Visual biases in judging body weight

Katri K. Cornelissen[1], Lucinda J. Gledhill[2], Piers L. Cornelissen[1] and Martin J. Tovée[2]

[1] Department of Psychology, Northumbria University, UK
[2] Institute of Neuroscience, Newcastle University, UK.

*Requests for reprints should be addressed to Dr. Tovée, Institute of Neuroscience, Faculty of Medical Sciences, Newcastle University, NE2 4HH, UK. (e-mail: m.j.tovee@ncl.ac.uk).

Abstract

Objectives: There has been a steady rise in obesity levels in Western countries and a contributory factor is people's failure to recognise weight gain. Two important visual perceptual biases, which have hitherto been ignored in the obesity literature, could contribute to this problem; contraction bias and Weber's law. Contraction bias predicts that the weight of obese bodies will be under-estimated and the degree of underestimation will increase as body mass index (BMI) increases. Weber's law predicts that change in the body size will become progressively harder to detect as their BMI increases.

Methods: In Experiment 1, twenty-nine women participants estimated the weight of 120 women varying in their body mass. In Experiment 2, twenty-eight women participants judged which body was the heavier in a 2-alternative forced choice paradigm.

Results: In Experiment 1, as predicted the participants showed a progressive under-estimation of over-weight and obese bodies, $\beta_1 = 0.71$, $t = 26.96$, $p < .0001$. For Experiment 2, there was a significant effect of the BMI of the bodies being judged on the just noticeable difference needed to discriminate between them: $F(1,196) = 89.39$, $p < .0001$ for 3D bodies and $F(1,86.5) = 44.57$, $p < .0001$ for digital photographs.

Conclusions: Normal visual perceptual biases influence our ability to determine body size: contraction bias and Weber's law mean that as bodies become overweight and obese it is harder to judge their weight and detect any increase in size. These effects may therefore compromise people's ability to recognise weight gain and undertake compensatory weight control behaviours.

Key words: Obesity, contraction bias, Weber's Law, JND, body image.

Introduction

In Western countries there has been an inexorable rise in obesity levels with a concomitant pressure on public health resources (Ogden et al., 2006; Swinburn et al., 2011). A recent report from the McKinsey Global Institute put the costs of obesity to the world economy at £1.3 trillion, and the cost to the UK at £47 billion (McKinsey-Global Institute, 2014). Obesity can take up to 8 years off a person's life expectancy and cause decades of ill-health (Grover et al., 2014). A potential contributory factor to the rise in obesity is the failure of people to recognise weight gain. The Health Belief Model (HBM) postulates six stages in the development of changes in behaviour, the first of which is perceived susceptibility which in this context corresponds to the perception of weight and weight gain (Rosenstock, Strecher & Becker, 1988). If overweight individuals cannot accurately judge their weight and are unable to detect weight gain, the subsequent stages of HBM will not occur and their behaviour will remain unchanged. Previous studies suggest that over-weight and obese people seem to under-estimate their size and weight, and may not detect weight gain (Kuchler & Variyam, 2003; Kuskowska-Wolk & Rössner, 1989; Maximova et al., 2008; Oldham & Robinson, 2015; Robinson & Kirkham, 2013; Rahman & Berenson, 2012; Truesdale & Stevens, 2008; Wetmore & Modkdad, 2012).

Of potentially equal importance is the ability to detect obesity and weight change in others. For example, healthcare professionals are advised to screen and offer weight control help to overweight and obese patients, but this rarely happens despite the fact that most people will see their GP at least once a year (NICE, 2006; NHS England, 2015). This maybe because GPs under-estimate the body mass index (BMI) of over-weight and obese patients, and the decision to initiate a weight control discussion is based on these impaired visual judgements (Robinson, Parretti & Aveyard, 2014). Another important group of people who

3

need to make accurate judgements about body size are the parents of school-aged children. There is a high level of child obesity in Western populations and it is important that their parents initiate changes in their child's diet and behaviours (Jackson-Leach & Lobstein, 2006; Olds et al., 2011). However, a significant proportion of parents under-estimate their child's BMI and fail to undertake weight control measures (Jones et al., 2011; Duncan, 2011). Thus, being able to accurately judge other people's BMI is extremely important.

Previous studies suggest that we make judgements about complex stimuli, such as bodies, by reference to a template based on the average of all that class of stimuli that we have seen - our visual diet (Leopold, O'Toole, Vetter & Blanz, 2001; Winkler & Rhodes, 2005). So if someone has seen many high BMI bodies then their internal reference (what they perceive as a normal, representative body size) will be shifted towards a heavier body size. So an individual's failure to recognize their own obesity may be because they compare themselves to the newly adapted reference point, which is much closer to their current body size, as opposed to an absolute comparison point which is not subject to adaptation (Robinson & Kirkham, 2013; Oldham & Robinson, 2015). However, this explanation does not consider two naturally occurring visual perceptual biases in magnitude estimation that could also contribute to the under-estimation of body size: contraction bias and Weber's law.

According to Poulton (1989), "the contraction biases can affect any kind of quantitative judgement or rating. In the absolute version of the stimulus contraction bias, magnitudes larger than the observer's reference magnitude are underestimated. Magnitudes smaller than the observer' reference magnitude are over-estimated" (p. 155). As mentioned above, there is evidence that humans learn a reference body size / body weight based on an average of all the bodies they have seen over the course of their life with a weighting towards more recent images (e.g., Winkler & Rhodes, 2005; Rhodes, Jeffery, Boeing & Calder, 2013). If so, contraction bias predicts that body size judgements should be most accurate

when comparing a body similar in size to the reference, and increasingly less accurate as the two diverge. Thus bodies much larger than the reference should be underestimated and bodies smaller than the reference should be overestimated. Quantitatively, sufficient evidence to suggest the presence of contraction bias can be derived from a plot of estimated weight (y-axis) as a function of stimulus weight (x-axis). We should observe: (i) a slope for the regression of estimated weight on actual weight which is statistically significantly less than 1 – i.e. the less than the slope of the line of equivalence where estimated weight agrees perfectly with actual weight; (ii) that this regression line crosses the line of equivalence approximately at the mean weight of the reference population. In the current study, Experiment 1 tests this prediction for White adult women in the UK.

The foregoing focuses on the problem of observers judging the body size of others, so what about judgements of self? Clearly, it is not straightforward to carry out an experiment in the way that contraction bias is usually conceived, because this would require an individual to make multiple judgements about themselves at a variety of different body weights (presumably) over time. While logically possible, this is impractical. The alternative is to ask many observers of varying body weights, to make judgements about themselves at one point in time. We then can ask whether the pattern of estimated weights across such a sample of observers fulfils the criteria above, and is therefore consistent with contraction bias. In a recent study, 100 non-eating disordered adult women estimated their own body size using the method of constant limits (Cornelissen et al., 2015). The accuracy of their estimates was indeed consistent with a contraction bias explanation.

While contraction bias could explain the under-estimation of over-weight and obese bodies, there is another perceptual phenomenon described by Weber's law, which means it also gets progressively *more* difficult to detect an increase in body weight *as we put weight on*. Weber's law states that the just noticeable difference (JND) between two stimuli will be a

constant proportion of their magnitude, leading to a constant Weber fraction over the stimulus range (i.e. $\Delta I / I = K$, where I = stimulus magnitude and K = constant) (Gescheider, 1997). This means that it is easier to notice, for example, a one BMI unit difference between two low BMI bodies than between two high BMI bodies. Over the full range of BMI, discriminating between higher BMI bodies requires progressively larger differences in BMI between stimuli. This means that as we get heavier it gets progressively harder to detect an increase in body mass.

To test our hypotheses, we carried out two experiments to test whether contraction bias and Weber's law do indeed apply to the perception of human body size, as we would expect, and thereby justify their inclusion in the debate about obesity. In the first experiment, participants estimated the weight of a set of bodies varying in weight and in the second, the participants had to discriminate between pairs of bodies of different weights. In the first experiment, contraction bias predicts that the participants will over-estimate low weight bodies and under-estimate high weight bodies. In the second experiment, Weber's law predicts the task will become progressively more difficult as the weight of the pairs of bodies increases.

Experiment 1: contraction bias

Methods

The experimental procedures and methods for participant recruitment for this study were approved by the local university ethics committee. Informed consent was obtained from all participants.

Participants

All participants for this study were recruited through the use of opportunity sampling from staff and students at [Blinded University] and [Blinded University] in the UK. For the pilot study for Experiment 1, we recruited 5 female participants from this population. Pilot testing showed that a sample of 22 participants would be sufficient to demonstrate the contraction bias effect with a power of 0.9, at an alpha level of .01 (see Results). For the full version of Experiment 1, to offset attrition in participant numbers and/or unexpected sources of variability, we recruited 29 female participants (age $M = 30.86$, $SD = 9.30$). None of the participants who took part in the pilot study also took part in Experiment 1. During recruitment we asked all potential participants whether they had a current diagnosis of an eating disorder or any history of such a disorder and excluded those individuals from this study.

Stimuli

One hundred and twenty digital photographs of female bodies were used as stimuli in this study. They were selected from the database of images reported in Tovée, Maisey, Emery and Cornelissen (1999). To generate the images, consenting women were photographed standing in a set pose, front-facing, against a fixed dark background, at a standard distance

from the camera, wearing tight grey leotards and leggings so that adiposity could easily be assessed by the observers. The images were stored as 24-bit colour pictures. The faces of the women in the images were blurred to protect anonymity. The women in the set of images for the current study varied in: weight from 28.2kg to 104.9 kg ($M = 59.14$, $SD = 13.44$); height from 1.47m to 1.83m ($M = 1.66$, $SD = 0.07$); BMI from 11.5 to 41.1 ($M = 21.61$, $SD = 5.75$).

Procedure

Before testing began, participants were informed that they were going to be shown a series of bodies varying in adiposity which they were required to estimate for weight using either kilograms or stones. They were free to use either scale depending on which one they were most comfortable with. Participants were then presented all 120 stimulus images, one at a time and in a randomized order, on a 19" flat panel LCD screen (1280w x 1024h pixel native resolution, 32-bit colour depth). Each image appeared on a plain black background beneath which was a linear scale ranging from ~25kg-115kg (first mark 30kg, last mark 110 kg, 10kg increment) above the line and ~4st-18st (first mark 5st, last mark 17st, 1st increment) below the line. Each image appeared on a plain black background beneath which was a linear scale ranging from 30kg- 110 kg above the line and 5st- 17st below the line. On each trial, with no time limitations, participants had to move a slider along the scale, and click with a mouse button when the participant judged that the slider had reached a weight corresponding to the weight of the woman in the image.

Results

Experiment 1 pilot data

From a statistical point of view, necessary evidence for contraction bias is a slope for the regression of estimated weight on actual weight that is statistically significantly less than 1. Therefore, we estimated this regression for each of the participants in the pilot study separately, in order to obtain an estimate of the mean beta coefficient for estimated weight across the five participants ($M = 0.77$) as well as its standard deviation ($SD = 0.24$). We then used G*Power v3.1.9.2 to compute that 22 participants would be required to demonstrate a statistically significant t-test of the difference of this mean from 1, at $\alpha = .01$ and power (1-β) 0.9.

Experiment 1 main data

In Figure 1 the dotted line represents veridical performance, if observers estimated body weight with perfect accuracy. However, Figure 1 shows clearly that there is a systematic departure from veridical performance. Images of women weighing ~70kg, i.e. the population average for White women in the UK (Health Survey for England, 2012), are estimated the most accurately. Thereafter, as the weight of women in the stimuli decreases, so participants systematically over-estimate their weight. Above the population average, observers systematically under-estimate body weight as the weight of the women in the stimuli increases.

***** Figure 1 about here ******

We used PROC REG in SAS v9.3 to compute an ordinary least squares regression of mean estimated weight for each image as a function of actual weight. The assumptions for linear regression were met, and this model explained 85% of the variance in estimated weight. The overall model fit was statistically significant, $F(1,118) = 726.88$, $p < .0001$. (NB

9

all p-values reported henceforth are for two-sided tests). The regression parameters, $\beta_0 =$ 19.60, $t = 11.99$, $p < .0001$, CI [16.37 - 22.84] and $\beta_1 = 0.71$, $t = 26.96$, $p < .0001$, CI [0.66 - 0.76], showed a statistically significant, positive linear relationship between stimulus weight and estimated weight. However, the slope of this relationship was significantly less than 1, $F(1,118) = 124.18$, $p < .0001$. Because we had also measured the height of the women in the stimulus images, we could carry out an equivalent analysis of estimated BMI as a function of actual BMI. This model explained 82% of the variance in estimated BMI. The overall model fit was statistically significant, $F(1,118) = 552.51$, $p < .0001$. The regression parameters, $\beta_0 =$ 5.39, $t = 7.01$, $p < .0001$, CI [3.87 – 6.91] and $\beta_1 = 0.79$, $t = 23.51$, $p < .0001$, CI [0.73 - 0.86], showed a statistically significant, positive linear relationship between stimulus weight and estimated weight. The slope of this relationship was also significantly less than 1, $F(1,118) = 36.93$, $p < .0001$. These data therefore demonstrate convincing evidence for contraction bias when women judge the body weight of other women.

Experiment 2: just noticeable difference

Methods

Participants

   The first part of Experiment 2 uses a mixed-design with one between-groups factor (CGI model: 2 levels) and one within participants factor (BMI: 8 levels). Therefore, we recruited 4 female participants to pilot this experiment, and we assigned two participants to each group. None of the participants who took part in this pilot study also took part in Experiment 2.

Pilot testing showed that a sample of 24 participants would be sufficient to quantify the relationship between JND and stimulus BMI, for the two different CGI models, with a power of 0.9, at an alpha level of 0.01. To offset attrition in participant numbers and/or unexpected sources of variability, we recruited 28 female participants (age $M = 31.81$, $SD = 7.83$) for this study from staff and students at [Blinded University] and [Blinded University] in the UK. As in Experiment 1, we asked all potential participants whether they had a current diagnosis of an eating disorder or any history of such a disorder and excluded those individuals from this study.

Stimuli

Experiment 2 comprised two parts. In the first part, we wanted to identify the smallest change in BMI that observers could detect, i.e. the just noticeable difference (JND), at eight separate points along the BMI continuum from ~12 to ~45. The 8 points correspond to the boundaries between BMI categories as well as points within each category. To create stimulus images which correctly represent how an individual body shape changes as a function of changing BMI is difficult, because these changes are highly non-linear (Crossley, Cornelissen & Tovée, 2012; Wells, Treleaven & Cole, 2007). One method that has been used previously is the video-distortion technique (VDT) (Gardner & Bokenkamp, 1996; Probst, Vandereycken & Van Coppenolle, 1997) in which 2D images of people are stretched or compressed in the horizontal dimension. However, this linear method is problematic as it creates shape changes particularly in the shoulder and hip regions which tend to be unrealistic (Cornelissen Bester, Cairns, Tovée & Cornelissen, 2015). Alternative methodologies can be used, such as morphing between images of high and low BMI bodies. While this is an improvement on the VDT in principle, because it is a non-linear method, it is nevertheless extremely difficult to maintain the combination of high body feature definition and stable

identity of the person in the morphed images across a wide range of BMI values. Inevitably, some form of averaging or smoothing is required which reduces the realism in the resultant images. For these reasons, we instead used film industry computer-generated imagery (CGI) methods to create graded 3D images of two individuals where: i) the identity of the person in the image is clearly maintained over a wide BMI range; ii) the body shape changes at different BMI levels are extremely realistic and iii) the 3D rendered stimulus images are high definition and photorealistic. In addition, we made precise estimates of the BMI of the 3D models in a particular image. 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 ~3000 women in the UK, aged between 18 and 45. Because our CGI models exist in an appropriately scaled 3D world, having set the height of our models (1.6m) we can therefore measure their waist and hip circumferences, and compare these with our HSE calibration curves in order to compute their BMI (Cornelissen, Tovée & Bateson, 2009).

In the second part of Experiment 2 we identified the smallest change in BMI that observers could detect (i.e. the JND) at BMI values around 15, 18, 21 and 24, using the same database of images of real women that we used in Experiment 1. Clearly, by using natural images, there is a much wider variety of shape and height, and we wanted to be sure that effects from the CGI stimuli could be replicated in images of real women.

Procedure

For both parts of Experiment 2, in order to measure observers' JNDs, we used a 2-alternative forced choice (2AFC) 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 thinner body. In the first part of Experiment 2, where we used the CGI stimuli, we presented observers 8 blocks of stimuli corresponding to the 8 points along the BMI continuum. Within each block, we presented pairs of images at each of 10 levels of BMI difference between the left and the right images. The set of differences in BMI between the image pairs was 0 to +/- 2.5 BMI units in 0.25 BMI steps. The stimulus image pairs were therefore drawn from the 8 BMI ranges: 14.5-18.5; 16.5-20.5; 20-24; 23-27; 24.5-30.5; 27-33; 32-38; 37-43. 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.

In the second part of Experiment 2, we used the same database of images as was used in Experiment 1. We used a similar logic for the experimental procedure, but this time drawing images from only four BMI ranges; 15.0-17.5; 18-20.5; 21-23.5; 24-26.5. All four image sets covered a BMI range of 0 to 2.5 BMI units in 0.5 BMI steps. Clearly this represents a coarser grained procedure than was the case for the CGI images, but it was necessary because of the wide natural variation in shape and BMI in real images of women.

For both the CGI and real image experiments, we randomized the order in which stimuli within a given BMI block were presented to participants, as well as the order of presentation of the BMI ranges themselves. For each participant, we used probit analysis to fit psychometric functions and we defined the JND as the BMI difference between image pairs at which observers correctly identified the larger body 75% of the time. These values were compared across participants, as a function of BMI, to test for Weber's law behaviour.

Results

Experiment 2 pilot data

We wanted to estimate sample size from an F test derived from the repeated measures, within-between interaction analysis of variance (ANOVA) option in G*Power v3.1.9.2. Therefore, we used PROC GLM in SAS v9.3 to compute a mixed-design ANOVA of the pilot participants' JNDs. The Type III test of the fixed effect of BMI was statistically significant, $F(7,16) = 8.65$, $p < .0005$ (p < .05 with a Greenhouse-Geisser correction of ε = 0.18). We then used the variance explained by BMI and the error variance to compute an effect size according to Cohen (1988) of $f(V) = 2.08$. At α = .05 and a power (1- β) of 0.9, G*Power returned a sample size of n = 24 for Experiment 2.

Experiment 2 main data

***** Figure 2 about here ********

Figure 2A shows the mean JND at the 75% correct response rate plotted as a function of the centre BMI of the 8 BMI ranges for the two CGI images. We used PROC MIXED in SAS v9.3 to quantify the relationship between JND, reference BMI and model. The threshold data required transforming to ensure that they conformed to a normal distribution. We permitted individual slope and intercept variation for each subject by specifying an 'unstructured' variance-covariance structure for the G-matrix. The mixed model explained 81% of the variance in mean JND. The Type III test of the fixed effect of BMI was statistically significant, $F(1,196) = 89.39$, $p < .0001$, CI [0.02 - 0.03]. However, neither the fixed effect of model nor the interaction between model and BMI were statistically significant: $F(1,71.2) = 3.12$, $p = .08$, CI [-0.03 - 0.40] and $F(1,196) = 1.98$, $p = .16$, CI [-

14

0.01 - 0.003], respectively. For the real images in Experiment 2B, the equivalent mixed model explained 81% of the variance in mean JND. The Type III test of the fixed effect of BMI was statistically significant, $F(1,86.5) = 44.57$, $p < .0001$, CI [0.05 - 0.11].

Discussion

In Experiment 1, our participants estimated the weight of 120 women varying in their body mass. Their estimates clearly show contraction bias with bodies below 70Kg being increasingly over-estimated and bodies above 70Kg being increasingly under-estimated. For example, an observer who judges the weight of a 100kg woman will under-estimate her weight by ~10kg. A value of 70 kg is the average body weight for adult women in the UK (Health Survey for England, 2012), and its adoption as a reference value against which to judge other women's bodies would be consistent with people's visual diet shaping their reference body so that it reflects the population norm. As the height and weight of the women in the photographs is known, we could calculate both their actual BMI and the BMI of their bodies based on the participants' estimation of their weight. These data show the same pattern of contraction bias, with a BMI of 27 being the most accurately judged, again consistent with a reference template based on the average BMI for adult women in the UK (Health Survey for England, 2012). Contraction bias can also be used to explain the accuracy of judgements of own body size over a wide range of BMI values from emaciated to obese (Cornelissen, Johns & Tovée, 2013; Cornelissen et al., 2015) and can also explain the fact that previous studies have consistently shown that obese people under-estimate their size relative to normal weight people, and why it gets progressively harder to detect weight increase as the overall weight of the individual increases (Kuchler & Variyam, 2003; Kuskowska-Wolk & Rössner, 1989; Maximova et al., 2008; Rahman & Berenson, 2012; Robinson & Kirkham, 2013; Truesdale & Stevens, 2008; Wetmore & Modkdad, 2012).

15

An additional problem in judging when a body is becoming over-weight/obese is the reduced ability to discriminate size change in heavier bodies. In Experiment 2, participants had to judge which body was the heavier in a 2-AFC paradigm. It demonstrated that Weber's law applies to body size judgements (i.e. a larger rise in BMI is required in higher BMI bodies to be detected than in lower BMI bodies). For example, for a height of 1.6 metres, women with BMIs of ~22 (normal) and ~35 (obese) would just be able to detect increases in body weight of ~3kg and ~5.5kg respectively. Experiment 2 used both CGI bodies (study 1) and real bodies (study 2). The CGI bodies accurately simulate the pattern of fat shown in real bodies (Cornelissen et al., 2015) and are judged in the same way as real bodies (Tovée, Edmonds & Vuong, 2012). The use of artificial bodies allows features such as height, proportions and skin colour and texture to be held constant and the effect of increasing body fat to be directly modelled. This approximates to an individual gaining weight and tests the participant's perceptual ability to discriminate fat addition. By using 2 different body models and two different body fat simulations, we minimised the possibility that the results are an artefact of the CGI fat simulation, a conclusion supported by the same pattern of results in real bodies. The real bodies vary in a number of dimensions, including proportions and height, and illustrate the difficulties that a health professional encounters in making a visual judgement of whether someone has become obese (i.e. are they above or below the overweight/obese category). Both parts of Experiment 2 therefore showed that the ability to discriminate a difference became progressively worse as the BMI of the bodies being judged increased, as we should expect from Weber's law. Moreover, we also confirmed that the Weber fraction (i.e. the JND in BMI units divided by the BMI of the bodies being judged) remained reasonably constant across the BMI range tested.

Our results suggest that two purely perceptual factors make it harder to detect both being obese and weight increase when obese, but this does not rule out more cognitive factors

also playing an important role. Judging bodies has both a perceptual and cognitive component. The perceptual component is the ability to accurately estimate the shape and size of a body and the cognitive component is how this estimation is interpreted. What is regarded as an acceptable body size or weight is influenced by cultural and media values, it is not just the sizes of the bodies we see every day both in real life and in the media (visual diet), but also the positive or negative social values we put on them (visual valence) and the context in which we see them (Tovée, Swami, Furnham, & Mangalparsad, 2006; Boothroyd, Tovée & Pollet, 2012; Bateson, Tovée, George, Gouws & Cornelissen, 2014; Boothroyd et al., 2016). Additionally, the accuracy of the judgements may be modulated or influenced by the ethnic or social group of the observer, their own anthropometric or psychological characteristics (including their contraceptive or hormonal status) (e.g. Cornelissen, Johns & Tovée, 2013; Cornelissen, Bester, Cairns, Tovée & Cornelissen, 2015; Rahman & Berenson, 2012; Robinson & Hogenkamp, 2014).

A potential limitation is that we are asking women to judge women's bodies, and thus do not have data on men's judgements or judgements of men's bodies. However, we are testing the existence of basic perceptual phenomena which should apply to anybody making a judgement about a body of whatever gender or age. Consistent with this hypothesis, studies which have tested the accuracy of the weight estimation of men's bodies have shown the same pattern of under-estimation as reported here, although they interpreted their results differently (e.g. Robinson, Parretti & Aveyard, 2014; Oldham & Robinson, 2015; Robinson & Kirkham, 2013). Additionally, both men and women seem to rate men's and women's bodies in the same way and within a particular culture have the same preference for the ideal body size for men and women which suggest a common pattern of assessment across both genders (e.g. Crossley et al., 2012; Smith, Cornelissen & Tovée, 2006; Swami & Tovée, 2005).

Of course there are multiple potential cues to weight gain. For example, as you add weight your clothes become tighter and you can quantify any weight change by stepping on the bathroom scales. However, it is easy to rationalise tighter clothes (they have shrunk in the wash) and the scales themselves may not be very accurate and provide unreliable feedback on body weight (e.g. Yorkin, Spaccarotella, Martin-Biggers, Quick & Byrd-Bredbenner, 2013). Additionally, many people, particularly men (see e.g. Striegel-Moore et al., 2009), may not check their weight unless they had a reason to do so. If you look in the mirror and do not detect any weight change it is quite possible that you will not have any reason to check your weight through standing on the scales.

The problems of detecting change in body size also represent a problem in people seeking to lose weight. Weber's law means that people who are obese have to lose a significantly larger amount of weight for it to be perceptible than someone of lower weight. For example as calculated above, someone who has a BMI of 35 would have to lose at least 5.5kg for it to be visually detectable. So they receive no positive visual feedback until they have lost a comparatively large amount of weight. As improving appearance is one of the reasons commonly given by overweight and obese people for trying to lose weight (Clarke, 2002; Dixon, Dixon & O'Brien, 2002; Hankey, Leslie & Lean, 2002), and as appearance is principally assessed visually, the lack of change in apparent appearance despite weight loss could undermine and potentially demotivate people in weight loss programmes.

It is also important to accurately detect obesity and weight change in others. As we discussed in the introduction, GPs should routinely screen their patients for being over-weight or obese. As patients visit a busy surgery, visual assessment of their weight status maybe a default way of screening their adiposity. However, our results suggest that as people's weight increases an observer will increasingly under-estimate their body size. This may explain the discrepancy in the proportion of patients being reported as being over-weight

or obese relative to the proportion in the general population (Robinson, Parretti & Aveyard, 2014). This may also be a reason why parents don not seem to recognise their children are over-weight and that they are getting heavier (Jones et al., 2011; Duncan, 2011).

Our results clearly point to the potential for perceptual factors contributing to problems with detecting obesity and weight increase. Overcoming these perceptual biases, however, may not be straightforward. The size and weight of an observer's internal reference body can be altered by selectively viewing bodies of a particular BMI. So the presentation of lighter bodies can recalibrate the perceptual "normal" size to a lower BMI, but this would also cause an increase in the under-estimation of body size through contraction bias (as the difference between the body being judged and the reference body would increase). However, the ability to detect body size increase (the JND) can potentially be improved in straightforward training programs. Previous studies have suggested that training in discriminating feature change can significantly reduce JND values; the expertise effect (Ball & Sekuler, 1982; Gauthier & Tarr, 1997; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Such a training program in the form of a downloadable app for smart phone, tablet or PC could therefore play an important role in improving the ability of health professionals to visually detect whether their patients are over-weight or obese. Thus in the case of a GP seeing a patient in the normal course of treatment, he or she would be able to accurately judge their patient's BMI and initiate a discussion on weight loss options if their BMI was in the overweight or obese range. It could also support weight loss programs. It would make over-weight and obese people more sensitive to weight gain and thus more likely to undertake weight control behaviours, and it would also make them more sensitive to weight loss and provide positive reinforcement for people taking part in weight loss programs. Finally, such a program incorporating child bodies rather than adults (Jones et al., 2015), could provide

training for parents in detecting whether their children are becoming overweight and whether they need to change their family's lifestyle to compensate.

References

Ball, K., & Sekuler, R. (1982) A specific and enduring improvement in visual motion discrimination. *Science, 218,* 697-698. doi: 10.1126/science.7134968

Bateson, M., Tovée, M.J., George, H.R., Gouws, A. & Cornelissen, P.L. (2014). Humans are not fooled by size illusions in attractiveness judgements. *Evolution & Human Behavior*, 35, 133-139. doi:10.1016/j.evolhumbehav.2013.11.007.

Boothroyd LG, Jucker JL, Thornborrow T, Jamieson M, Burt DM, Barton R, Evans EH, Tovée MJ. (2016). Television exposure predicts body size preferences in rural Nicaragua. *British Journal of Psychology* 2016, (in press).

Boothroyd, L.G., Tovée, M.J., & Pollet, T.V. (2012). Visual diet versus associative learning as mechanisms of change in body size preferences. *PLoS One, 7,* e48691. doi: 10.1371/journal.pone.0048691

Clarke, L.H. (2002). Older women's perceptions of ideal body weights: the tensions between health and appearance motivations for weight loss. *Ageing and Society 22*, 751 − 773. doi: 10.1017/S0144686X02008905

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.), New Jersey: Lawrence Erlbaum Associates.

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, 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, 7, e50601. doi: 10.1371/journal.pone.0050601.

Dixon, J.B., Dixon, M.E. & O'Brien, P.E (2002) Body image: Appearance orientation and

evaluation in the severely obese. Changes with weight loss. *Obesity Surgery, 12*, 65-

71. doi: 10.1381/096089202321144612

Duncan, D.T. (2011). Parental misperception of their child's weight status: clinical

implications for obesity prevention and control. *Obesity, 19*, 2293.

doi:10.1080/17477160600586614

Gauthier, I., & Tarr M.J. (1997). Becoming a "greeble" expert: exploring mechanisms for

face recognition. *Vision Research, 37,* 1673–168. doi: 10.1016/S0042-

6989(96)00286-6

Gauthier, I., Tarr, M.J., Anderson, A.W., Skudlarski, P., & Gore, J.C. (1999). Activation of

the middle fusiform 'face area' increases with expertise in recognizing novel objects.

Nature Neuroscience, 2, 568-573. doi:10.1038/9224

Gardner, R.M., & Bokenkamp, E.D. (1996). The role of sensory and nonsensory factors in body size estimations of eating disorder subjects. *Journal of Clinical Psychology, 52,* 3-15. doi: 10.1002/(SICI)1097-4679(199601)52

Gescheider, G.A. (1997). *Psychophysics: The Fundamentals*. (3rd Ed.) Lawrence Erlbaum, New Jersey, USA.

Grover, S.A., Kaouache, M., Rempel, P., Joseph, L., Dawes, M., Lau, D.C.W., et al. (2014). Years of life lost and healthy life-years lost from diabetes and cardiovascular disease in overweight and obese people: a modelling study. *Lancet Diabetes Endocrinology, 3,* 114–122. doi: 10.1016/S2213-8587(14)70229-3

Hankey, C.R., Leslie, W.S. & Lean, M.E.J. (2002). Why lose weight? Reasons for seeking weight loss by overweight but otherwise healthy men. *International Journal of Obesity*, 26, 880 – 882. doi:10.1038/sj.ijo.0801999

Health Survey for England (2012). National Centre for Social Research and University College London. Department of Epidemiology and Public Health. UK Data Archive, Colchester, Essex, UK.

Jackson-Leach, R. &, Lobstein, T. (2006). Estimated burden of paediatric obesity and co-morbidities in Europe. Part 1. The increase in the prevalence of child obesity in Europe is itself increasing. *International Journal of Pediatric Obesity, 1*, 26-32. doi:10.1080/17477160600586614.

Jones, A.R., Parkinson, K.N., Drewett, R.F., Hyland, R.M., Pearce, M.S., Adamson, A.J., et al. Parental perceptions of weight status in children: the Gateshead Millennium Study. *International Journal of Obesity*, 35, 953-962. doi:10.1038/ijo.2011.106

Kuchler, F., & Variyam, J.N. (2003). Mistakes were made: misperception as a barrier to reducing overweight. *International Journal of Obesity, 27*, 856–886. doi:10.1038/sj.ijo.0802293

Kuskowska-Wolk, A., & Rössner, S. (1989). The "true" prevalence of obesity: a comparison of objective weight and height measures versus self-reported and calibrated data. *Scandinavian Journal of Primary Health Care,7,* 79-82. doi: 10.3109/02813438909088651

Leopold, D.A., O'Toole, A.J., Vetter, T. & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*, 89–94. doi: 10.1038/82947

Maximova, K., McGrath, J.J., Barnett, T., O'Loughlin, J., Paradis, G., & Lambert, M. (2008). Do you see what I see? Weight status misperception and exposure to obesity among children and adolescents. *International Journal of Obesity, 32,* 1008-1015. doi: 10.1038/ijo.2008.15.

McKinsey Global Institute (2014). *Overcoming obesity: An initial economic analysis.* McKinsey & Company (http://www.mckinsey.com/insights/economic_studies/how_the_world_could_better_fight_obesity).

Ogden, C.L., Carroll, M., Curtin, L.R., McDowell, M.A., Tabak, C.J., & Flegal, K.M. (2006). Prevalence of overweight and obesity in the United States, 1999-2004. *JAMA, 295,* 1549–1555. doi:10.1001/jama.295.13.1549.

Oldham, M. & Robinson, E. (2015). Visual weight status misperceptions of men: Why

overweight can look like a healthy weight. *Journal of Health Psychology, 20,* 1–10.

doi: 10.1177/1359105314566257

Olds, T,I,M,, Maher, C., Zumin, S.H.I., Peneau, S., Lioret, S., Castetbon, K., et al. (2011).

Evidence that the prevalence of childhood overweight is plateauing: data from nine

countries. *International Journal of Pediatric Obesity* 6: 342-360. doi:

10.3109/17477166.2011.605895

Poulton, E.C. (1989). *Bias in Quantifying Judgements*. Erlbaum, Hove, UK.

Probst, M., Vandereycken, W., & Van Coppenolle, H. (1997). Body-size estimation in eating

disorders using video distortion on a life-size screen. *Psychotherapy and

Psychosomatics, 66,* 87-91. doi: 10.1159/000289114

Rahman, M. & Berenson, A.B. (2012). Self-Perception of Weight Gain Among Multiethnic

Reproductive-Age Women. *Journal of Women's Health, 21,* 340-346.

doi:10.1089/jwh.2011.2982

Rhodes, G., Jeffery, L., Boeing, A., & Calder, A. (2013). Visual coding of human bodies:

Perceptual aftereffects reveal norm-based, opponent coding of body identity. *Journal

of Experimental Psychology: Human Perception and Performance*, 39, 313-317.

doi:10.1037/a0031568.

Robinson, E., & Kirkham, T.C. (2013). Is he a healthy weight? Exposure to obesity changes

perception of the weight status of others. *International Journal of Obesity, 38,* 663-

667. doi:10.1038/ijo.2013.154.

Robinson, E. & Hogenkamp, P.S. (2015). Visual perceptions of male obesity: a cross-cultural study examining male and female lay perceptions of obesity in Caucasian males. *BMC Public Health, 15*, 492. doi:10.1186/s12889-015-1821-3

Robinson, E., Parretti, H. & Aveyard, P. (2014). Visual identification of obesity by healthcare professionals: an experimental study of trainee and qualified GPs. *British Journal of General Practice, 64*, e703-e708.doi:10.3399/bjgp14X682285.

Rosenstock, I.M., Strecher, V.J. & Becker, M.H. (1988). Social Learning Theory and the Health Belief Model. *Health Education Quarterly, 15*, 175-183. doi: 10.1177/109019818801500203

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

Striegel-Moore, R.H., Rosselli, F., Perrin, N., DeBar, L., Wilson, G.T., May, A. & Kraemer, H.C. (2009). Gender difference in the prevalence of eating disorder symptoms. *International Journal of Eating Disorders, 42,* 471–47. doi: 10.1002/eat.20625

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

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

Truesdale, K., & Stevens, J. (2008). Do the obese know they are obese? *North Carolina Medical Journal, 69,* 188–194.

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

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

Tovée, M.J., Swami, V., Furnham, A., & Mangalparsad, R. (2006). Changing perceptions of attractiveness as observers are exposed to a different culture. *Evolution and Human Behaviour,* 27, 443-456. doi: 10.1016/j.evolhumbehav.2006.05.004

Wells, J.C., Treleaven, P., & Cole, T.J. (2007). BMI compared with 3-dimensional body shape: the UK national sizing survey. *American Journal of Clinical Nutrition, 85,* 419-425. doi:10.1038/sj.ijo.0803727.

Wetmore, C., & Modkdad, A.H. (2012). In denial: misperceptions of weight change among adults in the United States. *Preventative Medicine, 56,* 93–100. doi:10.1038/sj.ijo.0803727

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

Yorkin, M., Spaccarotella, K., Martin-Biggers, J., Virginia Quick, V. & Byrd-Bredbenner, C. (2013). Accuracy and consistency of weights provided by home bathroom scales. *BMC Public Health*, 13:1194. doi:10.1186/1471-2458-13-1194

Figure Legends

Figure 1: Scatterplot depicting the relationship between the actual weight of the women in the images (kg) and the mean of the participants' estimations of their weight (i.e. one data point represents one image). The red line represents the linear regression of estimated weight on actual weight and the pink shaded region its 95% confidence limits. The dotted black line represents the line of equality (i.e. slope unity, intercept zero).

Figure 2: Figure 2A is a plot of mean JND as a function of the reference BMI value for each of the BMI ranges for the two CGI 3D models. The two models are indicated by the upward and downward pointing cyan triangles, respectively. Error bars represent 1 s.e. of the mean. The solid red line represents the main effect of BMI on JND derived from the mixed models (see text for details), and the pink shading its 95% confidence band. Figure 2B is an equivalent plot of mean JND as a function of the reference BMI value for each of the BMI ranges for the photos of real women. The Weber fractions in A and B were ~0.06 and ~0.1 respectively.
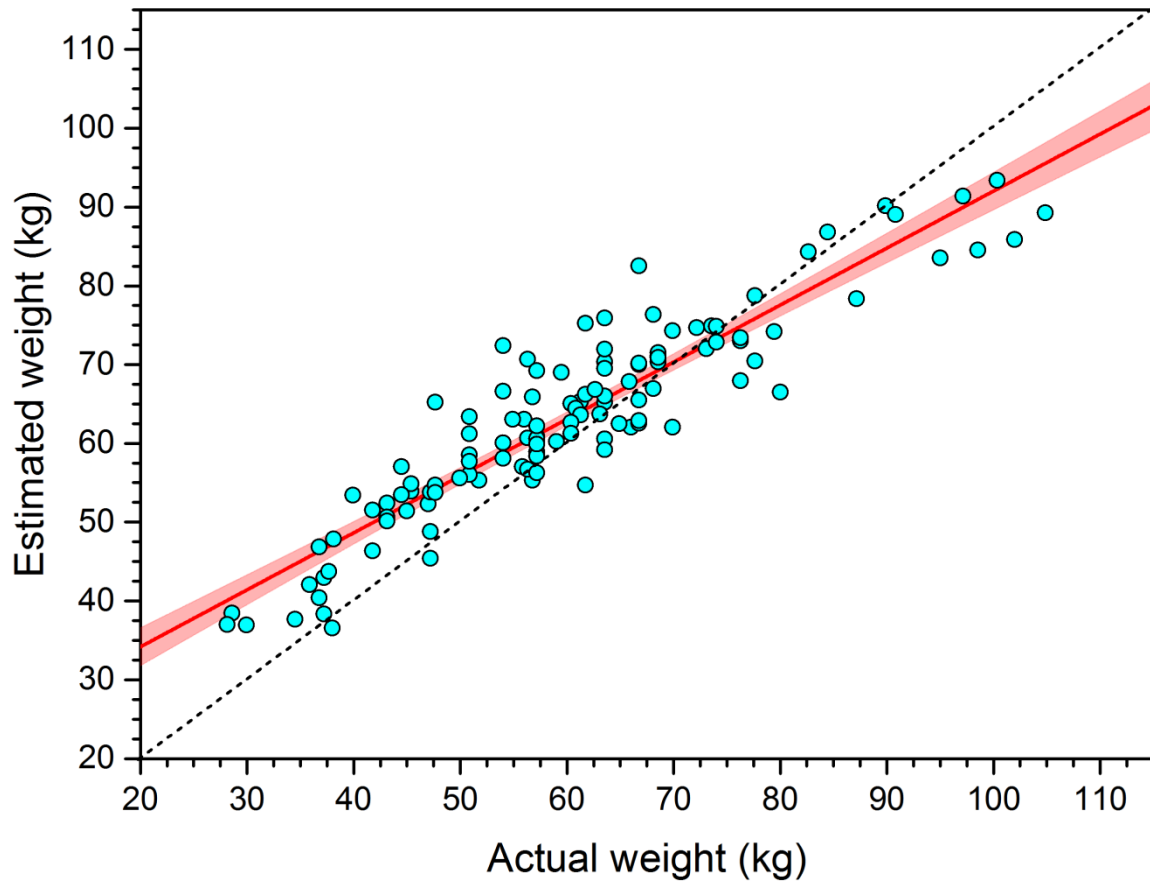
Figure 1

Figure 2