistent with this proposal, Figure 3c and 3d show silhouettes of the three-quarter and front view stimuli at BMIs ~20 and ~36, respectively. Qualitatively, it can be seen on inspection that a component of the anterior-posterior bulge (sagittal plane) is much more easily seen along the left torso edge in three-quarter view than front view.

While the foregoing discussion illuminates why the front view may carry less precise information about body size and adiposity, compared to the three-quarter view, it does not explain why having less information from a front view might lead participants to over-estimate their body size. We suggest that an additional line of argument is required to account for this directionality. Specifically, there is a wealth of evidence that recognition and/or discrimination of a novel view of an object is achieved by comparing it to a stored prototype, referred to as a canonical view (Edelman & Duvdevani-Bar, 1997; Niimi & Yokosawa, 2009; Palmer, Rosch, & Chase, 1981; Ullman, 1996). Stimulus viewpoints similar to, or the same as, the internal representation or representations give rise to the best task performance. Therefore, if it is true that the three-quarter view of our body stimuli is the most closely aligned to the canonical view of bodies that participants hold in memory, this may contribute to their more accurate performance with the three-quarter view stimuli. Over-estimation with the front view stimuli may arise as a compensatory mechanism, whereby participants are having to choose a more 'inflated' front view stimulus (i.e., one with a higher BMI) to best match some key component of the canonical representation. It is also conceivable that other cognitive mechanisms may cause this over-estimation. Clearly, these ideas are highly speculative, and require further research to test.

4.5. Limitations

This study has used CGI stimuli for three reasons: (a) the images are high resolution and photorealistic, (b) the identity of the model in the stimulus can easily be held constant over a wide range of BMIs, and (c) the stimuli can be calibrated for BMI based on their waist and hip circumferences. The calibration equations we used are derived from a multiple regression analysis of over 5000 individuals in the HSE (2003, 2008) datasets, and account for ~90% of the variability in BMI of UK citizens (Cornelissen, 2016b). Moreover, since the error residuals in these analyses were normally distributed, it is unlikely that there are systematic, directional calibration errors within the equations themselves. However, CGI model creation involves a degree of artistic judgement, and so it is possible that systematic biases exist in waist and hip circumferences that may have been introduced by the CGI modeller, giving rise to errors in apparent BMI. We are aware of only two ways to resolve such potential difficulties in the future. The first is always to use photographs of real people, with known BMIs, as stimuli. The problem here is that for a typical psychophysical paradigm we would have to run considerably more trials in order to offset the marked shape variation between real people who have the same BMIs, and this may not be logistically feasible. An alternative approach is to generate stimuli which are modelled from large datasets of individuals in whom accurate body composition measurements have been made coupled with high resolution 3D body shape scans (cf. Groves et al., 2019). This approach would allow an accurate statistical mapping to be made between body composition and body shape, leading to the production of correctly calibrated stimuli (cf. Maalin, Mohamed, Kramer, Cornelissen, & Tovée, 2020).

In conclusion, this follow-up study to Cornelissen at al. (2018) has shown that, at least for these particular CGI stimuli, self-estimates of body size made with front facing stimuli lead to less precise judgements, and over-estimates of body size around ~1.7 BMI units, compared to the same estimates made with three-quarter view stimuli. Not only do these results need to be replicated, but we also need to know whether they generalize to other classes of stimuli, such as line drawings, photographs, and personalized 3D avatars based on 3D body shape scans. Nevertheless, they do suggest that careful consideration needs to be given to stimulus orientation in future studies.

Acknowledgements: LB was funded by Northumbria University as an RA.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author. DOI: 10.1176/appi.books.9780890425596

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., (1961). An inventory for measuring depression. *Arch. Gen. Psychiatry*, 4, 561–571. DOI: 10.1001/archpsyc.1961.01710120031004.

Biener, A., Cawley, J., & Meyerhoefer, C. (2018). The impact of obesity on medical care costs and labor market outcomes in the US. *Clinical Chemistry*, 64(1), 108-117. DOI: 10.1373/clinchem.2017.272450

Billock, V. A., & Tsou, B. H. (2011). To honor fechner and obey stevens: relationships between psychophysical and neural nonlinearities. *Psychological Bulletin*, 137(1), 1–18. http://dx.doi.

org/10.1037/a0021394.

Boring, E. G. (1950). *A History of Experimental Psychology* (2nd ed.). New York, USA: Appleton-Century-Crofts.

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., & Brown, T. A. (1987). Body image in anorexia nervosa and bulimia nervosa: A review of the literature. *Behavior Modification,* 11(4), 487-521. DOI: 10.1177/01454455870114005

Cash, T.F., & Deagle, E.A. (1997). The nature and extent of body-image disturbances in anorexia nervosa and bulimia nervosa: A meta-analysis. *International Journal of Eating Disorders*, 22,107-126. DOI: 10.1002/(SICI)1098-108X(199709)22:2<107::AID-EAT1>3.0.CO;2-J

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, 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: 10.1016/j.bodyim.2015.01.001

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

Cornelissen, K. (2016b). *What does it mean to have distorted body image in anorexia nervosa*? (Doctoral thesis). Retrieved from http://nrl.northumbria.ac.uk/30330/

Cornelissen, K. K., McCarty, K., Cornelissen, P. L., & Tovée, M. J. (2017). Body size estimation in women with anorexia nervosa and healthy controls using 3D avatars. *Scientific Reports*, 7, 15773. DOI: 10.1038/s41598-017-15339-z.

Cornelissen, P. L., Cornelissen, K. K., Groves, V., McCarty, K., & Tovée, M. J. (2018). View-dependent accuracy in body mass judgements of female bodies. *Body Image*, 24, 116-123. DOI: 10.1016/j.bodyim.2017.12.007

Cornelissen, K. K., Widrington, H., McCarty, Pollett, T. V., Tovée, M. J., & Cornelissen, P. L. (2019). Are attitudinal and perceptual body image the same or different? Evidence from high-level adaptation. *Body Image*, 31, 35-47. DOI: 10.1016/j.bodyim.2019.08.001

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.

Efron, B., Hastie, T. J., Johnstone, I. M., & Tibshirani, R. (2004). Least Angle Regression. *Annals of Statistics*, 32, 407–499.

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<315::aid-eat2260130310>3.0.co;2-3

Fairburn, C.G., & Beglin, S.J., (1994). Assessment of eating disorders: interview or self-report questionnaire? *Eating Disorders,* 16(4), 363–370. DOI: 10.1002/1098-108X

(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-#.

Fairburn, C., Cooper, Z., & O'Connor, M. (2008). *Cognitive behavior therapy and eating disorders*. Guilford Press, New York

Gardner, R. M. (1996). Methodological issues in assessment of the perceptual component of body image disturbance. *British Journal of Psychology*, 87, 327–337.

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

Gardner, R. M., & Brown, D. L. (2010). Body image assessment: A review of figural drawing scales. *Personality and Individual Differences,* 48 (2), 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

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

Grafen, A., & Hails, R. (2002). *Modern Statistics for the Life Sciences*. Oxford University Press, Oxford.

Groves, V., Cornelissen, P. L., McCarty, K., Mohamed, S., Maalin, N., Tovée, M. J., & Cornelissen, K. K. (2019). How inaccurate are self-estimates of male body size? *Frontiers in Psychiatry*. DOI: 10.3389/fpsyt.2019.00720

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

Health Survey for England (2003). National Centre for Social Research and University College London. Department of Epidemiology and Public Health. UK Data Archive, Colchester, Essex. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjc1OPRxsTjAhXSKM0KHb3lApUQFjAAegQIARAC&url=https%3A%2F%2Fsp.ukdataservice.ac.uk%2Fdoc%2F5098%2Fmrdoc%2Fpdf%2F5098userguide.pdf&usg=AOvVaw1M0gC8lXbXE7tWobAF2Eqa

Health Survey for England (2008). National Centre for Social Research and University College London. Department of Epidemiology and Public Health. UK Data Archive, Colchester, Essex, UK. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwjH9IWrx8TjAhWUdc0KHSzSDdIQFjAAegQIARAB&url=https%3A%2F%2Fdigital.nhs.uk%2Fdata-and-information%2Fpublications%2Fstatistical%2Fhealth-survey-for-england%2Fhealth-survey-for-england-2008-physical-activity-and-fitness&usg=AOvVaw0dAW9_0fZ5tvr57Of_gLxI

Hurvich, C.M., & Tsai, C.L. (1990). The impact of model selection on inference in linear regression. *American Statistician*, 44, 214– 217.

Irvine, K. R., McCarty, K., McKenzie, K., Pollet, T.V., Cornelissen, K. K., Tovée, M. J., & Cornelissen, P. L. (2019a). Distorted body image influences body schema in individuals with negative bodily attitudes. *Neuropsychologia*, 122, 38-50. DOI: 10.1016/j.neuropsychologia.2018.11.015

Irvine, K. R., McCarty, K., Pollet, T. V., Cornelissen, K. K., Toveé, M. J., & Cornelissen, P. L. (2019b). The visual cues that drive the self-assessment of body .size: dissociation between fixation patterns and the key areas of the body for accurate judgement. *Body Image*, 29(c), 31-46. DOI: 10.1016/j.bodyim.2019.02.006

Kazis, L.E., Anderson, J.J., & Meenan, R.F. (1989). Effect sizes for interpreting changes in health status. *Medical Care, 27*, S178-S189.

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

Krech, D., & Crutchfield, R. S. (1958). *Elements of psychology*. New York, USA: Knopf.

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

Maalin, N., Mohamed, S., Kramer, R. S. S. et al. (2020). Beyond BMI for self-estimates of body size and shape: A new method for developing stimuli correctly calibrated for body composition. *Behavioural Research Methods*. DOI: 10.3758/s13428-020-01494-1

Mölbert, S. C., Thaler, A., Mohler, B. J., Streuber, S., Romero, J., Black, M. J., . . . Giel, K. E. (2018). Assessing body image in anorexia nervosa using biometric self-avatars in virtual reality: attitudinal components rather than visual body size estimation are distorted. *Psychological Medicine, 48*, 642–653. DOI: 10.1017/S0033291717002008

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

Murnen, S. K., & Smolak, L. (2019). The Cash effect: Shaping the research conversation on body image and eating disorders. *Body Image*, 31, 288-293. DOI: 10.1016/j.bodyim.2019.01.001

Musich, S., MacLeod, S., Bhattarai, G. R., Wang, S. S., Hawkins, K., Bottone Jr, F. G., & Yeh, C. S. (2016). The impact of obesity on health care utilization and expenditures in a medicare supplement population. *Gerontology and Geriatric Medicine*, 2, 2333721415622004. DOI: 10.1177/2333721415622004

Niimi, R., & Yokosawa, Z. (2009). Three-quarter views are subjectively good because object orientation is uncertain. *Psychonomic Bulletin & Review*, 16(2), 289-294. DOI: 10.3758/PBR.16.2.289

Osborne, M. R., Presnell, B., & Turlach, B. A. (2000). A New Approach to Variable Selection in Least Squares Problems. *IMA Journal of Numerical Analysis*, 20, 389–404.

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

Poulton, E. C. (1989). *Bias in quantifying judgements*. Hove, UK: Erlbaum.

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 10.1016/j.evolhumbehav.2008.08.007

Seidell, J. C. (1995). The impact of obesity on health status: some implications for health care costs. International journal of obesity and related metabolic disorders. *Journal of the International Association for the Study of Obesity*, 19, S13-6.

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

Steyerberg, E.W., Eijkemans, M.J.C. & Habbema, J.D.F. (1999). Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. *Journal of Clinical Epidemiology*, 52, 935– 942.

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., 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(8), 1752-1761. DOI: 10.1016/j.paid.2008.02.002

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

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

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B*, 58, 267–288.

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

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

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

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