NOVEL METHODS OF MEASURING AND MODIFYING PERCEPTION OF AND SATISFACTION WITH BODY SIZE: A CLINICAL TRANSLATIONAL APPROACH

HELEN BOULD

WORCESTER COLLEGE

Thesis submitted for the degree of Doctor of Philosophy

Department of Psychiatry

University of Oxford

Trinity Term, 2018

ABSTRACT

Novel methods of measuring and modifying perception of and satisfaction with body size: a clinical translational approach

Dr Helen Bould, Worcester College

Submitted for the degree of Doctor of Philosophy, Trinity Term 2018

Body dissatisfaction is a risk factor for subsequent eating disorders, weight gain and low mood. It is resistant to treatment and, untreated, predicts relapse to eating disorders. Measurement of its perceptual and cognitive components currently relies predominantly on self-report questionnaires.

This thesis describes the development and testing of a novel brief intervention designed to change perception of and satisfaction with own body size. Three studies demonstrate and replicate the finding that whether one-off, or repeated over the course of a week, exposure to over- or normal- weight images leads to participants subsequently perceiving others' bodies as smaller, in comparison to those shown underweight bodies. One-off exposure to images of overweight, compared to normal or underweight, women also leads to participants viewing their own bodies as smaller, and being more satisfied with their body size; however, repeated exposure over the course of a week does not have this effect, and in fact leads to women perceiving their own bodies as larger. Such repeated exposure to over- rather than underweight bodies does however increase women's ideal body size.

The thesis also describes two studies concerning the development and testing of novel measures of body size and satisfaction with body size in healthy women and those

with an eating disorder. Using an Avatar to demonstrate real and ideal body size, and a novel behavioural measure of body satisfaction using outfits, both correlated well with existing measures of body size and satisfaction, and there were differences on these measures between individuals with and without eating disorders. The novel implicit tasks (Implicit Association Tasks and a Lexical Decision Task) also correlated with existing measures of body size and satisfaction, however, there were no differences on these measures between individuals with and without eating disorders.

The findings suggest that exposure to images of women of different sizes is effective in changing perception of what constitutes a normal weight, but has less predictable effects in relation to perception of and satisfaction with own size. Implicit measures of body size and satisfaction do not appear to be useful in distinguishing between women with and without eating disorders, but a novel behavioural measure using outfits, and using Avatars to demonstrate perceived own size, does show promise.

TABLE OF CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

ACKNOWLEDGEMENTS

Thank you first to my supervisors, Catherine Harmer, Matthew Broome and Rebecca Park for all their help and support. Thanks also to Marcus Munafò and his lab group, and to Ian Penton-Voak, Andy Skinner, Glyn Lewis and the team at the Centre for Academic Mental Health at the University of Bristol for getting this process started, and for their wisdom along the way.

Thanks to the PERL lab group and the Child Psychiatry group led by Alan Stein for helping when I got stuck, stopping for lunch, and supporting me through the highs and lows of life during the DPhil.

Thanks to the Child and Adolescent Psychiatry consultants and registrars in both Bristol and Oxford for supporting me in not forgetting how to be a clinician whilst I tried to also become a researcher.

Thanks to my in laws Angela and Stephen Lightfoot, especially for the Freddie (and Kitty) Fridays, and thank you to my parents Bruce and Tree Bould for this and more, and for more years than I care to mention.

Thanks to Freddie, for making sure I left work on time from the beginning, and to Kitty, whose arrival was responsible for the nine-month gap in the middle of Study 5.

And of course most thanks of all to my husband Jonathan who moved house (twice) and spent many hundreds of hours commuting across the country so that I could do this. I couldn't have done it without you.

DECLARATION

Some of the work in this thesis has previously been published and is included in the Appendices.

Chapter 2

Helen Bould, Rebecca Carnegie, Heather Allward, Emily Bacon, Emily Lambe, Megan Sapseid, Katherine S. Button, Glyn Lewis, Andy Skinner, Matthew R. Broome, Rebecca Park, Catherine J. Harmer, Ian S. Penton-Voak, Marcus R. Munafò. "Effects of exposure to bodies of different sizes on perception of and satisfaction with own body size: two randomized studies" Royal Society Open Science. 2018 5 171387; DOI: 10.1098/rsos.171387

Chapter 3

(Published as a preprint; joint first authorship with Andy Skinner)

Andy Skinner, **Helen Bould,** Peter J. Etchells, Ian S. Penton-Voak, Andrew Clark, Catherine Harmer, Rebecca Park, Matthew Broome, Marcus Munafò. 2017. "The Morphed Photographic Figure Scale: Creation and Validation of a Novel Set of Realistic Female Body Stimuli." OSF. January 23. doi:10.17605/osf.io/UGJHA

ABBREVIATIONS

95% CI	95% Confidence Interval
AN	Anorexia Nervosa
BED	Binge Eating Disorder
BN	Bulimia Nervosa
BDI	Beck Depression Inventory
BMI	Body Mass Index
BSSS	Body Shape Satisfaction Scale
CBT-ED	Cognitive Behavioural Therapy for Eating Disorders
EDE	Eating Disorders Examination
EDE-Q	Eating Disorders Examination – Questionnaire
IAT	Implicit Association Task
IQ	Intelligence Quotient
LDT	Lexical Decision Task
NART	National Adult Reading Test
OR	Odds Ratio
OSFED	Other Specified Feeding and Eating Disorder
PANAS	Positive and Negative Affect Scale
PHQ-9	Patient Health Questionnaire – 9
RSES	Rosenberg Self-Esteem Scale
SCID	Structured Clinical Interview for DSM-IV
SD	Standard Deviation
STAI	State-Trait Anxiety Inventory
VAS	Visual Analogue Scale

STATEMENT OF CONTRIBUTIONS

Study 1

Dr Rebecca Carnegie (RC) assisted with data collection. Dr Katherine Button (KB) programmed the ePrime task used in this study and Study 2, and held the randomisation codes for both studies. Dr Angela Attwood assisted with checking the data generated in this study and Study 2.

Study 2

Ms Emily Lambe (EL), Ms Heather Allward (HA), Ms Megan Sapseid (MS) and Ms Emily Bacon (EB) assisted in data collection for Study 2; Prof Ian Penton-Voak (IPV) co-supervised MS, EB and HA in their dissertations, which made use of subsets of the data collected during this study.

Study 3

Dr Andy Skinner (AS) used Psychomorph software to generate images of the same women morphed to appear of different weights. Dr Andy Wood (AW) created an online version of the *Morphed Photographic Figure Scale* and the *Outfit Choice Questionnaire* for use in this study

Study 4

Mr Jeffery Dejong (JD) created a user interface for the Avatar programme used in this study and Study 5, programmed the training task and assigned numbers to the three trial groups for Study 4. Dr Katharine Noonan (KN) assisted with data collection, data entry and with double checking that data entry was correct. Andrea Reinecke (AR) assisted with programming the LDT and IAT tasks in ePrime.

Study 5

Marianne O'Connor (MO'C) assisted with finalising EDE diagnoses where there was diagnostic uncertainty. Ms Alex Pike (AP) assisted with recruitment and with double checking data entry.

CHAPTER 1: INTRODUCTION

Body dissatisfaction is a key risk factor in the development, maintenance and relapse of eating disorders (Jacobi, Hayward, de Zwaan, Kraemer, & Agras, 2004; Neumark-Sztainer, Paxton, Hannan, Haines, & Story, 2006; E. Stice, Marti, & Durant, 2011). It also increases the risk of low mood (Paxton, Neumark-Sztainer, Hannan, & Eisenberg, 2006) and subsequent weight gain (Quick, Wall, Larson, Haines, & Neumark-Sztainer, 2013b; Sonneville et al., 2012; van den Berg & Neumark-Sztainer, 2007). It has been identified as a key target area for the development of novel interventions (Image, 2012).

This thesis will describe a series of experiments designed to test whether it is possible to alter perception of own and others' body size, and thus modify body dissatisfaction. It will go on to describe the development and testing of novel measures aimed at improving the measurement of different aspects of perception of and satisfaction with own and other's body size. Finally, it will use these novel measures to explore differences in body perception and satisfaction in women with and without a current eating disorder.

The introduction will set the scene by describing body dissatisfaction, and current research into whether it is possible to alter perception of and satisfaction with own and other's body size, including a consideration of possible mechanisms involved in perception of body size. It will give an overview of the current measures of perception and satisfaction with own size. It will then give a brief overview of eating disorders, describing their epidemiology, classification, core psychopathology including the role

of body dissatisfaction, existing treatments, and the need for novel treatments. It will also describe the evidence for differences between women with and without an eating disorder in terms of how they perceive their bodies. It will conclude with an overview of the experiments described in the chapters that constitute the rest of the thesis.

1. Body Dissatisfaction

Body dissatisfaction is common in the general population, with prevalence studies finding that 35% of adolescent girls (Bearman, Martinez, Stice, & Presnell, 2006), and two thirds of adult women (Frederick, Peplau, & Lever, 2006) are dissatisfied with their bodies. Dissatisfaction with body size is a key part of this: over 75% of adolescent girls (Ricciardelli & McCabe, 2001) and over 95% of adult women (V. Swami et al., 2012) report a wish to be thinner, and 74% of adult women report a wish to lose weight (Muennig, Jia, Lee, & Lubetkin, 2008).

These high rates of body dissatisfaction are often ascribed to "internalisation of the thin ideal". "Internalisation of the thin ideal" refers to the idea that individuals, to differing amounts, try to conform to media-driven, socially-defined ideas of attractiveness, including body size (Eric Stice, Schupak-Neuberg, Shaw, & Stein, 1994; Thompson, 1999). Although it is sometimes described as a "normative discontent", body dissatisfaction has many negative consequences, increasing the risk of eating disorders, low mood, and subsequent weight gain (Jacobi et al., 2004; Neumark-Sztainer, Eisenberg, Fulkerson, Story, & Larson, 2008; Paxton et al., 2006; Quick, Eisenberg, Bucchianeri, & Neumark-Sztainer, 2013; Sonneville et al., 2012; E. Stice et al., 2011; van den Berg & Neumark-Sztainer, 2007).

Body dissatisfaction is a key component of eating disorders, with "disturbance in the way in which one's body weight or shape is experienced" being part of the DSM-5 (American Psychiatric Association, 2013) criteria for diagnosing Anorexia Nervosa, and over-concern with body shape and weight part of the diagnostic criteria for Bulimia Nervosa in ICD-10 (World Health Organisation, 1992). As an aspect of eating disorder psychopathology, body dissatisfaction is particularly hard to treat, often persisting after otherwise successful treatment (Exterkate, Vriesendorp, & de Jong, 2009), and, if untreated, predicts maintenance of and relapse to eating disorders (Jacobi et al., 2004; E. Stice et al., 2011).

The concept of body dissatisfaction is often subdivided into two components: (1) **perception** of one's own body size, and (2) an attitudinal or cognitive component of **dissatisfaction with body size** and shape (Cash & Deagle, 1997).

1.1 Perception of body size

Conceptualising body dissatisfaction as having perceptual and cognitive components suggests the possibility that if one could change perception of body size, it may be possible to change satisfaction with size. If this could be achieved, it would give rise both to potential new treatments for eating disorders, and to strategies to prevent body dissatisfaction at a public health level.

1.1.1 Perception of others' body size

There is some evidence that simply exposing people to images of bodies of different sizes can alter their perception of *others'* body size. The body size viewed as being most "normal" becomes thinner after exposure to thinner bodies, and larger following exposure to larger bodies (Brooks, Mond, Stevenson, & Stephen, 2016; Glauert,

Rhodes, Byrne, Fink, & Grammer, 2009; Hummel et al., 2013; Hummel, Rudolf, Untch, Grabhorn, & Mohr, 2012; Robinson & Christiansen, 2014; Robinson & Kirkham, 2014; Sturman, Stephen, Mond, Stevenson, & Brooks, 2017; Winkler & Rhodes, 2005). Viewing photographs of males whose body mass index (BMI) is in the obese, as opposed to normal, category leads to both male and female participants judging an image of a man with a BMI in the overweight range to have a healthy, normal weight (Robinson & Christiansen, 2014; Robinson & Kirkham, 2014). The body size perceived as most attractive also becomes smaller following exposure to thinner as opposed to larger bodies (Glauert et al., 2009; Winkler & Rhodes, 2005). Effects have been found in a range of types of images, including line drawings (Winkler & Rhodes, 2005), computer-generated images (Glauert et al., 2009), computer-manipulated photographs (Brooks et al., 2016; Hummel et al., 2013; Hummel et al., 2012; Mohr, Rickmeyer, Hummel, Ernst, & Grabhorn, 2016; Sturman et al., 2017) and photographs of human bodies (Robinson & Christiansen, 2014; Robinson & Kirkham, 2014). Effects have also been seen separately for muscle and fat (Sturman et al., 2017).

1.1.2 Perception of own body size

Previous studies of whether exposure to bodies of different sizes can also change perception of *own* body size predating this thesis had been small (Ns of 14 and 16) (Hummel et al., 2013; Hummel et al., 2012). Participants in these two studies were shown "thin" or "fat" modified photographs of self or others, and were then asked to pick the most accurate photograph of themselves: those shown thinner-than-actual photographs of self or others judged a thinner-than-actual picture of self as most accurate; those shown fatter-than-actual photographs of self or others subsequently

judged a fatter-than-actual photograph as most accurate. This suggests that photographs of others may change perception of own size, but the studies were too small to be confident that these results did not arise by chance. Concurrently with the work described in this theses, Brooks (Brooks et al., 2016) found that exposure to images of either thinner versions of self or thinner versions of others led to subsequent images of self and others being viewed as larger (and vice versa). However, they found that although effects did transfer to some extent between images of own and others bodies, the effects were larger when the exposure and testing were with the same body type (e.g. both images were of self).

1.1.3 Mechanisms for perception of body size

Whilst this thesis does not set out to test the possible mechanisms underlying perception of body size, it is useful to consider how changes in perception of the size of own and others bodies might arise.

The majority of the studies to date assume that the underlying mechanism for the observed changes in perception of the size of others' bodies following exposure to bodies of different sizes is *visual adaptation*. Adaptation occurs when a subject has prolonged exposure to a particular stimulus. They then experience aftereffects of this exposure when they subsequently view other, similar, stimuli. A classic example of this is of focussing on an object moving in one direction; stationary stimuli viewed immediately afterwards appear to be moving in the opposite direction (Addams, 1834).

Visual adaptation has been investigated with regard to how faces are perceived: for example, it has been shown that adapting participants to a series of distorted faces (features contracted to the centre of the face) leads to subsequent (undistorted) faces

being perceived as distorted in the opposite direction (features expanded from the centre) (Rhodes et al., 2011). Visual adaptation has also been shown to have potential therapeutic effects: repeated presentation of angry faces leads to individuals being more likely to perceive subsequent, ambiguous, faces as happy, and to have lower levels of self-reported anger (Penton-Voak et al., 2013).

With regard to perception of body size, visual adaptation would predict that repeated or extended exposure to smaller images would lead to participants viewing subsequent images as larger, and vice versa. This is what has been seen in the literature described above; however these findings do not in themselves confirm that the underlying mechanism is visual adaptation.

A related mechanistic explanation is that of *"norm-based" coding*. Again, this idea has been more thoroughly explored in the literature on the perception of faces. The central concept is that neural representations of faces are "norm-based" (Rhodes & Jeffery, 2006; Webster & MacLeod, 2011). That is to say, we have an internal representation of "an average face functioning as a norm against which identity-related deviations can be coded" (Rhodes, Jeffery, Boeing, & Calder, 2013). It is postulated that each dimension of the face is coded by a pair of neural channels tuned to opposite extremes, and that "normal" is represented by equal activation in both channels. "Norms" are believed to be updated by experience, so as to allow "fine-tuning of person perception coding mechanisms to our perceptual diet" (Rhodes et al., 2013). This mechanism has also been tested in relation to body identity (Rhodes et al., 2013), and similar findings to those from studies of faces led the authors to conclude that body identity, like facial identity, is coded by pairs of neural channels, tuned to

extreme values. Extended to perception of body size, this idea would suggest that repeated exposure to "underweight" bodies would lead to the body weight "norm" shifting, so that what is perceived as "normal" becomes more underweight. Again this mechanism could help to explain the results found to date in the literature.

1.2 Satisfaction with own body size

Another line of work, following on from the "internalisation of the thin ideal" theory, has explored the impact of exposure to different types of images on body satisfaction. Meta-analysis of these studies (Grabe, Ward, & Hyde, 2008) concludes that in general women feel less satisfied with their bodies after viewing photographs of thin models compared to neutral images or normal-sized models.

However, of the 80 studies included in this meta-analysis, 75 either had no control group, or compared viewing thin models with neutral images of some sort, either unrelated to bodies, or, in two studies, images of older men and women, and infants and children playing (Kalodner, 1997; Wegner, Hartmann, & Geist, 2000). They therefore do not give insight into what aspects of the women's bodies might be having this effect. Of those using women's bodies in the control group, one study used images which were chosen to vary according to attractiveness rather than weight (Cash, Cash, & Butters, 1983) and one used television footage of "thin and attractive" women in the experimental group and compared this with "not thin or attractive" women in the control group (Cattarin, Thompson, Thomas, & Williams, 2000). This leaves uncertainty as to whether the differences found were due to factors relating to attractiveness or to weight. Two studies used clothing adverts or fashion models from women's magazines in the experimental group, but compared them with "pictures

from specialised women's clothing catalogs" for "more amply furnished women" in one case (Wilcox & Laird, 2000) and "overweight bordering on obese women" from other magazines in the other (Ogden & Mundray, 1996). It seems likely that such images would have varied in terms of the attractiveness and glamour of the images as well as their size and weight. One study attempted to look at the effects on body satisfaction of photographs which varied on weight alone, using the face of one woman pasted onto either a relatively slim or an overweight body (Lin & Kulik, 2002). This study found lower levels of body satisfaction in the group shown a single underweight image. However, the study did not measure baseline levels of body satisfaction, so it is possible that these differences between groups were pre-existing. Interpreting these results is further complicated by the study scenario, in which the women were told that they were about to meet a male participant who would choose whether he would rather go on a date with them, or the other woman, whose photograph they were purportedly being shown. One further small study (N=40, randomised into four groups) used a single black and white print advert for Kraft cheese, featuring an "attractive female model" in two conditions, and a version using the model altered to appear overweight using Adobe Photoshop (no further details of the manipulation are reported) in a further two conditions; participants rated self-attractiveness after viewing the image for 15 seconds. Self-attractiveness ratings were lower following exposure to the "overweight" models, but again no baseline measurements were taken (Crouch & Degelman, 1998).

1.3 Measuring body size perception and satisfaction

The measures used to assess body size perception and satisfaction are key to the design and interpretation of studies attempting to alter them.

1.3.1 Measuring perception of own body size

Multiple ways of assessing perception of own body size have been developed. The various different methodologies have been summarized and thoroughly critiqued in the literature (Farrell, Lee, & Shafran, 2005; Gardner & Brown, 2014). The simplest methods of body size estimation have involved asking participants to choose the body size most similar to their own from a set of different bodies (Stunkard, Sorensen, & Schulsinger, 1983; Viren Swami, Salem, Furnham, & Tovée, 2008). However, line drawings (Stunkard et al., 1983) do not provide realistic representations of human form, and photographic scales have used different individuals, who therefore vary in characteristics other than simply weight (Viren Swami et al., 2008).

More complex methods of body size estimation include the use of "analogue scales" such as moving calipers to demonstrate the size of body parts, and "optical distorting methods", such as distorting photographs, videos, or mirror images (Farrell et al., 2005; Gardner & Brown, 2014). Perhaps the most problematic aspect of these methods is that they are simply unwieldy to perform with large numbers of participants.

Most recently, 3D modelling software has been used to allow participants to adjust multiple dimensions in order to create their ideal body size (Crossley, Cornelissen, & Tovee, 2012). However, such avatars are not constrained by biological plausibility, and this limits their ecological validity, which has been highlighted as a potential limitation of body size estimation methods (Farrell et al., 2005).

1.3.2 Measuring satisfaction with body size

Multiple methods for assessing body satisfaction have also been developed. These involve questionnaire measures asking participants how satisfied they feel with various body parts (e.g., Body Shape Satisfaction Scale (Pingitore, Spring, & Garfield, 1997) or how their feelings about their shape and weight influence their behaviour (e.g., Shape and Weight subscales of the Eating Disorder Examination Questionnaire (EDE-Q) (C. G. Fairburn & Beglin, 1994). These questionnaires have good data demonstrating their validity and reliability (C. G. Fairburn & Beglin, 1994; Neumark-Sztainer et al., 2006; Pingitore et al., 1997), but there are limitations to questionnaire measures. These include the possibility that subjects may intentionally conceal their true views by responding in a way that they see as being socially acceptable, or as adhering to social norms. Furthermore, people may lack insight into their own beliefs, which can be a problem both generally (Nisbett & Wilson, 1977) and specifically in people with eating disorders (Konstantakopoulos, Tchanturia, Surguladze, & David, 2011).

More recent literature has proposed the use of tasks designed to assess implicit or unconscious beliefs, such as Implicit Association Tests (IAT) (Greenwald, McGhee, & Schwartz, 1998) to more fully understand people's views of their own body size. A small number of trials have used different versions of the IAT and related measures to explore implicit attitudes to underweight or overweight bodies in participants with and without eating disorders. Results suggest that in comparison with healthy controls, participants with eating disorders have more negative views of overweight bodies (Cserjesi et al., 2010; Spring & Bulik, 2014) and may (Ahern, Bennett, & Hetherington, 2008), or may not (Spring & Bulik, 2014) have more positive attitudes

towards underweight bodies. Smaller implicit ideal body size is also associated with higher levels of concern about body shape (Khan & Petroczi, 2015). One study has found that implicit and explicit measures of weight identity are associated in women (Grover, Keel, & Mitchell, 2003).

Another relatively unexplored way of studying satisfaction with own body size is to attempt to measure behavioural outcomes associated with body size satisfaction. This might be measured by choices about what clothes to wear, an important but under-researched part of women's appearance management (Tiggemann & Lacey, 2009). Qualitative research suggests that women are more likely to choose more closer fitting or revealing clothes when they feel more confident with their bodies (Grogan, Gill, Brownbridge, Kilgariff, & Whalley, 2013); and women who view themselves as overweight are more likely to report using clothing as "camouflage" (Tiggemann & Andrew, 2012).

Another such behavioural outcome is the amount a participant chooses to eat in an experimental situation; results from previous studies have been mixed as to whether body satisfaction is associated with eating more or less (Boyce, Kuijer, & Gleaves, 2013).

2. Eating disorders

Key to this thesis is the role of body dissatisfaction in the psychopathology of eating disorders. This section therefore describes the importance of eating disorders, their core psychopathology with particular regard to body dissatisfaction, and the evidence base with regard to current treatments. It highlights the need for novel treatments, and particularly those targeting body dissatisfaction.

2.1 Importance

Eating disorders are associated with high mortality: a meta-analysis found that in patients with anorexia nervosa, rates of death are 5.9 times higher than would be expected in an age and sex matched population, and in patients with bulimia nervosa and other eating disorders, rates are 1.9 times higher (Arcelus, Mitchell, Wales, & Nielsen, 2011). Eating disorders are relatively common: a Dutch community study of adolescents girls using DSM-5 criteria found a lifetime prevalence among 19 year old women of 5.7% compared to 1.2% in males (Smink, van Hoeken, Oldehinkel, & Hoek, 2014). Incidence rates of eating disorders in UK primary care increased from 32.3 per 100,000 in 2000 to 37.2 in 2009 (Micali, Hagberg, Petersen, & Treasure, 2013). Moreover, population studies consistently find that many people meeting diagnostic criteria do not receive any kind of treatment (around 56% in Holland (Smink et al., 2014); 28% in the US (Swanson, Crow, Le Grange, Swendsen, & Merikangas, 2011).

2.2 Classification, diagnosis, and core psychopathology

Eating disorders are a group of conditions in which negative beliefs about body shape, weight and eating, accompany behaviours including restricting eating, binge eating, excessive exercise, vomiting and laxative use. In anorexia nervosa (AN), weight is "less than minimally expected" and dietary intake is restricted. A key part of the diagnosis, and part of the diagnostic criteria, is that there are disturbances in the way that patients with AN experience their weight and shape, described in ICD-10 as "an overconcern with body shape and weight" (World Health Organisation, 1992) and in DSM 5 "self-evaluation is unduly influenced by body shape and weight" (American Psychiatric Association, 2013). Bulimia Nervosa (BN) comprises binge eating with

compensatory behaviours aimed at reducing weight, such as vomiting or excessive exercise. Here again, weight and shape is central in the diagnosis, with "self-evaluation…unduly influenced by body shape and weight" (American Psychiatric Association, 2013) and "overconcern with body shape and weight" (World Health Organisation, 1992). Atypical eating disorders, which make up more than half of diagnosed eating disorders, have been variously named "Eating Disorders Not Otherwise Specified" (EDNOS) (DSM 4), "Other Specified Feeding and Eating Disorders" (OSFED) (DSM 5) and "Atypical" eating disorders (ICD-10). Here, the patient fulfils some but not all of the full criteria for AN or BN.

The transdiagnostic approach to eating disorders (C. G. Fairburn & Harrison, 2003) takes the approach that "far more unites the…eating disorders than separates them" (C. G. Fairburn & Harrison, 2003). AN, BN (and the "Atypical" eating disorders) are "united by a distinctive core psychopathology…patients overevaluate their shape and weight" (C. G. Fairburn & Harrison, 2003). Adding weight to the argument that there is a common core psychopathology is the fact that patients often move between diagnostic sub-categories during the course of their illness.

One part of this core common central psychopathology shared between AN, BN and OSFED, of over-evaluation of shape and weight, is that this over-evaluation of their body shape and weight does not result in feeling positive about it: patients with eating disorders have high levels of body dissatisfaction (E. Stice, 2001).

2.3 Existing treatments

There has been a recent NICE review of the evidence for treatments for eating disorders in adults, children and adolescents (NICE, 2017). For AN in adults, NICE

recommends three possible treatments, in no particular order (as they found evidence of no difference between the three treatments): Cognitive Behavioural Therapy for Eating Disorders (CBT-ED), the Maudsley Model of Anorexia Nervosa Treatment for Adults (MANTRA) and Specialist Supportive Clinical Management (SSCM). They noted that study quality was either low or very low in all the studies comparing these treatments, and made a recommendation for research on the effective interventions for the treatment of anorexia nervosa. A recent large multi-centre randomised control trial comparing these three NICE recommended treatments found that only 60% of participants completed their treatment, and only 28.3% of participants were in remission from their illness at 12 month follow up (Byrne et al., 2017); rates of drop out and remission this low are not unusual for trials in this field.

For AN in children and young people, NICE recommended Family Therapy for Anorexia Nervosa, and where this is unacceptable, contraindicated, or ineffective, CBT-ED. Again they noted the low quality of the evidence of the treatment trials.

For BN, NICE recommended BN-focussed guided self-help, and if unacceptable, ineffective or contraindicated, CBT-ED. Again they described the evidence as being mostly of very low quality. For children and young people with BN they recommend Family Therapy for BN and if unacceptable, contraindicated or ineffective, CBT-ED, and again noted the low or very low quality of the evidence. They recommended research on how effective family therapies are in comparison with CBT-ED.

Psychotropic drugs are not recommended for the treatment of anorexia nervosa. A meta-analysis found no difference between placebo and either antidepressant or anti-psychotic medication on weight gain in anorexia nervosa (de Vos et al., 2014) and no

differences in anorexia nervosa symptomatology between anti-psychotics and placebo (Dold, Aigner, Klabunde, Treasure, & Kasper, 2015). There is very little evidence for pharmacotherapy in adolescents with bulimia nervosa, but there is some evidence to suggest that high dose antidepressants might improve outcomes in adults with bulimia nervosa (Bacaltchuk & Hay, 2003; Bacaltchuk, Hay, & Trefiglio, 2001; Mitchell, Roerig, & Steffen, 2013).

In view of this underwhelming evidence base for treatments, it is perhaps unsurprising to find that literature review suggests that remission rates in those with AN eight to ten years after diagnosis are around 50%; for BN they are around 75% (Keel & Brown, 2010). Relapse is also common: a recent systematic review found that between 10 and 50% of patients with anorexia nervosa relapse following recovery: the majority of studies reviewed found relapse rates of around 25% (Khalsa, Portnoff, McCurdy-McKinnon, & Feusner, 2017).

Undoubtedly, there is a need for a better evidence base for treatments for eating disorders in general. With regard to body dissatisfaction in particular, it appears to be quite resistant to current treatments. For example, a Dutch study (Exterkate et al., 2009) found that although there was an overall improvement in patients with AN, BN and EDNOS during the course of their treatment, they made little improvement in terms of their ongoing negative feelings about their body size. Untreated, body dissatisfaction predicts maintenance of and relapse to eating disorders (Jacobi et al., 2004). Meta-analysis of interventions aiming to improve body satisfaction in itself found that initial apparently moderate effect sizes became much smaller after

accounting for small sample sizes and publication biases (Alleva, Sheeran, Webb, Martijn, & Miles, 2015).

2.4 Novel potential treatments for body dissatisfaction

One recent paper (Gledhill et al., 2017) examined whether giving feedback on participants' judgement as to the size of others' bodies can reduce clinical eating disorder symptom scores. Forty participants with high baseline body dissatisfaction were shown a series of CGI images of women's bodies and asked to rate each of them as "fat" or "thin". They were then randomised to either complete a training task in which they were given feedback regarding whether their response was correct or not, with the aim of shifting the threshold where they perceived a body as fat, or a control task where the feedback was consistent with their baseline judgements. The training task successfully shifted their perception of the cut off between "fat" and "thin" bodies within the task, and was associated with decreased levels of eating disorder symptoms (shape, weight and eating concerns), which were sustained for two weeks. Results were replicated in 20 patients with atypical anorexia nervosa, raising the possibility that such interventions may be a useful novel adjunct in eating disorder treatment.

2.5 Perception of body size in eating disorders

From the literature, presence of a diagnosed eating disorder is not in itself predictive of greater overestimation of others' Body Mass Index (BMI). Differences between individuals in estimation of others' BMI seem to be driven by observer BMI rather than eating disorder diagnosis, with those with lower BMI overestimating other's BMI to a greater extent (Tovee, Emery, & Cohen-Tovee, 2000).

However, it has long been argued that in people with an eating disorder, the perceptual component of body dissatisfaction includes misperception and overestimation of own actual size (Farrell et al., 2005; Gardner & Brown, 2014). The review of the literature by Farrell (Farrell et al., 2005) found that of 52 studies published in this area between 1973 and 2002, around half found that those with anorexia nervosa overestimated their own body size, but half found either no such overestimation, or indeed an underestimation. However, the range of different methodologies used in the studies made it difficult to interpret whether the differences were in fact primarily due to variation in quality or type of assessment. The review of the literature from 2002-2013 (Gardner & Brown, 2014) found that nine new studies had been conducted. Of these new studies, seven found that those with AN overestimated their body size to a greater extent than control participants, although again a range of methodologies was used and there has been no attempt at meta-analysis of the results, or to investigate whether there is any apparent publication bias.

3. An overview of the thesis experiments and chapters

Chapter 2 concerns Studies 1 and 2, designed to test and replicate the effects of a single 15-minute intervention on perception of and satisfaction with own size. In each study, participants were randomised into three groups and exposed to one of three sets of images (photographs of women stretched to appear either slightly overweight, normal weight, or slightly underweight according to their randomisation group), with pre- and post- measures of perception of own and other's size, and satisfaction with own size.

Chapter 3 describes the development of the *Morphed Photographic Figure Scale*, a novel set of images of women modified to appear either under, over or neither over- nor underweight, using Psychomorph software (Tiddeman, 2001), and a novel *Outfit Choice Questionnaire* measure of body satisfaction. It then describes Study 3, an online study gathering population data on how these new images were perceived, and how the new measure of body satisfaction compared with existing measures.

Chapter 4 describes the development of novel implicit tasks to measure perception of and satisfaction with own size. Results are presented from Study 4a, comparing results from these newly-developed measures with existing measures of body satisfaction and eating disorder psychopathology.

Chapter 5 concerns Study 4b, a week-long study testing the effect of a five minute, twice-daily task on perception of and satisfaction with own size. Women were randomised into three groups and repeatedly completed a task using one of the three sets of the novel images of women described in Chapter 3 (photographs modified to appear either under, over or neither over- nor under-weight).

Chapter 6 concerns a Case Control study comparing perception of own and other's size, and satisfaction with own size, in a "Case" group of women with a current eating disorder, and a "Control" group of women without a current or previous eating disorder.

Chapter 7 draws together the results from the previous chapters and discusses their implications for future research and practice.

HYPOTHESES

1. That following exposure to images of "underweight", as opposed to "normal" or "overweight" bodies, women

   a. would judge others' bodies as larger, in both a single 15-minute intervention (Studies 1 & 2, Chapter 2) and in a week-long, repeated intervention (Study 4b, Chapter 5).
   b. would judge their own bodies as larger, in both a single 15-minute intervention (Studies 1 & 2, Chapter 2) and a week-long, repeated intervention (Study 4b, Chapter 5).
   c. would be less satisfied with their body size, in both a single 15-minute intervention (Studies 1 & 2, Chapter 2) and a week-long, repeated intervention (Study 4b, Chapter 5).
   d. would eat fewer snacks, in both a single 15-minute intervention (Studies 1& 2, Chapter 2) and a week-long, repeated intervention (Study 4b, Chapter 5).

2. That changes in perception of others' body size, own body size and satisfaction with own body size would persist to the following day (Study 2, Chapter 2).

3. That male participants, and those who were older, had higher body mass index (BMI), were more satisfied with their own weight, and did not have a history of an eating disorder would judge more photographs of women as under or normal weight, and that conversely women, and those with lower BMI, who were less satisfied with their own weight, and who had a history of an eating disorder would judge more photographs of women as being overweight (Study 3, Chapter 3).

4. That in healthy women, novel behavioural and implicit measures of body size perception and satisfaction with body size would correlate with existing questionnaire measures of own size and satisfaction with own size (Study 4a, Chapter 4).

5. That women with a current eating disorder (Cases), in comparison with Controls (Study 5, Chapter 5) would

a. judge more bodies as being overweight and fewer as underweight
b. overestimate their body size to a greater extent than controls
c. have a smaller ideal body size
d. feel less satisfied with their bodies across a range of existing and novel measures, including implicit measures.

CHAPTER 2: EFFECTS OF EXPOSURE TO BODIES OF DIFFERENT SIZES ON PERCEPTION OF AND SATISFACTION WITH OWN BODY SIZE: TWO RANDOMISED STUDIES

1. Introduction

As discussed in Chapter 1, body dissatisfaction is a risk factor for eating disorders (Jacobi et al., 2004), low mood, obesity and weight gain in adolescents and young adults of both genders (Haines, Neumark-Sztainer, Wall, & Story, 2007; Quick, Wall, Larson, Haines, & Neumark-Sztainer, 2013a; Sonneville et al., 2012; van den Berg & Neumark-Sztainer, 2007), as well as being a component of eating disorders which is both resistant to treatment (Alleva et al., 2015) and predictive of relapse (Jacobi et al., 2004).

One main component of body dissatisfaction concerns dissatisfaction with body size. If it were possible to alter perception of own body size, this could potentially provide a novel method of reducing body dissatisfaction and its negative consequences in both the general population and those with a current eating disorder.

Previous studies suggest that exposing people to images of bodies of different sizes can alter their perception of what constitutes a "normal" body size in others. Exposure to images of thinner as opposed to larger bodies leads to thinner bodies being judged most "normal", and vice versa. This has been show in both female (Brooks et al., 2016; Glauert et al., 2009; Hummel et al., 2013; Hummel et al., 2012; Winkler & Rhodes, 2005) and male bodies (Robinson & Christiansen, 2014; Robinson & Kirkham, 2014).

An important question leading on from this is whether exposure to bodies of different sizes can also alter perception of *own* size. As discussed at greater length in Chapter 1, research investigating this suggests that it may be possible, but the two studies conducted prior to the ones described here were very small (Hummel et al., 2013; Hummel et al., 2012). Relatedly, it is also important to understand whether exposure to tightly-controlled images varying only in size can alter *satisfaction* with own size. As discussed at greater length in Chapter 1, a reasonable body of evidence suggests that women feel less satisfied with their bodies after viewing photographs of thin models (Grabe et al., 2008). However, the studies in this area either have no control group, a control group comprising neutral images, a control group comprising women who differ on attractiveness as well as size, or have other methodological limitations such as an absence of baseline measures.

No studies to date have investigated the *duration* of any effect on perception of body size, or whether exposure to bodies of different sizes has an impact on any *behavioural* outcome, such as food consumption. Whether food consumption increases or decreases might enhance our understanding of how body dissatisfaction can influence subsequent eating disorders or weight gain.

This chapter therefore describes two studies. Study 1 comprised a group of 18-30 year old women with a BMI in the normal range. Study 2 was an attempt to replicate findings from Study 1 in a group of women with high baseline body dissatisfaction, more analogous to a clinical group. In each study, participants were randomised to see one of three weight categories of images of women's bodies ("under-", "normal" or "over-weight").

In both studies, the hypothesis was that following exposure to images of "underweight", as opposed to "normal" or "overweight" bodies, participants would (1) judge others' bodies as larger (replicating previous research), (2) judge their own bodies as larger (this was the primary outcome), and (3) be less satisfied with their body size. The further hypothesis was that increase in satisfaction with body size would lead to (4) increased chocolate consumption following the task, possibly due to decreased dissatisfaction with own weight. In Study 2, there was an additional hypothesis (5): that changes would persist to the following day.

2. Methods: Study 1

2.1 Participants

Female 18-25 year olds were recruited from the general population, by email invitation to a database of volunteers willing to be contacted about psychology research at the University of Bristol (consisting of University of Bristol students), and through advertisements on the University of Bristol Precinct. Potential participants completed a brief electronic screening questionnaire giving their age, sex, fluency in English, height, weight, and whether they had a previous diagnosis of an eating disorder; they were not asked for details of their ethnicity or socio-economic status.

Participants were eligible to take part if they were female, English (or fluent in English), had no history of a diagnosed eating disorder, and had a body mass index (BMI) of 19-25 kg/m2 (normal range). Participants were offered £5 of "Love to Shop" vouchers for their time.

2.2 Measures

Participants completed the Body Dissatisfaction subscale of the Eating Disorder Inventory (EDI) (Garner, 1991) online before attending the session. The Body Dissatisfaction subscale comprises nine items, and was scored as suggested for a non-clinical population (Schoemaker, van Strien, & van der Staak, 1994): four from 1 (Never) to 6 (Always) (e.g. "I think that my thighs are too large"), and five from 1 (Always) to 6 (Never) (e.g. "I think that my hips are just the right size"). The Body Dissatisfaction subscale of the EDI has an internal reliability of 0.90 in a clinical, and 0.94 in a non-clinical, sample (Schoemaker et al., 1994). Participants completed the PHQ-9 (Patient Health Questionnaire) (Kroenke, Spitzer, & Williams, 2001) during the session, before completing the task, as mood may impact on body dissatisfaction (Baker, Williamson, & Sylve, 1995; Taylor & Cooper, 1992). The PHQ-9 is a 9 item depression measure consisting of questions which score the DSM-IV depression criteria as "0" (not at all) to "3" (nearly every day). The scale has an internal reliability of 0.89.

Before and after completing the task, participants selected a number to represent the size of four normal-weight stimuli on the computer screen, on a scale from slightly underweight (1) to slightly overweight (7). Two of these stimuli were from the same set of images as the task stimuli (although not images which had been used in the tasks), and were dressed like the task stimuli in a sports bra and shorts, and two were fully clothed. Participants also completed Visual Analogue Scales (VAS) of own size, from 0 (too thin) to 10 (too fat), and satisfaction with own size, from 0 (not at all satisfied) to 10 (extremely satisfied), both before and after the task.

2.3 Stimuli

The stimuli used for the task were created from full body photographs of ten different women aged between 18 and 25 years with a BMI of 22 to 23 kg/m^2 and of White ethnicity, selected from a set of images of women with known BMI held in the School of Experimental Psychology, University of Bristol. These stimuli were used for the control group, "normal weight". The stimuli were modified by increasing the width of the images of women's bodies (initially 880 pixels) by 150 pixels to create an "overweight" group, and decreasing the width of the women's bodies by 150 pixels to create an "underweight" group. Images were cropped so the width of all photographs

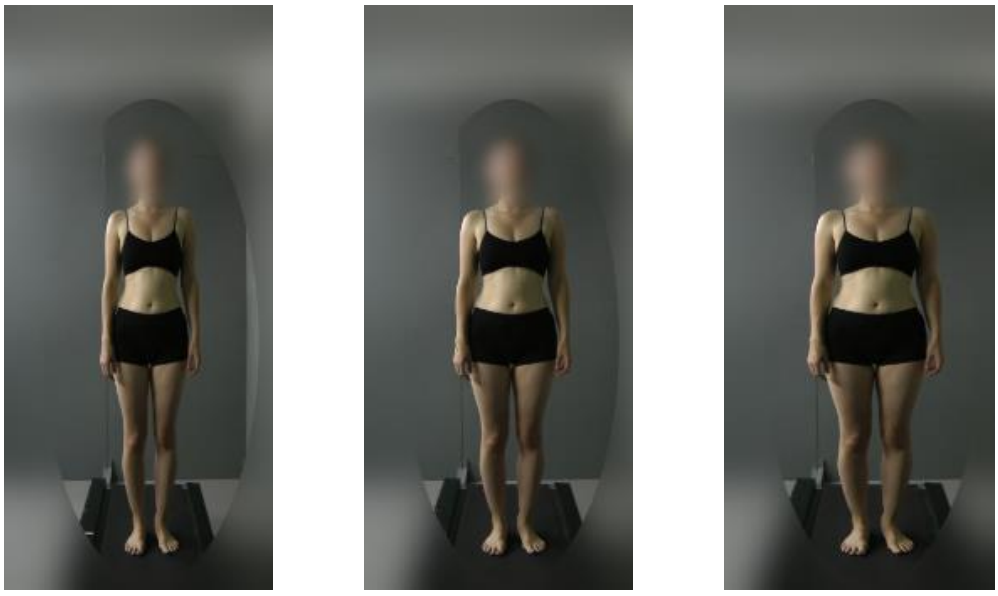


Figure 1: Example of stimuli used (left: "underweight"; middle: "normal weight"; right: "overweight")

remained 880 pixels. Faces were blurred to preserve the privacy of the women in the photographs, and to prevent subjects using faces to recognise the stimuli. An example of the stimuli used is given in Figure 1.

2.4 Procedure

Participants attended the University of Bristol, dressed in close-fitting clothing, as requested in the Participant Information Sheet and in the email inviting them. They provided informed consent, consenting to the tasks they would undertake, but remaining blind to the hypothesis and purpose of the study. They were tested individually, including by a colleague, RC (see Statement of Contributions). They were randomised to one of the three groups by a computer-generated random sequence to ensure balance across groups. KB, a colleague not involved in any testing or participant contact, assigned letters to the three training conditions (see Statement of Contributions); experimenters and participants were blind to which training condition was which.

Participants completed the pre-task measures described above. They then completed a single session of the training. Following this, they were asked to remove outer layers of clothing (e.g., cardigans, jumpers, jackets, coats), look at themselves in a full-length mirror, and then complete the post-task measures. The purpose of using the mirror was to ensure that the participant saw their own body in relation to the bodies to which they had been exposed. The experimenter then weighed them and measured their height without shoes. They were then asked to take a seat outside the room, the experimenter explaining "I need to check the computer has registered the tasks correctly", and they were offered a bowl containing 300g chocolate ("*Maltesers"*)

whilst they waited. *"Maltesers"* were chosen as they are a palatable snack food that most people like, and because they are quite small, so people might choose to eat a variable quantity. After 2-3 minutes, they were invited back into the room, the remaining chocolate was weighed and they were debriefed.

The training session consisted of a two-back working memory task. The stimuli (photographs of the 10 women adjusted to be either "over", "under" or "normal" (BMI 22-23kg/m^2) weight) were presented sequentially (and repeatedly) to participants, who were instructed to respond as to whether the image was a 'target' (i.e., had the same identity as the individual presented two trials previously) or a 'non-target' (i.e., had a different identity from the individual presented two trials previously). This was to ensure that participants paid attention to the images displayed, as attention has been shown to enhance training effects (Rhodes et al., 2011). In the "underweight" condition, the 10 images modified to appear "underweight" were displayed. In the overweight condition, the 10 images modified to appear "overweight" were displayed. In the control condition, the 10 images of "normal weight" women were displayed.

The tasks were programmed by KB (see Statement of Contributions) using E-Prime version 2.0 software (Psychology Software Tools Inc., Pittsburgh PA, USA). Each image was displayed for 2000 ms, with a 20 ms inter-trial-interval. To prevent low-level visual adaptation to the images, they were spatially jittered by an amount varying randomly from the centre of the screen by up to +/- 51 pixels in the x coordinate and +/- 38 pixels in the y coordinate (approximately 10% in each direction of the 1024 by 768 pixel display). Image size was 562 by 765 pixels, with a viewing distance of approximately 1 m. Failure to respond within the 2000 ms presentation resulted in no

response being recorded for that trial. Each task block consisted of 48 trials, in which 8 'targets' were presented. A total of 7 blocks, giving a total of 336 trials, were presented, taking approximately 15 minutes for each participant.

2.5 Statistical analysis

Data were checked to ensure that none were missing or incorrectly entered, with assistance from AA (see Statement of Contributions). The primary outcome was judgement of own body size (VAS from too thin [0] to too fat [10]) post-task, adjusted for pre-task score. The secondary outcome was participant satisfaction with their body size after the intervention (VAS from not at all satisfied [0] to extremely satisfied [10]). Other outcome measures included perception of size of the images of women with BMI 22-23kg/m^2 shown after the task, and weight of "*Maltesers"* taken by participants after finishing the task.

For each outcome measure, analysis was conducted using linear regression, treating randomisation group as the categorical exposure variable and adjusting for the participants' baseline score on each measure, by including it in the regression analysis. All analyses were repeated adjusting for the potential confounders of age, BMI and PHQ-9 score by including them in the regression analyses. The data are available with restricted access from the data.bris Research Data Repository (http:/data.bris.ac.uk/data) DOI: 10.5523/bris.rydgtj6qorrflndeh2e3c74g8, with permission.

Scores for individual measurements that were more than three standard deviations from the mean were judged to be outliers, and sensitivity analyses were conducted excluding such outliers.

Power calculation shows that three groups of 30 gave 90% power at a 5% alpha level to detect an effect size of d = 0.38.

2.6 Ethics

Ethics approval was sought and obtained from the Faculty of Science Research Ethics Committee at the University of Bristol (reference 270912588). Participants gave written consent to participate. The study was conducted according to Good Clinical Practice guidelines.

3. Results: Study One

3.1 Descriptive data

Ninety women were recruited between 11th November 2012 and 16th December 2013. Study recruitment is shown in the CONSORT diagram (Figure 2) and descriptive data is shown in Table 1.

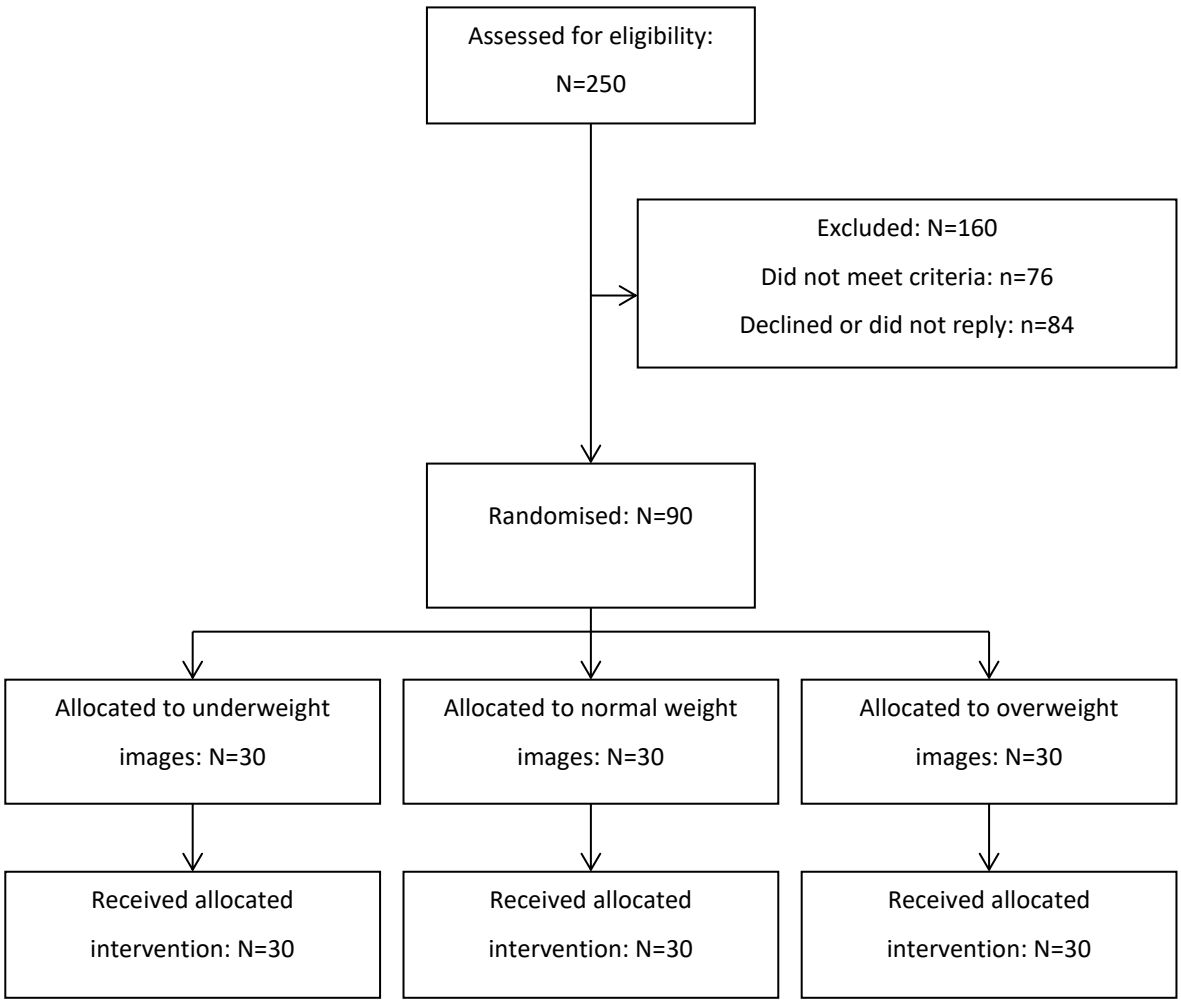


Figure 2: CONSORT diagram for Study 1

	All Groups (mean, SD)	Adapted to "underweight" (mean, SD)	Adapted to "normal weight" (mean, SD)	Adapted to "overweight" (mean, SD)
Age	20.4 (1.86)	20.6 (1.94)	20.1 (1.55)	20.3 (2.09)
Body Dissatisfaction score (0-54)	30.9 (7.2)	30.7 (5.44)	32.3 (7.71)	29.8 (8.20)
Measured height (m)	1.66 (0.06)	1.66 (0.06)	1.66 (0.07)	1.65 (0.05)
Measured weight (kg)	60.0 (6.52)	61.2 (4.99)	59.6 (7.72)	59.1 (6.60)
Measured BMI (kg/m^2)	21.8 (1.77)	22.3 (1.60)	21.5 (1.77)	21.7 (1.89)
PHQ-9 score	3.8 (3.05)	3.4 (2.39)	5.0 (3.60)	2.9 (2.73)
Baseline rating of size of computer images (1-7)	4.78 (0.84)	4.80 (0.72)	4.86 (0.89)	4.68 (0.90)
Post-training rating of size of computer images (1-7)	4.54 (0.93)	5.03 (0.84)	4.48 (0.82)	4.13 (0.92)
Baseline rating of own size (0-10)	5.99 (1.01)	5.87 (0.96)	6.12 (0.97)	5.99 (1.13)
Post-training rating of own size (0-10)	5.74 (1.07)	5.88 (1.09)	5.79 (0.97)	5.53 (1.15)
Baseline satisfaction with own size (0-10)	5.28 (2.06)	5.60 (2.06)	4.77 (1.64)	5.48 (2.37)
Post-training satisfaction with own size (0-10)	5.79 (1.96)	5.77 (2.12)	5.65 (1.61)	5.96 (2.14)
Chocolate consumed post-training(g)	7.71 (6.16)	6.10 (5.6)	8.30 (6.10)	8.73 (6.64)

Table 1: Descriptive data on groups at baseline & post-training (Study 1)

3.2 Perception of size of others' bodies (Hypothesis 1)

There was evidence of a difference between the means for the three groups in their judgment of the size of the images of other women's bodies shown following exposure to different sized images, with those exposed to "overweight" images during the task perceiving subsequent images as smaller than those exposed to normal or "underweight" images (p<0.001) (Figure 3). Adjustment for potential confounders (baseline PHQ-9, age, BMI) did not alter these results substantially. Full adjusted and unadjusted results for all measures are given in Table 2.

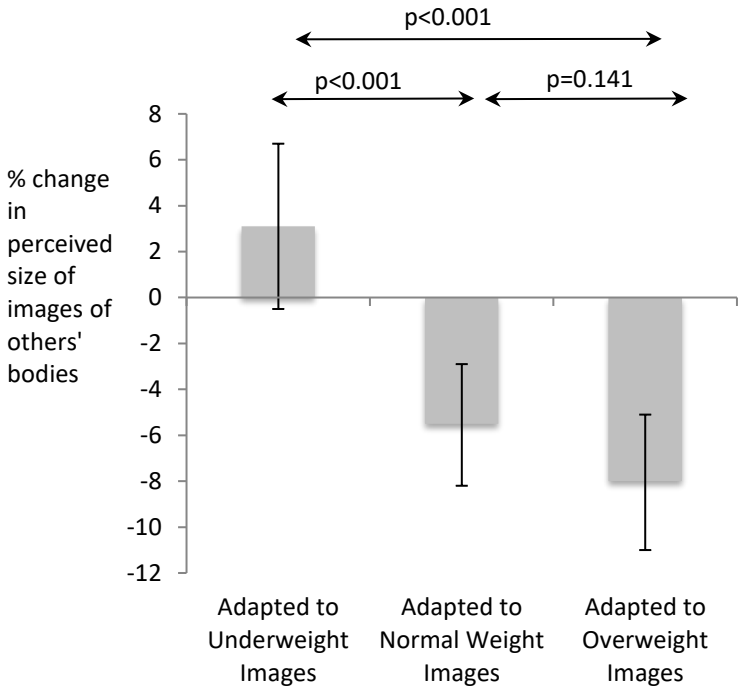


Figure 3: Mean percentage perceived change in size of others' bodies, pre to post task, by group (error bars represent 95% confidence intervals) – Study 1

3.3 Perception of own body size (Hypothesis 2)

There was only weak evidence to support a difference between the three groups in their mean judgment of their own size following exposure to different sized images (p=0.073) (Figure 4). Adjusting for potential confounders did not alter these results substantially.

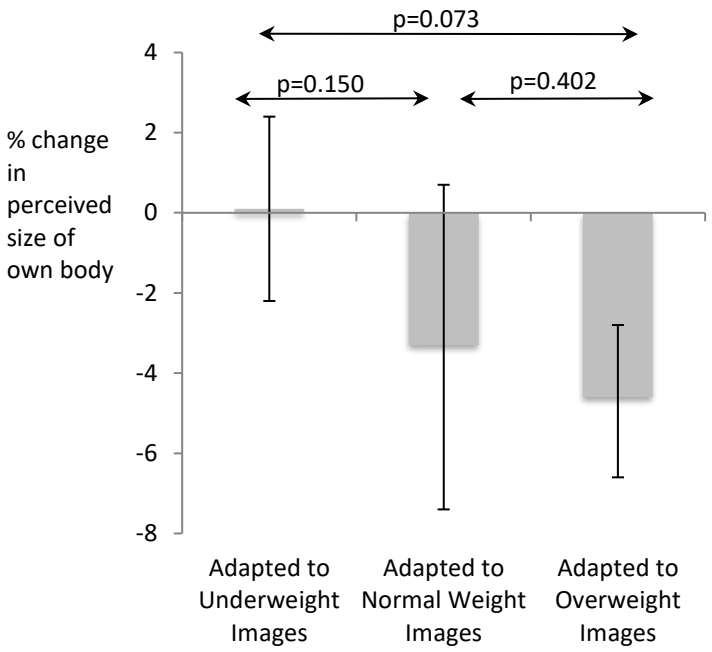


Figure 4: Mean percentage perceived change in size of own body, pre to post task, by group (error bars represent 95% confidence intervals) – Study 1

3.4 Satisfaction with own body size (Hypothesis 3)

There was evidence for group differences in mean satisfaction with own size following exposure to different sized images (p=0.046), with those exposed to "underweight" images being less satisfied than those exposed to "normal" or "overweight" images (Figure 5). Adjusting for potential confounders made no substantial difference to the results.

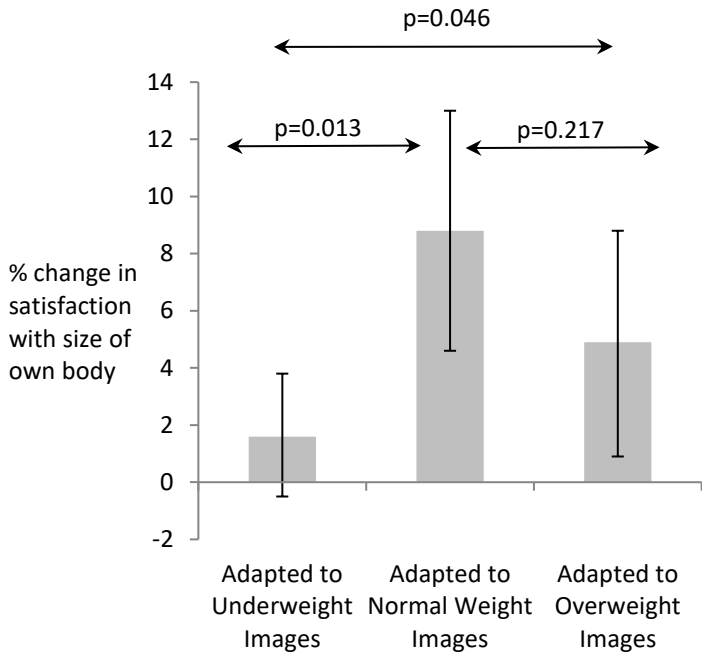


Figure 5: Mean percentage change in satisfaction with own body size, pre to post task by group (positive numbers indicate higher levels of satisfaction following task) (error bars represent 95% Confidence Intervals) – Study 1

3.5 Chocolate consumption (Hypothesis 4)

The weight of chocolate consumed was not normally distributed, so a binary variable of eating 'no' versus 'any' chocolate was created. Treating the three experimental groups as a categorical variable, logistic regression found no differences in chocolate consumption between groups (p=0.203) (Figure 6).

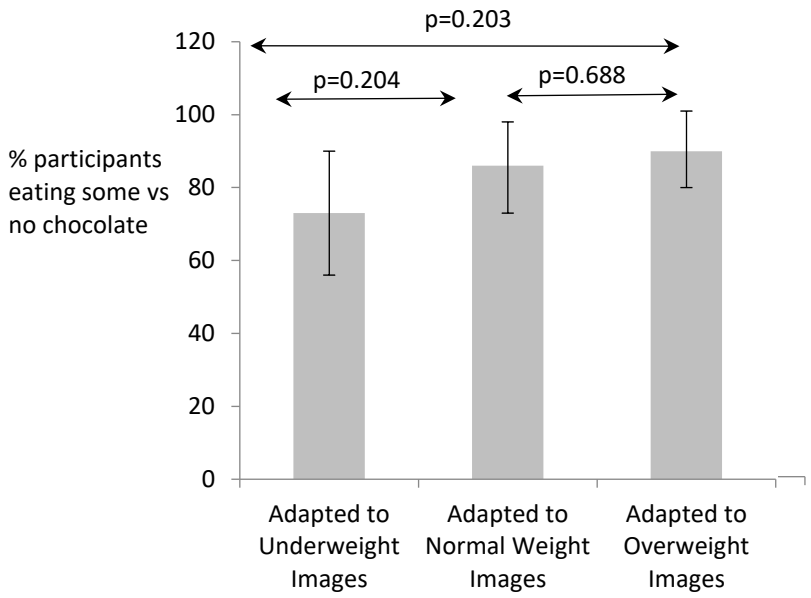


Figure 6: Percentage of participants eating some vs no chocolate following the task (error bars represent 95% Confidence Intervals) – Study 1

	Comparison with those adapted to normal weight images*			Comparison with those adapted to normal weight images, adjusted for age, BMI and PHQ-9 score*		
	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups
Perceived size of computer images (post-task rating adjusted for pre-task rating)	0.59 (0.30, 0.87), p<0.001	-0.21 (-0.50, 0.07), p=0.141	p<0.001	0.66 (0.37, 0.96), p<0.001	-0.17 (-0.46, 0.13), p=0.269	p<0.001
Perceived own size (post-task rating adjusted for pre-task rating)	0.28 (-0.10, 0.66), p=0.150	-0.16 (-0.54, 0.22), p=0.402	p=0.073	0.28 (-0.13, 0.68), p=0.180	-0.12 (-0.53, 0.28), p=0.546	p=0.132
Satisfaction with own size (post-task rating adjusted for pre-task rating)	-0.60 (-1.07, -0.13), p=0.013	-0.29 (-0.77, 0.18), p=0.217	p=0.046	-0.67 (-1.16, -0.18), p=0.008	-0.34 (-0.82, 0.15), p=0.170	p=0.028

**(Linear Regression Analyses with Randomisation Group as the categorical exposure variable)*

Table 2: Perception of size of others, own size and satisfaction with own size – Study 1

3.6 Sensitivity analysis excluding outliers

Analyses were repeated removing extreme individual measures which may not have been representative of the overall changes since they were more than three standard deviations from the mean. Removing outliers resulted in excluding four participants with scores for individual measurements that were more than three standard deviations from the mean: one participant who reported her own size at baseline as 10 on the 10-point VAS scale, one who reported a change in her perception of her own size of 4 points on the 10-point VAS scale, and two who reported an increase in satisfaction of 3.5 and 3.85 respectively on the 10-point VAS scale. Analyses were repeated as described above, but excluding these outliers (Table 3).

There were no differences in results for judgement of the size of the computer images before and after the task. In judgement of own size, the statistical evidence for a difference between groups strengthened considerably (p=0.008). There were no major differences in satisfaction with own size following removal of outliers. Adjusting for confounding variables did not substantially alter any results.

	Unadjusted comparison with those adapted to normal weight images*			Comparison with those adapted to normal weight images, adjusted for age, BMI and PHQ-9 score*		
	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups
Perceived size of computer images (post-task rating adjusted for pre-task rating)	0.59 (0.29, 0.88), p<0.001	-0.23 (-0.54, 0.07), p=0.126	p<0.001	0.66 (0.35, 0.96), p<0.001	-0.20 (-0.51, 0.12), p=0.213	p<0.001
Perceived own size (post-task rating adjusted for pre-task rating)	0.42 (0.09, 0.71), p=0.011	-0.04 (-0.36, 0.27), p=0.793	p=0.008	0.38 (0.05, 0.70), p=0.023	-0.03 (-0.36, 0.30), p=0.871	p=0.022
Satisfaction with own size (post-task rating adjusted for pre-task rating)	-0.54 (-0.97, -0.11), p=0.014	-0.31 (-0.75, 0.13), p=0.164	p=0.047	-0.61 (-1.06, -0.17), p=0.007	-0.33 (-0.79, 0.12), p=0.150	p=0.027

*(Linear Regression Analyses with Randomisation Group as the categorical exposure variable)

Table 3: Changes in perception of size of others, own size and satisfaction with own size after removing outliers – Study 1

4. Methods: Study Two

Study Two used the same methodology as that used in Study One, in a newly-recruited group with high baseline body dissatisfaction, in order to see whether results could be replicated in a group with more relevance to clinical populations. There were some small differences in methodology, which are outlined below. The study protocol was pre-registered on the Open Science Framework (https://osf.io/sngb6/).

4.1 Participants

Inclusion criteria required a baseline body dissatisfaction score of >35 on the Body Dissatisfaction subscale of the Eating Disorder Inventory (Garner, 1991); this value was chosen as it corresponded to the top quartile of scores on the EDI in Study 1. Criteria were also widened to include those with a reported body mass index (BMI) of 18 to 25kg/m^2, as it was anticipated that it would be harder to recruit sufficient numbers with a high body dissatisfaction score, and in Study 1 many participants were excluded because they had a BMI between 18 and 19. Reimbursement was increased to £10, due to the requirement to attend follow up the next day.

4.2 Measures and Stimuli

Measures were identical to Study One; and participants returned to complete the same measures at 24 hour follow up. The stimuli used were identical to Study One.

4.3 Procedure

The data collection was conducted by EL, HA, EB and MS, all of whom were co-supervised by IPV (see Statement of Contributions). Day 1 study procedure was identical, except that participants were not debriefed. When they attended their 24-hour follow-up, participants were asked to look at themselves in a full-length mirror, and then completed the same measures as previously. Before leaving, they repeated the computer task using the "normal" weight women, in order to reverse any possible increase in body dissatisfaction caused by exposure to "underweight" women. They were then debriefed.

4.4 Statistical analysis

Outcomes and analysis were identical to Study One, with additional analysis of results from 24-hour follow-up, to investigate whether between-group differences in perception of computer images, and perception of and satisfaction with own size were present after 24 hours. The data forming the basis of the results presented are available open access from the data.bris Research Data Repository (http:/data.bris.ac.uk/) DOI: 10.5523/bris.o1pqic3gep9n12c5mh0q0ojwj. A power calculation showed that three groups of 31 participants gives 90% power to detect between randomisation group differences of 0.5 on a 10-point scale at a significance level of 0.05.

A series of "mega-analyses" were also conducted, in which all the analyses were repeated, incorporating the results from both Study One and Study Two into each analysis. The analyses were identical to those described for the individual

studies, except for inclusion of a variable to indicate whether each data point was from Study One or Study Two.

In addition, the mega-analysis tested whether there was an association overall between change in perceived size of computer images and change in perceived own size, and between change in perceived own size and satisfaction with own size.

4.5 Ethics

Ethics approval was obtained from the Faculty of Science Research Ethics Committee at the University of Bristol (reference 2706148085).

5. Results: Study Two

5.1 Descriptive data

96 women were recruited between 29th July 2014 and 9th August 2015 (see CONSORT diagram Figure 7). Four participants were excluded as they had been erroneously recruited with Body Dissatisfaction scores ≤35, despite inclusion criteria requiring a score >35. One participant withdrew her consent following participation. Table 4 shows descriptive data.

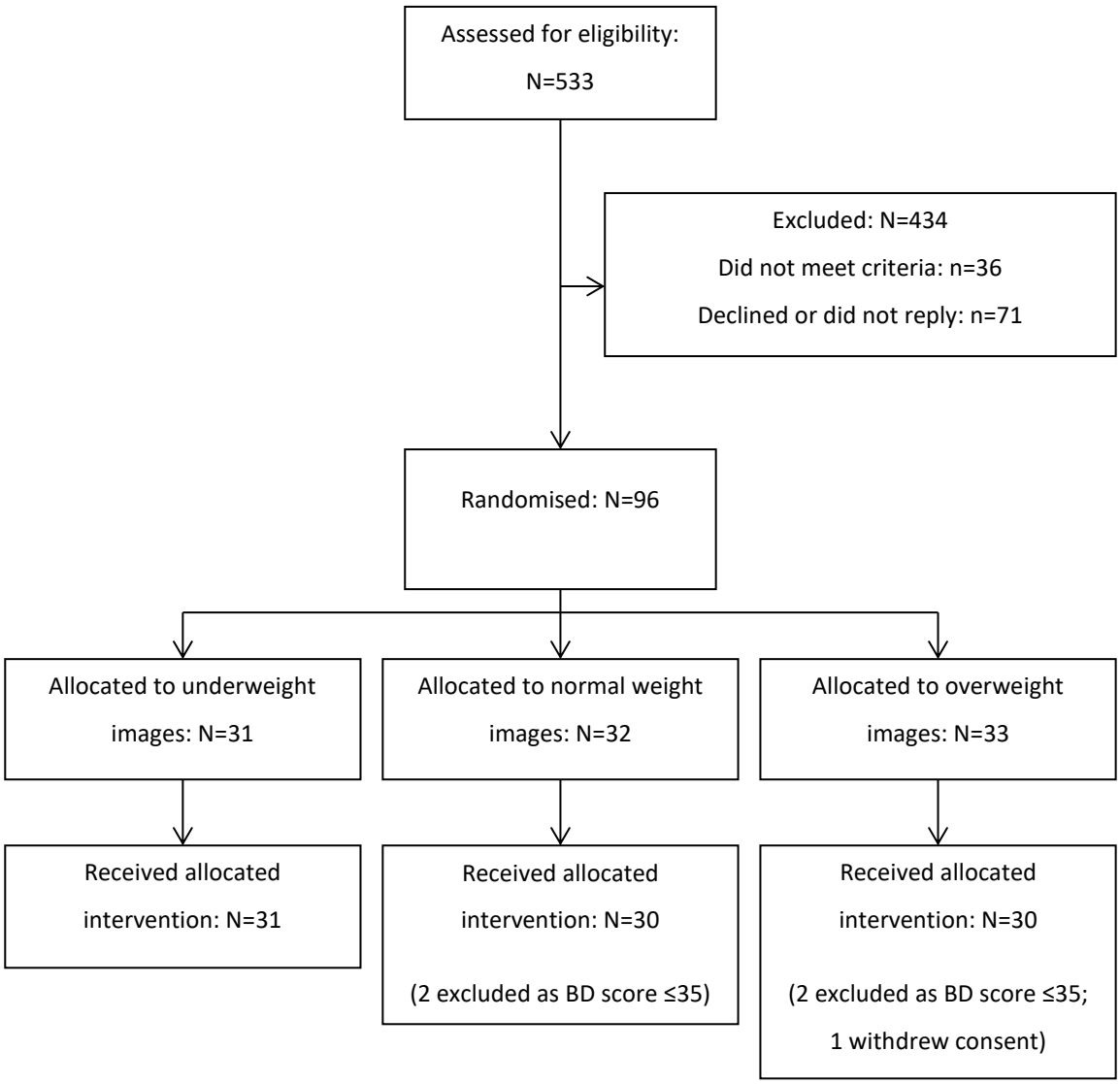


Figure 7: CONSORT diagram for Study 2

	All Groups (mean, SD)	Adapted to "underweight" (mean, SD)	Adapted to "normal weight" (mean, SD)	Adapted to "overweight" (mean, SD)
Age	21.4 (2.03)	21.5 (2.08)	21.2 (2.16)	21.5 (1.89)
Body Dissatisfaction score (0-54)	39.4 (3.89)	40.0 (4.94)	39.1 (3.06)	39.1 (3.43)
Measured height (m)	1.66 (0.07)	1.65 (0.08)	1.68 (0.07)	1.65 (0.06)
Measured weight (kg)	60.5 (7.04)	60.9 (8.26)	61.7 (7.77)	60.1 (6.47)
Measured BMI (kg/m^2)	22.1 (1.85)	22.3 (1.98)	21.9 (1.84)	22.1 (1.77)
PHQ-9 score	6.5 (4.17)	7.65 (4.83)	6.47 (4.31)	5.3 (2.90)
Baseline rating of size of computer images (1-7)	5.09 (0.99)	4.94 (0.91)	5.19 (1.04)	5.14 (1.05)
Post-training rating of size of computer images (1-7)	4.57 (0.91)	4.85 (0.77)	4.69 (0.64)	4.16 (1.12)
Day 2 rating of size of computer images (1-7)	4.83 (0.77) (N=84)	5.02 (0.75) (N=28)	5.0 (0.63) (N=26)	4.52 (0.82) (N=30)
Baseline rating of own size (0-10)	6.73 (1.01)	7.01 (0.96)	6.85 (1.02)	6.33 (0.96)
Post-training rating of own size (0-10)	6.26 (1.20)	6.82 (1.03)	6.27 (1.22)	5.68 (1.09)
Day 2 rating of own size (0-10)	6.58 (1.04) (N=84)	6.89 (0.97) (N=28)	6.73 (1.10) (N=26)	6.15 (0.91) (N=30)
Baseline satisfaction with own size (0-10)	3.92 (1.68)	3.72 (1.87)	3.85 (1.62)	4.20 (1.56)
Post-training satisfaction with own size (0-10)	4.45 (1.71)	3.81 (1.66)	4.32 (1.55)	5.25 (1.63)
Day 2 satisfaction with own size	3.99 (1.58) (N=84)	3.54 (1.53) (N=28)	3.73 (1.50) (N=26)	4.63 (1.54) (N=30)
Chocolate consumed post-training(g)	6.00 (7.57)	3.9 (4.85)	8.48 (9.15)	5.86 (7.82)

Table 4: Descriptive data on groups at baseline & post-training – Study 2

5.2 Perception of size of others' bodies (Hypothesis 1)

There was evidence of a difference between the means for the three groups in their judgment of the size of the images of others' bodies they saw following exposure to different sized images, consistent with Study 1. Those exposed to "overweight" images during the task viewed subsequent images as smaller than those shown normal or "underweight" images ($p<0.001$) (Figure 8; detailed results from Study 2 are also shown in Table 5). Adjusting for potential confounders (baseline PHQ-9, age, BMI) did not alter these results substantially.

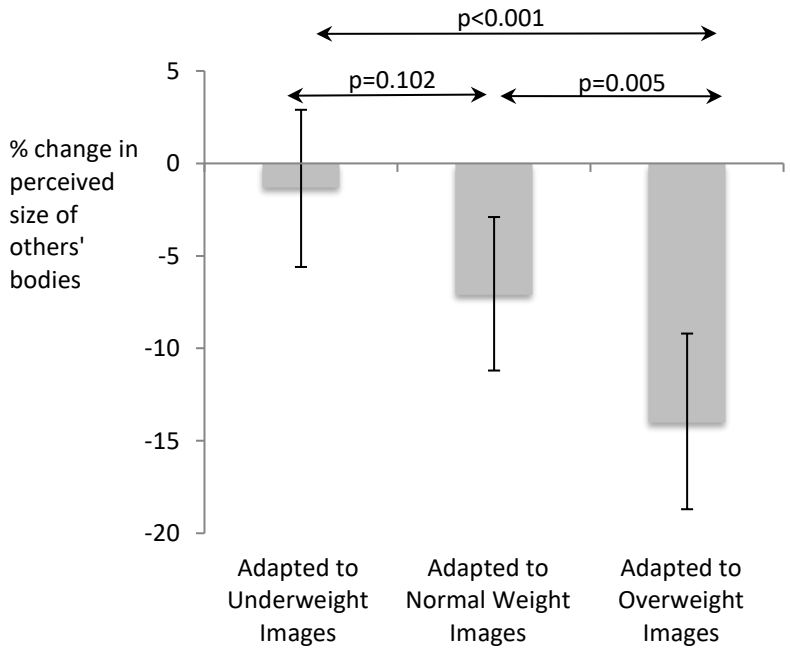


Figure 8: Mean percentage perceived change in size of others' bodies, pre to post adaptation, by group (error bars represent 95% confidence intervals) – Study 2

5.3 Perception of own body size (Hypothesis 2)

There was evidence of a difference between the means for the three groups in their judgment of their own size following exposure to different sized images, with those shown "overweight" images subsequently judging their own body as smaller than those shown "normal" or "underweight" images (p=0.018) (Table 5 and Figure 9). Adjusting for potential confounders did not alter these results substantially.

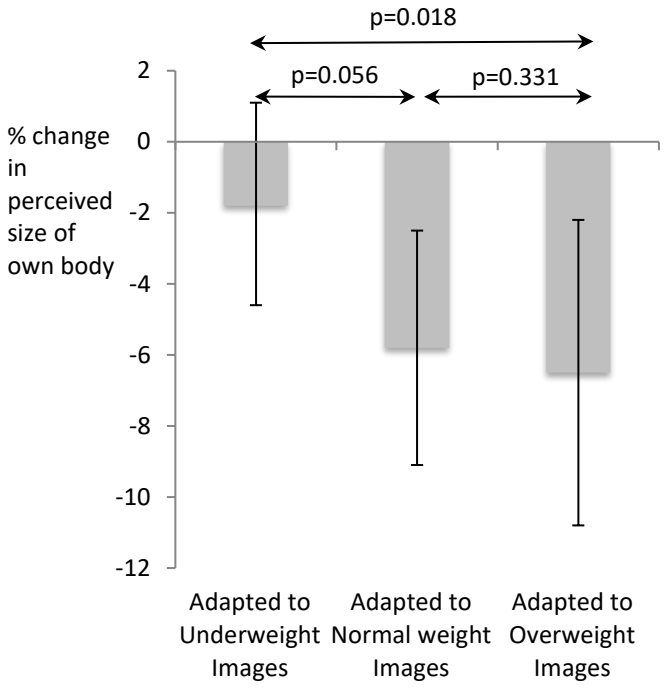


Figure 9: Mean percentage perceived change in own body size, pre to post task, by group (Error bars represent 95% Confidence Intervals) – Study 2

5.4 Satisfaction with own body size (Hypothesis 3)

There was evidence for group differences in mean satisfaction with own size following exposure to different sized images, with those shown overweight images subsequently being more satisfied with their own bodies than those shown "underweight" or "normal weight" images (p=0.004) (Figure 10, Table 5). Adjusting for potential confounders made no substantial difference to the results.

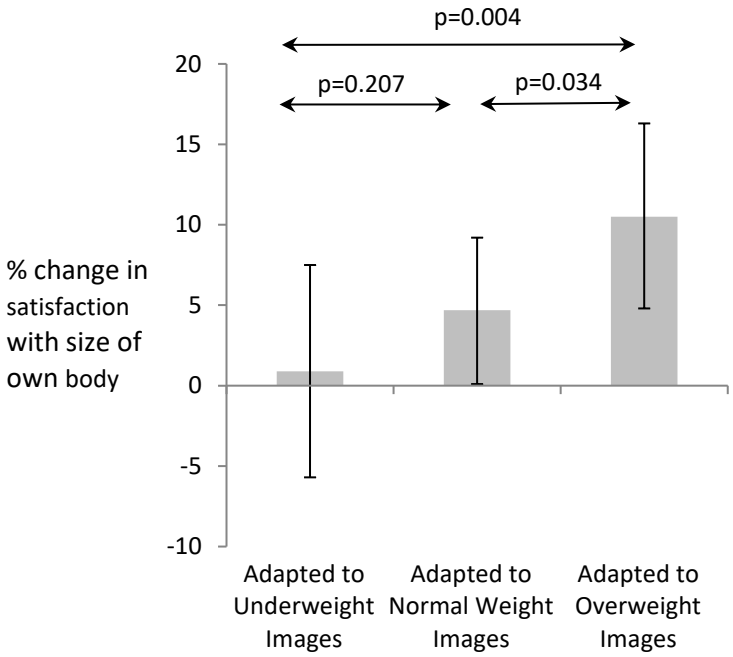


Figure 10: Mean percentage change in satisfaction with own body size, pre to post task by group (positive numbers indicate higher levels of satisfaction following task) (error bars represent 95% Confidence Intervals) – Study 2

5.5 Chocolate consumption (Hypothesis 4)

The weight of chocolate consumed was not normally distributed, so a binary variable of eating 'no' chocolate versus eating 'any' chocolate was created. Treating the three groups as a categorical variable, logistic regression analysis found weak evidence for between-group differences in chocolate consumption, with those in the "underweight" and "overweight" groups both eating less chocolate than those in the "normal weight" group (p=0.081) (Figure 11, Table 5).

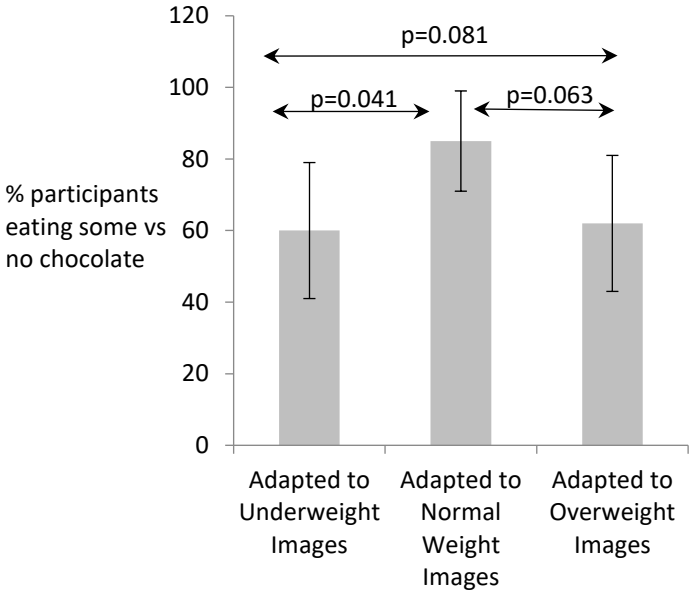


Figure 11: Percentage of participants eating some vs no chocolate following the task (error bars represent 95% Confidence Intervals)

5.6 Twenty-four hour follow-up (Hypothesis 5)

There was some loss to follow up: 84 of the 91 participants (92%) attended follow up at 24 hours, of whom 26 were in group shown "normal weight" images, 28 in the group shown "underweight" images and 30 in the group shown "overweight" images. Analysis using only data from those participants who attended on both days, found that between group differences in perceived size of computer images persisted at 24 hour follow-up ($p<0.001$), with differences in the same direction as in the post-task follow-up (Table 5). Results persisted following adjustment for confounding variables. There was no evidence that changes in perception of own size persisted after 24 hours ($p=0.311$). Differences in satisfaction with own size did persist ($p=0.032$), although the strength of the evidence for this difference reduced following adjustment for confounders ($p=0.080$). Repeating analyses including all participants, carrying forward the last available observation (i.e. those collected immediately after the task) in those who did not attend follow up, did not substantially alter these results.

(Linear Regression Analyses with Randomisation Group as the categorical exposure variable)

	Comparison with those adapted to normal weight images*			Comparison with those adapted to normal weight images, adjusted for age, BMI and PHQ-9 score*		
	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups	Adapted to underweight images	Adapted to overweight images	Test of Trend across all groups
Perceived size of computer images (post-task rating adjusted for pre-task rating)	0.29 (-0.06, 0.64), p=0.102	-0.51 (-0.86, -0.15), p=0.005	p=0.001	0.33 (-0.03, 0.68), p=0.069	-0.48 (-0.84, -0.13), p=0.008	p<0.001
Perceived own size (post-task rating adjusted for pre-task rating)	0.44 (-0.01, 0.90), p=0.056	-0.23 (-0.70, 0.24), p=0.331	p=0.018	0.40 (-0.06, 0.85), p=0.085	-0.30 (-0.77, 0.17), p=0.215	p=0.018
Satisfaction with own size (post-task rating adjusted for pre-task rating)	-0.44 (-1.12, 0.25), p=0.207	0.75 (0.06, 1.44), p=0.034	p=0.004	-0.34 (-1.03, 0.35), p=0.329	0.72 (0.03, 1.41), p=0.042	p=0.012
Perceived size of computer images at 24 hour follow up (adjusted for pre-task rating) (N=84)	0.12 (-0.18, 0.42), p=0.413	-0.47 (-0.77, -0.18), p=0.002	p<0.001	0.17 (-0.13, 0.47), p=0.271	-0.47 (-0.76, -0.18), p=0.002	P<0.001
Perceived own size at 24 hour follow up (adjusted for pre-task rating) (N=84)	-0.05 (-0.39, 0.49), p=0.823	-0.27 (-0.71, 0.17), p=0.229	p=0.311	-0.05 (-0.49, 0.39), p=0.825	-0.33 (-0.77, 0.11), p=0.143	p=0.286
Satisfaction with own size at 24 hour follow up (adjusted for pre-task rating) (N=84)	-0.18 (-0.89, 0.53), p=0.616	0.70 (-0.002, 1.40), p=0.051	p=0.032	0.002 (-0.69, 0.69), p=0.994	0.67 (-0.01, 1.35), p=0.053	p=0.080

Table 5: Perception of size of others, own size and satisfaction with own size - Study 2

5.7 Sensitivity analysis excluding outliers

Removing outliers resulted in excluding five participants with scores for individual measurements that were more than three standard deviations from the mean (one scoring 10 and one scoring 3 for pre-test own size, two with change in size scores of -3.9 and +2 and one with a mean post-task score for the computer images of 1). Analyses were repeated excluding these outliers. There was no difference in results for judgement of the size of the computer images before and after the task. In judgement of own size, statistical evidence for a difference between groups strengthened considerably (p=0.005) (Table 6). Adjusting for confounding variables did not substantially alter any of these results.

	Unadjusted comparison with those adapted to normal weight images			Comparison with those adapted to normal weight images, adjusted for age, BMI and PHQ-9 score		
	Adapted to underweight images	Adapted to overweight images	Test of Trend across all groups	Adapted to underweight images	Adapted to overweight images	Test of Trend across all groups
Perceived size of computer images (post-task rating adjusted for pre-task rating)	0.33 (0.03, 0.63), p=0.034	-0.39 (-0.70, -0.07), p=0.017	p<0.001	0.38 (0.08, 0.69), p=0.015	-0.34 (-0.66, -0.03), p=0.033	p<0.001
Perceived own size (post-task rating adjusted for pre-task rating)	0.50 (0.09, 0.90), p=0.017	-0.19 (-0.62, 0.23), p=0.364	p=0.005	0.48 (0.07, 0.89), p=0.021	-0.26 (-0.70, 0.17), p=0.228	p=0.003
Satisfaction with own size (post-task rating adjusted for pre-task rating)	-0.48 (-1.15, 0.20), p=0.161	0.57 (-0.13, 1.28), p=0.110	p=0.014	-0.39 (-1.08, 0.30), p=0.262	0.54 (-0.18, 1.25), p=0.140	p=0.043
Perceived size of computer images (at 24 hour follow up, adjusted for pre-task rating)	0.13 (-0.15, 0.42), p=0.353	-0.56 (-0.85, -0.27), p<0.001	P<0.001	0.17 (-0.11, 0.46), p=0.234	-0.56 (-0.85, -0.27), p<0.001	p<0.001
Perceived own size (at 24 hour follow up, adjusted for pre-task rating)	0.11 (-0.32, 0.55), p=0.605	-0.33 (-0.77, 0.11), p=0.143	p=0.127	0.01 (-0.43, 0.44), p=0.973	-0.40 (-0.85, 0.04), p=0.075	p=0.117
Satisfaction with own size (at 24 hour follow up, adjusted for pre-task rating)	-0.23 (-0.96, 0.50), p=0.530	0.73 (-0.01, 1.47), p=0.054	p=0.028	-0.03 (-0.74, 0.69), p=0.940	0.74 (0.02, 1.47), p=0.044	p=0.058

*(Linear Regression Analyses with Randomisation Group as the categorical exposure variable)

Table 6: Changes in perception of size of others, own size and satisfaction with own size after removing outliers – Study 2

5.8 Mega-analysis of Studies One and Two

Combining the results from Studies One and Two strengthened the findings throughout, and there was strong statistical evidence in support of hypotheses 1 to 3 (Table 7). Adjusting for confounders, and repeating the analyses excluding outliers (as defined separately for Study 1 and Study 2) did not change these findings (Table 8).

There was an overall association between change in perceived size of computer images of other women and change in perceived own size (0.238 (95% CI 0.080, 0.397), p=0.003. There was also an association between change in perceived own size, and change in satisfaction with own size, with satisfaction increasing as perceived size decreased (-0.295 (95% CI -0.384, -0.205), p<0.001).

	Comparison with those adapted to normal weight images*			Comparison with those adapted to normal weight images, adjusted for age, BMI and PHQ-9 score*		
	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups	Adapted to underweight images	Adapted to overweight images	Test of Trend across all groups
Perceived size of computer images (post-task rating adjusted for pre-task rating)	0.44 (0.22, 0.68), p<0.001	-0.37 (-0.60, -0.14), p=0.002	p<0.001	0.50 (0.27, 0.72), p<0.001	-0.34 (-0.58, -0.12), p=0.003	p<0.001
Perceived own size (post-task rating adjusted for pre-task rating)	0.36 (0.06, 0.65), p=0.018	-0.19 (-0.48, 0.11), p=0.217	p=0.001	0.31 (0.01, 0.60), p=0.04	-0.20 (-0.50, 0.10), p=0.197	p=0.004
Satisfaction with own size (post-task rating adjusted for pre-task rating)	-0.45 (-0.88, -0.03), p=0.038	0.24 (-0.19, 0.67), p=0.277	p=0.006	-0.44 (-0.87, -0.01), p=0.047	0.18 (-0.26, 0.61), p=0.43	p=0.017

*(Linear Regression Analyses with Randomisation Group as the categorical exposure variable)

Table 7: Perception of size of others, own size and satisfaction with own size (Studies 1 & 2 combined)

	Comparison with those adapted to normal weight images*			Comparison with those adapted to normal weight images, adjusted for age, BMI and PHQ-9 score*		
	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups	Adapted to underweight images	Adapted to overweight images	Test of Trend across all groups
Perceived size of computer images (post-task rating adjusted for pre-task rating)	0.47 (0.26, 0.68), p<0.001	-0.32 (-0.54, -0.10), p=0.004	p<0.001	0.52 (0.31, 0.73), p<0.001	-0.30 (-0.52, -0.08), p=0.007	p<0.001
Perceived own size (post-task rating adjusted for pre-task rating)	0.43 (0.18, 0.68), p=0.001	-0.12 (-0.38, 0.15), p=0.383	p=0.001	0.40 (0.14, 0.65), p=0.002	-0.14 (-0.41, 0.12), p=0.29	p<0.001
Satisfaction with own size (post-task rating adjusted for pre-task rating)	-0.42 (-0.83, 0.00), p=0.050	0.18 (-0.25, 0.61), p=0.41	p=0.017	-0.43 (-0.84, -0.07), p=0.046	0.10 (-0.33, 0.54), p=0.64	p=0.034

*(Linear Regression Analyses with Randomisation Group as the categorical exposure variable)

Table 8: Changes in perception of size of others, own size and satisfaction with own size after removing outliers (Studies 1 & 2 combined)

6. Discussion

6.1 Main findings

There was evidence in both Study 1 and Study 2 in support of Hypothesis 1, that exposing women to images of women's bodies of different sizes alters their perception of the size of subsequent bodies, with exposure to larger bodies leading to subsequent perception of other bodies as smaller. This is in line with findings from previous studies (Brooks et al., 2016; Glauert et al., 2009; Hummel et al., 2013; Hummel et al., 2012; Robinson & Christiansen, 2014; Robinson & Kirkham, 2014; Sturman et al., 2017; Winkler & Rhodes, 2005).

In Study 2, and in the mega-analysis of both studies, there was good evidence in support of Hypothesis 2, that exposing participants to images of "overweight" as opposed to "underweight" or "normal weight" bodies leads to them perceiving their own size as smaller. There was also evidence in both studies that satisfaction with own size changes following exposure to images of women's bodies of different sizes, with exposure to larger bodies resulting in increased satisfaction with own size. This finding was stronger in Study 2, in which participants were dissatisfied with their bodies at baseline. Differences in satisfaction with own body size, and perception of body size of others persisted to 24-hour follow up, although differences in perception of own size did not persist. There was no overall effect of exposure to images of different sizes on chocolate consumption, which may be due to individual variation in whether a change in satisfaction with own size leads to the consumption of more or less chocolate.

Studying the differences between groups more closely, changes in perception of own and others' size appear to be predominantly driven by decreases in perception of own size following exposure to normal or overweight images. Exposure to "underweight" images had less effect on subsequent perception of own and others' size. This may be because the "underweight" images used were in fact very similar in size to the media images surrounding us, and may be larger than average media images. For example, a study of Internet models found that over 50% had a BMI of ≤18 (Owen & Laurel-Seller, 2000). It is likely that body size norms (Rhodes et al., 2013) are strongly affected by previous exposure to media images, so further exposure to "underweight" images may make little difference.

6.2 Strengths and Limitations

These studies build on previous research in demonstrating in a larger sample than that previously used, that exposure to bodies of different sizes can alter perception of and satisfaction with own body, and that this occurs in all women, and to a greater extent in those with high baseline body dissatisfaction. Experimenters and participants were blinded and the study was tightly controlled, in that all groups saw images of the same women, manipulated to appear in different sizes. The second study replicated the methods of the first, and was preregistered with defined hypotheses and primary outcome measures.

However, there were some limitations to the study. Firstly, the images created were not very sophisticated, merely varying the overall width of the bodies rather than altering their proportions in any other way, and the images were not piloted

to see whether they were viewed as realistic. However, the images are more sophisticated than those used in some previous studies, in that they were photographs of real women, rather than 2D or 3D sketches. Secondly, at baseline, participants rated images with a BMI of 22 to 23kg/m^2 as being towards the overweight end of the scale (mean of 4.78 on a 7-point scale (Study 1); 5.09 (Study 2)), despite this BMI being in the middle of the healthy range of BMI. Therefore, those completing the "normal" weight version of the task may have been exposed to images they considered to be slightly overweight. This does appear to be the case: exposure to "normal" weight images led to viewing subsequent computer images as smaller, perceiving own size as smaller and being more satisfied with own size. Thirdly, the choice of chocolate as a snack may not have been ideal, both because it is a food that some people crave, and because the craving often involves both approach and avoidance, as well as feelings of guilt (Cartwright & Stritzke, 2008). Fourth, the purpose of the study may have been apparent to participants, and results may have been affected by demand characteristics. Finally, participants were almost all students at the University of Bristol, and they therefore may not be representative of the national population of 18-25 year olds.

6.3 Future directions

It was important to address the limitations of Studies 1 and 2 in order to develop an intervention that could eventually be trialled as an adjunctive treatment in women with current eating disorders.

In order to design such an intervention, it was first necessary to create and test a set of images of women's bodies which were ecologically valid, in that they had the appearance of being of different weights, rather than simply having been stretched to make them wider or narrower. It was also necessary to gather data from a large sample of women on how such novel images were perceived in terms of weight, as results from Studies 1 and 2 suggested that using a healthy BMI might not be a good way of selecting a set of images viewed as being of "normal" weight.

Furthermore, it would be useful to gather more information on the effect of perception of and satisfaction with body size on behavioural outcomes, in particular exploring the idea of clothing choice as a possible behavioural measure of satisfaction with body size.

Therefore, Chapter 3 describes the process of designing and validating a novel set of ecologically valid images of women of different weights, and of a new *Outfit Choice Questionnaire*.

CHAPTER 3: DESIGNING AND VALIDATING THE *MORPHED PHOTOGRAPHIC FIGURE SCALE* AND THE *OUTFIT CHOICE QUESTIONNAIRE*

1. Introduction

Following the promising results described in Chapter 2, the aim of the next study was to design ways to overcome the limitations with Studies 1 and 2. In order to do this, it was necessary to create a set of images of women's bodies that varied only in terms of weight, rather than simply using bodies that had been stretched or made narrower. Ideally such images would keep other features of the women, such as leg-to-body ratio, constant, but would show realistic changes in shape and body texture, including fat deposits and muscle tone, at various different weights. A set of images, the *Morphed Photographic Figure Scale,* was created, with technical support from AS (see Statement of Contributions).

The second aim of Study 3 was to create a behavioural measure of satisfaction with own body. Clothing choice is viewed as being an important but rather under-researched area in understanding how women manage their appearance (Tiggemann & Lacey, 2009). A small qualitative literature suggests that women are more likely to choose more closer fitting or revealing clothes when they feel more confident with their bodies (Grogan et al., 2013); and that women who view themselves as overweight are more likely to report using clothing as "camouflage" (Tiggemann & Andrew, 2012). A set of 60 images of different outfits was created, as described in greater detail in the Methods section, in order to develop a quantitative measure of clothing choice, the *Outfit Choice Questionnaire*.

It was then necessary to validate both the *Morphed Photographic Figure Scale* and the *Outfit Choice Questionnaire*, to gather data on how people perceived the weight status of the new images of women's bodies, and to validate the *Outfit Choice Questionnaire* against existing questionnaire measures of body satisfaction. This required the online rating Study 3.

In Study 3, participants were asked to rate each of the 90 morphed images in the *Morphed Photographic Figure Scale* as 'over-weight', 'under-weight', or 'neither over or under weight'. As well as gathering data about how each image was perceived, data from this study were also used to test for between group differences in judgments about weight status. The hypotheses were that male subjects, and those who were older, had higher body mass index (BMI), were more satisfied with their own weight, and did not have a history of an eating disorder would be more likely to perceive more images of women as being under or normal weight, and that conversely women, and those with lower BMI, who were less satisfied with their own weight, and who had a history of an eating disorder would be more likely to perceive more images as being overweight.

Participants were also asked to rate the outfits in the *Outfit Choice Questionnaire* in terms of how comfortable with their bodies they would feel wearing each of them. The hypothesis was that scores on this measure would correlate with existing measures of body satisfaction, and that factor analysis might enable satisfaction with different parts of the body to be assessed separately.

2. Methods

2.1 Participants

Participants were recruited via email and poster advertisements to University of Oxford students and staff. No eligibility criteria were set, because the online nature of the study meant that it was not possible to verify participant data; instead, participants were asked to give their demographic details. Everyone who completed the study was offered the opportunity to be entered into a prize draw to win an iPad Air; personal data to enable the winner to be contacted was separated from questionnaire responses, which were anonymised on collection.

2.2 Morphed Photographic Figure Scale

2.2.1 Scale Creation

The *Morphed Photographic Figure Scale* was created using PsychoMorph software by AS (see Statement of Contributions) (Skinner, 2017). In brief, full body photographs of women aged between 18 and 25 years, were selected from a set of images of women with known BMI held in the School of Experimental Psychology, University of Bristol. Two "reference" images were created, one "underweight reference body", the PsychoMorph generated average of photographs of the 10 bodies with the lowest Body Mass Index (BMI) available in the set of images (mean BMI = 17.3), and one "overweight reference body", the average of the 10 bodies with the highest BMI available in the set of images (mean BMI 25.2). Ten images were then selected from the same set of images, five with a BMI in the range 20-23kg/m^2, and five with a BMI in the range 16-19kg/m^2. AS then applied +20%, +40%, +60% and +80% of the PsychoMorph

underweight <-> overweight transform to make four larger versions of the individual, and -20%, -40%, -60% and -80% of the transform to make four slimmer versions. A total of 90 images were created, and an example sequence is shown in Figure 12.

Figure 12: An example sequence: The central image is the veridical image of the individual, BMI 16.9. The four to the left are morphed slimmer, the four to the right morphed larger

2.2.2 Online Rating

Participants were instructed "Look briefly at each of the following images, and decide whether they are over-weight, under-weight or neither over- nor under-weight". They were then shown the 90 images in a randomised order, and the accompanying text asked: "Is this woman over-weight, under-weight or neither over- nor under-weight?" Participants selected a weight category by clicking on one of the category words.

2.3 Outfit Choice Questionnaire

2.3.1 Scale Creation

A set of images of different outfits was created, with outfits varying in how close-fitting or revealing they were of different parts of the body, and in how much

they would provide “camouflage” for the wearer. Images of items of clothing were found online, and used to create a set of images of varying types and styles of outfits that might be worn in a variety of different settings, including swim/beach wear, sportswear, casual clothes, work outfits and party outfits. A“screen grab” tool was used to copy images of a range of items of clothing, and these images were pasted together to create a set of 60 different outfits, all modified to be of the same size. In order to enable respondents to focus on the shape of the clothes, rather than being distracted by other features, all outfits chosen were black. The images were informally piloted with a group of female colleagues in the Dept of Psychiatry, University of Oxford, and images viewed as being odd, or where the style of the outfit was generally disliked, were replaced.

2.3.2 Online Rating

Participants were given the following onscreen instruction: “Please look briefly at each of the following outfits and decide how comfortable you would feel wearing each of them in an appropriate public setting. Score from 1 (not at all comfortable) to 7 (very comfortable). You will be shown this range alongside each outfit. Please enter your score using the numbers on your keyboard.”

Participants were then shown each of the 60 outfits in a randomised order, and for each outfit were asked: “How comfortable with your body would you feel wearing this outfit in an appropriate public setting?” from 1 (not at all comfortable) to 7 (very comfortable).

As noted above, no exclusion criteria were set for the study, as it would not have been possible to verify whether people met the criteria. It was therefore

anticipated that a proportion of respondents would be male. Respondents who categorised themselves as male were given the following version of the task: "Please look briefly at each of the following outfits and decide how comfortable with her body a woman would have to feel to wear them in an appropriate setting. Score from 1 (She wouldn't have to be comfortable with her body to wear this) to 7 (She would have to be totally comfortable with her body to wear this). You will be shown this range alongside each outfit. Please enter your score using the numbers on your keyboard."

2.4 Other Measures

Two existing measures were used to collect data on participant satisfaction with own body. The Body Shape Satisfaction Scale (BSSS) (Pingitore et al., 1997) has an internal consistency of 0.92, and a two-week test-retest reliability of 0.68 to 0.77 in adolescents and young adults, and has been found to be predictive of poor health behaviours (Neumark-Sztainer et al., 2006).

The Eating Disorder Examination – Questionnaire EDE-Q (C. G. Fairburn & Beglin, 1994), is a questionnaire version of the Eating Disorder Examination, the gold standard measure of eating disorder pathology. The EDE-Q has four subscales: Weight Concern, Shape Concern, Eating Concern and Restraint. Since the hypothesis concerned participant satisfaction with their bodies, the Weight Concern and Shape Concern subscales of the EDE-Q were used. Test-retest reliabilities of 0.92 for Weight Concern and 0.94 for Shape Concern have been found in a sample of female undergraduates (Luce & Crowther, 1999).

Participants were also asked to give their age in years, gender (Male/Female), country of birth, ethnicity, number of hours since they last ate a meal, height, and current weight. They were also asked "Have you ever been diagnosed with an eating disorder? (Yes/No)", with the supplementary question: "If yes, it would be helpful if you could give brief details of your diagnosis".

2.5 Procedure

Participants followed a link within the email advertisements, or followed the URL given on the poster advertisements, which took them to the online study (implemented using the Xperiment online experimental platform, http://www.xperiment.mobi, by AW (see Statement of Contributions). Participants completed a consent process whereby they read an Information Sheet about the study, including experimenter contact details in case they wished to clarify any aspect of the study, and then completed an online consent form. They then answered a short series of questions concerning demographics. According to the gender they selected, they then completed either the female or male version of the *Outfit Choice Questionnaire*, followed by the BSSS and the two EDEQ subscales. Finally, they completed the ratings of the weight status for each of the images in the *Morphed Photographic Figure Scale.* Following this, participants were re-consented to check that they were still happy for their data to be used, and were then invited to make a note of their Unique Participant Code, which they were told they would need if they decided over the following two weeks that they wished to withdraw their data from the study. Finally, they

were debriefed as to the purpose of the study, and asked to enter a contact email address if they wished to be entered into the prize draw.

2.6 Statistical Analysis

2.6.1 Morphed Photographic Figure Scale

In order to create a set of images to use as stimuli in the next study, the percentages of female 18-30 year-olds, with a BMI of 18.5-25, and with no previous or current eating disorder, who coded each image as “underweight”, “neither under nor over weight” and “overweight” were calculated.

Percentages of images viewed as underweight, and percentage of images viewed as overweight, were calculated by gender, age, previous or current eating disorder, BMI, time since last meal, ethnicity, and three body satisfaction measures (EDEQ Weight Concern, EDEQ Shape Concern, Body Satisfaction Score), initially in a descriptive table and then using regression analysis to test for differences between groups. Regression analyses were repeated, adjusting for all variables except ethnicity. Analyses were conducted using Stata 15.

2.6.2 Outfit Choice Questionnaire

Using only the data from female participants, regression analyses was used to compare the results from the *Outfit Choice Questionnaire* with results from existing measures of body dissatisfaction, age, BMI and current or previous eating disorder diagnosis. Exploratory factor analysis was used to investigate the factor structure of the questionnaire.

2.7 Ethics

Ethics approval for the study was obtained from the Central University Research Ethics Committee at the University of Oxford (MSD-IDREC-C1-2015-020).

3. Results

3.1 Descriptive Data

The study ran from 7th April 2015 – 28th May 2015, during which time 337 participants completed the online task. Data on descriptive variables were complete for 337 people; however, two people had improbably high scores for BMI (>100 kg/m^2), and one was missing BMI, and they were therefore excluded from subsequent analyses (CONSORT diagram, Figure 13).

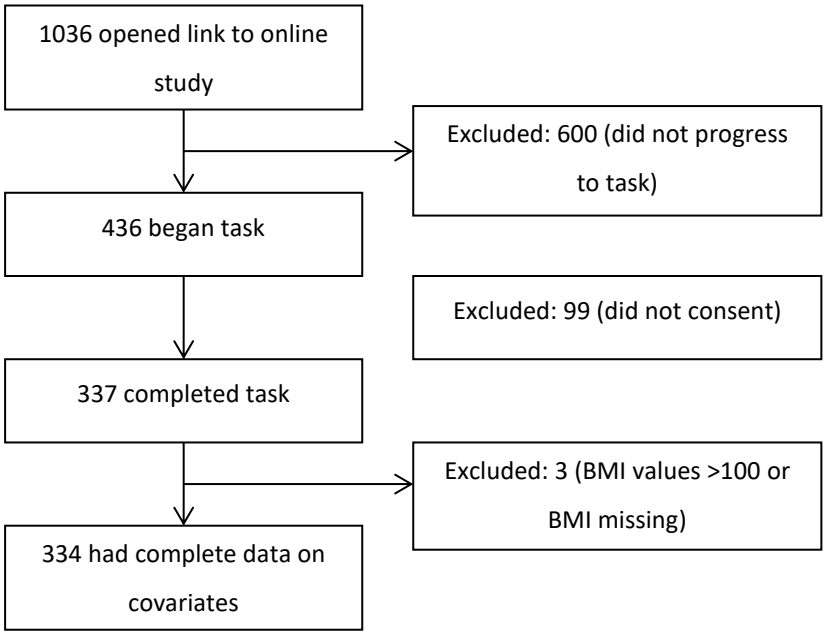


Figure 13: CONSORT diagram for Study 3

The mean age was 23.1 years old (SD 5.94), and 244 (73.1%) of participants were female. 280 (83.8%) of participants described themselves as White, and 230 (68.9%) were UK-born. The mean participant BMI was 22.4 kg/m^2 (Standard

Deviation 3.37), and 29 (8.68%) reported that they had a previous or current diagnosed eating disorder. Further details of demographics are given in Table 9.

Participant Characteristics		
Gender	Male (N (%))	90 (27.0)
	Female (N (%))	244 (73.1)
Ethnicity	Asian/Asian British (N (%))	23 (6.89)
	Black/African/Caribbean/Black British (N (%))	2 (0.60)
	Mixed/Multiple ethnic groups (N (%))	24 (7.19)
	Other (N (%))	5 (1.50)
	White (N (%))	280 (83.8)
Current/previous eating disorder	No (N (%))	305 (91.3)
	Yes (N (%))	29 (8.68)
Age (mean (standard deviation))		23.1 (5.94)
Mean hours since last ate (mean (standard deviation))		3.19 (3.34)
Mean BMI		22.4 (3.37)
EDEQ weight concern subscale	0-4 (N (%))	231 (69.2)
	>4 (N (%))	103 (30.8)
EDEQ shape concern subscale	0-4 (N (%))	259 (77.5)
	>4 (N (%))	75 (22.5)
Body Satisfaction Scale Score (mean (standard deviation))		3.29 (0.75)

Table 9: Descriptive data on questionnaire responders

3.2 Morphed Photographic Figure Scale

3.2.1 How images are perceived by the target group

The main purpose of designing these new images was to improve on the study methodology from Studies 1 and 2. The images used in Studies 1 and 2 were simply stretched and therefore looked either wider or narrower than the normal women, rather than more over- or underweight. The previous images were also not tested before use to see whether participants would be likely to perceive them as over, under or of normal weight.

It was important to investigate the perception of weight categories of the images by those in the target group that would be recruited in Study 4 (women aged 18-30, with a BMI of 18-25, with no previous or current eating disorder). Data from this smaller group meeting Study 4 criteria were therefore used (N=173; exact N varied by image, as not all participants gave a category for every image; missing data on each image ranged from 0 to 8.1%). For each image, the percentage of participants who viewed it as overweight, underweight and neither over- nor underweight was calculated. An example rating is shown in Figure 14.

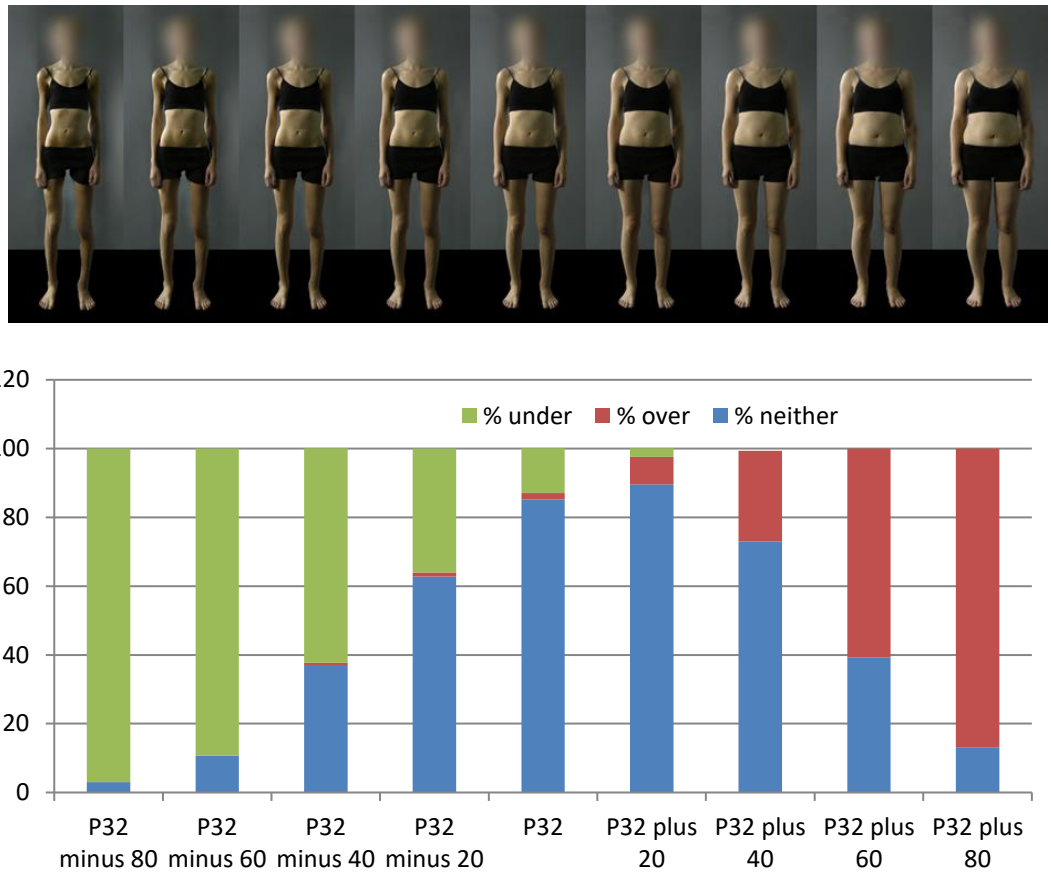


Figure 14: Top panel: Example sequence. The central image is the veridical image of the individual (BMI 16.9). The four to the left are morphed slimmer, the four to the right morphed larger. Bottom panel: Percentages of participants rating the images in the sequence as underweight, overweight or neither.

3.2.1 Differences between groups in perceived weight status

Between group differences in how the images were perceived were calculated to test the hypotheses. Means and standard deviation scores for the percentages of images coded as being under-, over- or neither under- nor overweight (Table 10) were calculated, both for the whole sample and by gender, age (converted into a binary variable of <21 years old/≥21 years as it was not normally distributed), ethnicity, BMI, history of a current or previous eating disorder, and three

measures of body satisfaction (EDE-Q Weight Concern and Shape Concern subscales, both used as binary scores with the cut off for clinical range of >4), and Body Satisfaction Score (grouped into “high”, “medium” and “low” tertiles) as it was not normally distributed.

			Percentage of images perceived as underweight		Percentage of images perceived as neither over nor underweight		Percentage of images perceived as overweight	
		N	Mean %	95% CI	Mean %	95% CI	Mean %	95% CI
Total		**334**	**18.5**	**17.6, 19.3**	**47.0**	**45.6, 48.4**	**30.6**	**29.3, 31.9**
By gender	Male	90	18.5	16.7, 20.3	48.0	45.2, 50.8	29.0	26.6, 31.5
	Female	244	18.4	17.4, 19.4	46.7	45.1, 48.3	31.1	29.6, 32.7
By ethnicity	Asian	23	20.6	17.1, 24.1	45.0	38.8, 51.2	27.6	21.1, 34.2
	Black	2	11.7	-9.5, 32.8	60.6	53.5, 67.6	27.2	20.2, 34.3
	Mixed	24	15.9	12.6, 19.2	49.1	43.8, 54.3	28.8	24.2, 33.4
	Other	5	13.3	6.6, 20.0	44.7	26.9, 62.5	37.6	16.7, 58.4
	White	280	18.6	17.7, 19.6	47.0	45.5, 48.5	30.9	29.5, 32.3
By current/previous eating disorder	No	305	18.5	17.6, 19.4	47.3	45.9, 48.8	30.3	28.9, 31.7
	Yes	29	18.1	15.0, 21.2	44.1	39.5, 48.8	33.3	28.7, 37.9
By age bands	<21	185	18.6	17.4, 19.8	46.7	44.9, 48.6	30.6	28.8, 32.4
	≥21	149	18.2	17.0, 19.5	47.4	45.3, 49.6	30.5	28.6, 32.4
By hours since last ate	0-1	118	17.7	16.2, 19.2	47.9	45.4, 50.5	30.7	28.5, 33.0
	1.5-3	114	19.2	17.8, 20.5	45.1	43.0, 47.2	31.1	29.1, 33.7
	3.5-20	102	18.5	16.8, 20.2	48.2	45.5, 50.8	29.5	27.1, 31.8
By BMI bands	Underweight (<18.5)	21	16.8	13.1, 20.6	49.6	42.9, 56.3	30.4	25.6, 35.2
	Normal (18.5-25)	262	18.2	17.3, 19.1	46.5	45.0, 48.0	31.3	29.8, 32.8
	Overweight (>25)	51	20.4	17.7, 23.1	48.7	44.4, 53.0	27.0	23.7, 30.3
By EDEQ weight subscale	0-4	231	18.4	17.3, 19.4	48.9	47.1, 50.6	28.8	27.3, 30.3
	>4	103	18.7	17.2, 20.2	42.9	40.9, 45.0	34.6	32.2, 36.9
By EDEQ shape subscale	0-4	259	18.2	17.2, 19.2	48.2	46.6, 49.8	30.0	28.5, 31.4
	>4	75	19.3	17.4, 21.2	43.1	40.8, 45.5	32.7	29.8, 35.5
By Body Satisfaction score	1-2.8	102	18.7	17.1, 20.3	45.1	42.6, 47.5	31.9	29.4, 34.5
	2.9-3.6	118	18.4	16.9, 19.8	46.9	44.6, 49.3	30.1	28.0, 32.2
	3.7-5	114	18.4	16.9, 19.8	48.9	46.5, 51.3	29.8	27.6, 32.0

Table 10: Percentage of Images Perceived as Over, Under or Neither Over- nor Underweight, by category

Full results of regression analyses comparing the different categories are given in Table 11. Those with a higher BMI were more likely to perceive women as being underweight, and less likely to perceive them as being overweight. This

difference was robust to adjustment for all the other variables listed in the table, and indicated that 0.3% more images were viewed as underweight with each 1 point increase in participant BMI (95% CI 0.02, 0.57, p=0.03), and 0.5% fewer images were perceived as overweight with each 1 point increase in participant BMI (95% CI -0.93, -0.11, p=0.01).

Higher EDEQ Weight Concern scores were strongly associated with rating a higher percentage of women as overweight, which again was robust to adjustment by all the other variables in the table: 7.33% more images are viewed as overweight by those with an EDEQ Weight Concern score of >4 than those with a score of 0-4 (95% CI 3.59, 11.06), p<0.001 (4 is the cut-off for clinical significance on each EDEQ subscale) (C. G. Fairburn & Beglin, 1994).

There was no evidence for any differences in the percentage of images coded as under or overweight by gender, age, presence or absence of current/previous eating disorder, time since last meal, EDEQ Shape Concern or Body Satisfaction score.

	Differences in percentage of images perceived as underweight		Differences in percentage of images perceived as overweight	
Difference in perceived weight scores	Unadjusted difference (95% confidence intervals, p value) (N=334)	Adjusted* difference (95% confidence intervals) (N=334)	Unadjusted difference (95% confidence intervals) (N=219)	Adjusted* difference (95% confidence intervals) (N=218)
Male	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
Female	-0.07 (-2.01, 1.86), p=0.941	0.22 (-1.88, 2.33), p=0.835	2.10 (-0.84, 5.03), p=0.161	-0.45 (-3.56, 2.66), p=0.776
No current/ previous ED	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
Current/ previous ED	-0.36 (-3.41, 2.69), p=0.816	-0.35 (-3.60, 2.88), p=0.830	2.99 (-1.64, 7.62), p=0.206	0.90 (-3.89, 5.69), p=0.711
Age <21	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
Age ≥21	-0.42 (-2.15, 1.30), p=0.629	-0.89 (-2.68, 0.91), p=0.331	-0.10 (-2.73, 2.53), p=0.942	0.76 (-1.89, 3.41), p=0.574
Hours since last ate: 0-1	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
Hours since last ate: 1.5-3	1.44 (-0.62, 3.50), p=0.170	1.30 (-0.80, 3.40), p=0.224	0.65 (-2.48, 3.79), p=0.679	1.00 (-2.10, 4.11), p=0.525
Hours since last ate: 3.5-20	0.79 (-1.33, 2.91), p=0.466	0.64 (-1.52, 2.79), p=0.563	-1.28 (-4.51, 1.95), p=0.436	-1.52 (-4.71, 1.67), p=0.349
BMI	**0.29 (0.04, 0.55), p=0.023**	**0.30 (0.02, 0.57), p=0.033**	**-0.43 (-0.82, -0.05), p=0.028**	**-0.52 (-0.93, -0.11), p=0.012**
EDEQ weight 0-4	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
EDEQ weight >4	0.32 (-1.54, 2.18), p=0.737	-0.34 (-2.87, 2.19), p=0.792	**5.80 (3.03, 8.55), p<0.001**	**7.33 (3.59, 11.06), p<0.001**
EDEQ shape 0-4	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
EDEQ shape >4	1.14 (-0.92, 3.19), p=0.278	1.19 (-1.57, 3.96), p=0.397	2.74 (-0.37, 5.86), p=0.084	-1.03 (-5.11, 3.06), p=0.621
BD score 1-2.8	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
BD score 2.9-3.6	-0.29 (-2.42, 1.84), p=0.789	0.26 (-2.18, 2.71), p=0.833	-1.84 (-5.06, 1.39), p=0.264	-0.26 (-3.87, 3.35), p=0.889
BD score 3.7-5	-0.31 (-2.45, 1.84), p=0.778	0.63 (-2.08, 3.35), p=0.647	-2.11 (-5.37, 1.13), p=0.202	0.75 (-3.26, 4.76), p=0.713

*Table 11: Between Group Differences in Perceived Weight Scores (positive scores indicate more bodies perceived as overweight) (*adjusted for all other values in the table)*

3.3 Outfit Choice Questionnaire

Amongst the 244 female respondents, the *Outfit Choice Questionnaire* had a mean score of 5.12 (SD 1.11) (range 1 (I would feel very uncomfortable with my body wearing this) to 7 (I would feel very comfortable with my body wearing this)).

3.3.1 Relationship to Existing Measures of Body Dissatisfaction

The *Outfit Choice Questionnaire* correlated well with other measures of body satisfaction, and with previous or current eating disorder, as shown in the regression analyses in Table 12. Higher BMI scores were also associated with lower mean scores on the *Outfit Choice Questionnaire*. After adjustment for all the other measures given in Table 12, the strongest associations were with presence versus absence of a current or previous eating disorder, and the Body Satisfaction Scale.

Existing Measure	Regression (95% CI, p)	Regression (95% CI, p), adjusted for age & BMI	Regression (95% CI, p), adjusted for age, BMI, EDEQ Shape and Weight Concern, BSSS and eating disorder status
EDEQ: Shape concern (used as a continuous score)	-0.67 (-0.82, -0.52), p<0.001	-0.35 (-0.43, -0.27), p<0.001	-0.13 (-0.26, 0.00), p=0.054
EDEQ: Weight concern (used as a continuous score)	**-0.64 (-0.83, -0.46), p<0.001**	**-0.24 (-0.32, -0.17), p<0.001**	0.05 (-0.05, 0.16), p=0.334
BSSS	**0.40 (0.33, 0.46), p<0.001**	**0.89 (0.73, 1.05), p<0.001**	**0.70 (0.48, 0.93), p<0.001**
BMI	**-0.79 (-1.18, -0.40), p<0.001**	**-0.73 (-1.11, -0.36), p<0.001**	-0.02 (-0.06, 0.01), p=0.190
Current or previous eating disorder (Y/N)	**-0.09 (-0.12, -0.05), p<0.001**	**-0.10 (-0.13, -0.06), p<0.001**	**-0.48 (-0.85, -0.11), p=0.010**
Age	-0.01 (-0.04, 0.01), p=0.289	0.001 (-0.02, 0.02), p=0.932	-0.02 (-0.04, 0.003), p=0.102

Table 12: Regression analyses of the Outfit Choice Questionnaire against existing measures of body satisfaction, current/previous eating disorder status, BMI and age

3.3.2 Factor Analysis

Factor analysis was conducted, to establish the structure of the *Outfit Choice Questionnaire*. The scree plot of the Eigenvalues from the unforced principal factors analysis and the numerical results from the rotated unforced principal factors analysis are presented in Appendix 1.

On the basis of these results, the factor analysis was then forced to 6 factors (as this was the point where the amount of variance explained dropped), and 8

factors (as after this point the variance dropped below 1). Studying the results in terms of the outfits grouped in each factor using these two analyses, the 8 factor analysis did not group the data in a more coherent way than the 6 factor analysis. The 6 factor analysis only had one outfit in the 6th factor, and the analysis was therefore rerun, forcing to 5 factors. The distribution of outfits from this are shown in Figure 15.

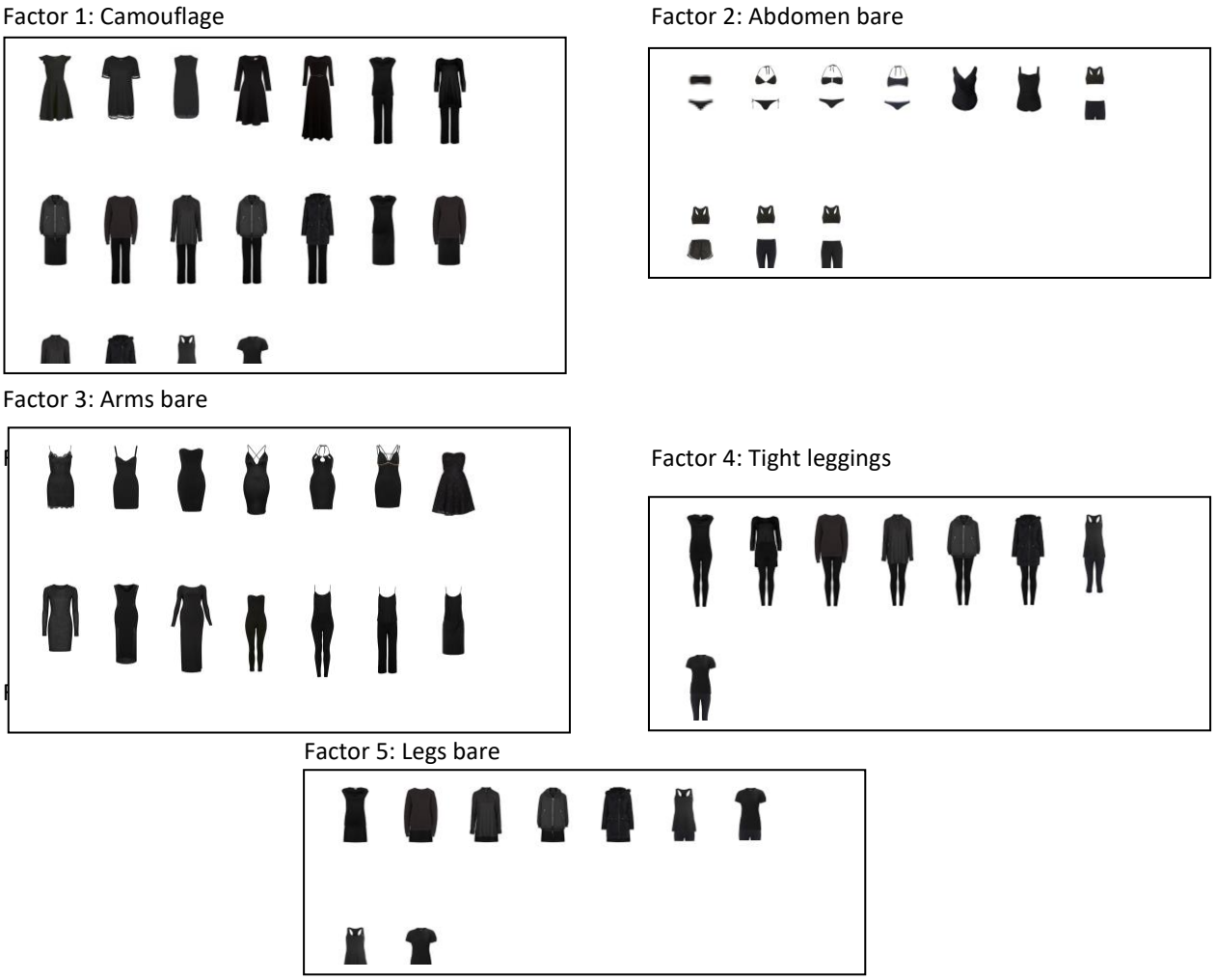


Figure 15: Factor Distribution of Outfit Choice Questionnaire (forced to 5 factors)

4. Discussion

4.1 Morphed Photographic Figure Scale

Each image in the *Morphed Photographic Figure Scale* was validated in terms of categorising its perceived weight category in a large online study. Ratings data for each image shows a clear rank ordering from underweight to overweight, indicating good construct validity.

Results from the online study were used to test whether perception of weight categories varied by various individual factors. As participant BMI increased, they judged more images of others as underweight, and fewer images as overweight. This is consistent with the existing literature which has found that subjects with a lower BMI are more likely to overestimate the BMI of others' bodies (Tovee et al., 2000).

In line with previous findings, presence of a diagnosed eating disorder was not in itself predictive of viewing more images as overweight (Tovee et al., 2000). However, participants with EDE-Q Weight Concern scores in the clinical range were more likely to judge images as being overweight. This contrasts with a previous study (Tovee et al., 2000), which found no relationship between accuracy of BMI ratings and any EDE-Q subscales. This difference may be due to the fact that this study asked for participant opinion on weight category, rather than an estimate of actual BMI. It may be that participants in the previous study would have classified lower BMI bodies as overweight had they been asked to categorise them, rather than purely to estimate their BMI.

There was no evidence for gender differences in perceived weight categories. This is in line with some previous work that has found no evidence of gender differences in what constitutes an ideal, attractive or healthy weight for women (Crossley et al., 2012). However, this contrasts with results from an earlier study using numerical descriptions of weight, which did find gender differences in what would be considered underweight, with men giving a lower estimates of what they would consider under- and overweight in a woman (Oakes, Slotterback, & Mecca, 2003). It may be that the use of numerical descriptions of weight rather than images in this previous study picked up gender differences in knowledge of what different weights would look like in a person, rather than actual difference in opinion regarding weight status.

Despite finding few differences in weight perception between groups, a data set was created, indicating how each image was perceived by women in the target group for the planned Study 4: those aged 18-30, with a BMI of 18.5 to 25, and no previous or current eating disorder. An example of this data for a set of images based on one veridical woman is presented in Appendix 2.

4.2 Outfit Choice Questionnaire

The *Outfit Choice Questionnaire* was validated against existing measures of body dissatisfaction and the measure correlated well. In terms of its factor structure, it divided up into five distinct factors: Factor 1 “Camouflage”: outfits which cover the majority of the body; Factor 2 “Abdomen bare”: outfits in which the abdomen is exposed; Factor 3 “Arms bare”: outfits in which arms are exposed; Factor 4

"Tight leggings": outfits including tight leggings; and Factor 5: "Legs bare" outfits, with short skirts or shorts exposing part of the thighs of the wearer.

4.3 Future Directions

Having validated a new set of ecologically valid images of women's bodies of different weights in the *Morphed Photographic Figure Scale*, and gathered data on how each was viewed in terms of weight category, a selection of these new images was ready for use in Study 4. The planned intervention in Study 4 was very similar to that in Studies 1 and 2, but, in order to make it more analogous to a possible clinical intervention, was given for 5 minutes twice a day for a week. A range of measures were collected at baseline and follow up in order to develop a fuller understanding of perception of and satisfaction with own body size.

Chapter 4 will describe the process of designing a series of measures to assess implicit beliefs about body size and satisfaction with body size, and how these and the *Outfit Choice Questionnaire* compared to existing measures in the baseline results collected in Study 4.

CHAPTER 4: NEW WAYS OF MEASURING PERCEPTION OF AND SATISFACTION WITH OWN BODY SIZE IN WOMEN

1. Introduction

This chapter will describe the process of designing a series of measures to assess beliefs about body size and satisfaction with body size, and how these new measures compared to existing measures in the baseline results collected in Study 4.

As discussed at greater length in Chapter 1, dissatisfaction with body size can be thought of as having two components: perception of body size, and cognitive beliefs about body size (Cash & Deagle, 1997). In order to fully understand and develop new interventions to treat body dissatisfaction, it is important to be able to accurately measure both components and understand how they relate to each other.

Existing measures of perception of own body size are either unwieldy to perform in large numbers of participants (Farrell et al., 2005; Gardner & Brown, 2014) use line drawings which are not very accurate representations of bodies (e.g (Stunkard et al., 1983), or use 3D modelling software not constrained by biological plausibility (Crossley et al., 2012).

Existing measures of satisfaction with body size rely on self-report questionnaires. Whilst useful, they leave open the possibility that people either lack insight into their own beliefs (Konstantakopoulos et al., 2011; Nisbett & Wilson, 1977), or conceal their true views in order to respond in a way they view

as being socially acceptable. Tasks using implicit measures are proving to be useful early predictors of treatment response, for example in the case of implicit measures of mood enabling very early prediction of response to treatment in people suffering from depression (Warren, Pringle, & Harmer, 2015). Implicit tasks have potential use in the area of body dissatisfaction and eating disorders, both as a way of investigating concealed beliefs, and as a potential starting point for developing new treatments.

Therefore, a new method of measuring perception of own size, and a range of measures designed to capture implicit views about body size and satisfaction with body size were developed. These new tasks, along with the *Outfit Choice Questionnaire* described in Chapter 3, were used to gather a comprehensive set of measures at baseline (Day 1) and follow-up (Day 8) in Study 4. In this Chapter, the development of these new measures is described, along with how they compare with existing questionnaire measures of body dissatisfaction and eating disorder pathology taken at baseline in Study 4. The hypothesis was that performance on the novel tasks would be associated with subjective ratings on the existing questionnaire measures.

2. Methods

2.1 Participants

Participants were recruited by e-mail advertisement (to undergraduate and graduate students at the University of Oxford via their College Junior and Middle/Graduate Common Rooms and University departments), word of mouth, posters in Departments and Colleges, and advertisements on Facebook.

Potentially interested participants were asked to complete a brief online screening questionnaire to ensure that they met inclusion criteria.

Inclusion criteria required participants to be female, aged 18 to 30 years old, speak English as a first language and have a Body Mass Index (BMI) of 18-25 kg/m^2. Participants were excluded if they had a current or previous eating disorder, were undergoing treatment for any other mental illness, or taking psychoactive medication that may affect cognitive function (e.g. street/illicit drugs, methylphenidate, St John's Wort), had any current serious neurological illness (e.g. MS, epilepsy), had Coeliac disease, were currently pregnant, smoked more than 10 cigarettes per day or had dyslexia.

Results presented here concern the measures taken during the baseline (Day 1) session of Study 4. Participants were asked to eat a light meal 2 hours before attending each session, and to avoid caffeine, alcohol and cigarettes in the 2 hours before the session. Participants were tested by one of two experimenters, HB and KN (see Statement of Contributions).

2.2 Existing Measures

Participants completed a series of questionnaire measures at baseline (Day 1), which are described here.

2.2.1 Body Shape Satisfaction Scale

The Body Shape Satisfaction Scale (BSSS) (Pingitore et al., 1997) has an internal consistency of 0.92, and a two-week test-retest reliability of 0.68 to 0.77 in adolescents and young adults, and has been found to be predictive of poor health

behaviours including unhealthy weight control behaviours, binge eating, lower levels of physical activity, and less fruit and vegetable intake (Neumark-Sztainer et al., 2006).

2.2.2 EDE-Q

The Eating Disorder Examination – Questionnaire EDE-Q (C. G. Fairburn & Beglin, 1994) is a questionnaire version of the Eating Disorder Examination (C. G. W. Fairburn, G.T. , 1993), the gold standard measure of eating disorder pathology. It has four subscales: Weight Concern, Shape Concern, Eating Concern and Restraint. Each subscale is scored from 0-6 and the cut off for a clinical score is ≥4. Test-retest reliabilities of 0.92 for Weight Concern, 0.94 for Shape Concern, 0.87 for Eating Concern and 0.81 for Restraint have been found in a sample of female undergraduates (Luce & Crowther, 1999).

2.2.3 Visual Analogue Scales

As in Studies 1 and 2 (Chapter 2), two simple 10-point Visual Analogue Scales were used, as a very simple measure of perception of and satisfaction with own size. The results from Studies 1 and 2 demonstrate that these scales are sensitive to change. The first scale asks participants to rate the size of their body at the moment, from too thin (0) to too fat (10). The second scale asks participants to rate how satisfied they feel about the size of their body at the moment, from very dissatisfied (0) to very satisfied (10). It was anticipated that these simple measures would correlate with more complex questionnaire measures of body satisfaction.

2.3 Avatar Measures

Software from the company BodyLabs (www.bodylabs.com) was used. This company has used machine learning algorithms and a large database of body shapes to create avatars that can be manipulated to different sizes, but ensures that the body shapes created by the manipulations are biologically plausible. Working with a developer (JD – see Statement of Contributions), a version of the female avatar was created, which enabled participants to alter the weight, bust, waist and hip measurements using sliding scales, without any numbers to guide them.

Using this new avatar, I was able to enter participant height and weight, and participants were then presented with two sequential starting images, both of which were the same height as the participant: the first of these was their body weight plus 10%, and the second was their own body weight minus 10%. They were told that the avatar was their height but not their weight. They were shown the sliders and that the avatar could be rotated to view her from different angles, and they were asked to take up to 2 minutes to "make this avatar your own size and shape", before clicking "next" to have a second attempt. Once finished, they were asked to look away whilst the experimenter noted the size and shape they had made the women. The hypothesis was that there would be a strong correlation between measured BMI and the size of this "own BMI" avatar (Figure 166).

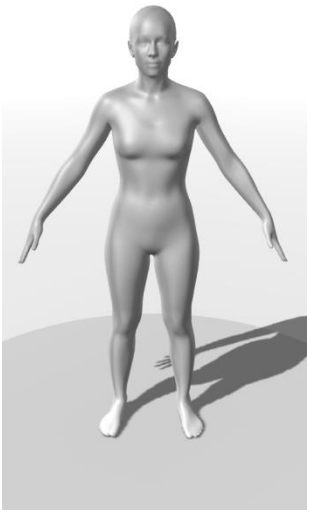


Figure 16: Example Avatar

Participants also completed the task with the instruction to make the avatars "the size and shape you would most like to be". The hypothesis was that there would be an association between explicit measures of body dissatisfaction and smaller ideal BMI, and also a correlation between higher levels of body dissatisfaction and greater discrepancy between the avatar measures of "own BMI" and "ideal BMI".

2.4 Behavioural Measures

2.4.1 Biscuit Consumption

Half way through the session, participants were told they had completed all the reaction time tasks (these tasks are described below in Section 2.5 "Novel Implicit Measures") and could therefore drink caffeine again. They were given a short break and offered a cup of tea or coffee, or a glass of water. The researcher brought them their drink and a bowl containing 100 g of small, sweet, oat biscuits. The researcher told them "I'll leave you to have a 5-minute break whilst I

set up the laptop to take home with you". The researcher returned after 5 minutes. The bowl of biscuits was weighed after the participant had left, to determine the weight of biscuits the participant had eaten.

2.4.2 Outfit Choice Questionnaire

The development of this task is described in detail in Chapter 3. Participants were shown the 60 outfits in a randomised order and for each outfit were asked "How comfortable with your body would you feel wearing this outfit in an appropriate public setting?" rating this between 1 (not at all comfortable) and 7 (very comfortable). A mean score was generated, indicating overall level of comfort in different outfits.

2.5 Novel Implicit Measures

Participants also completed a series of novel measures designed to capture implicit beliefs about body size and satisfaction with body size. These are described in more detail here.

2.5.1 Lexical Decision Task

Participants were primed to their own body by completing the Avatar "own BMI" task, in which they were asked to make the avatar their own size and shape. They then completed the Lexical Decision Task described by Smith (Smith, Joiner, & Dodd, 2014). The task is to respond as quickly as possible regarding whether the groups of letters presented are real words or not. The task contains real words in the categories of "beauty", "ugly", "neutral" and "positive". Reaction times were calculated for "beauty" minus "ugly" words, and for "neutral" minus "positive" words. It was hypothesised that faster reaction times for beauty than ugly words,

and for positive than neutral words would be associated with lower body dissatisfaction. The LDT task was presented in ePrime, which was programmed with assistance from AR (see Statement of Contributions).

2.5.2 Lexical Decision Task – Recall

Following two further tasks and a five-minute break, participants were given one minute to write down as many of the real words from the LDT as they were able to recall. Number of "beauty" minus number of "ugly" words, and number of "neutral" minus number of "positive" words recalled were calculated. It was hypothesised that recalling a greater number of "beauty" minus "ugly" words would be associated with lower body dissatisfaction, and recalling a greater number of "positive" versus "neutral" words would be associated with lower body dissatisfaction.

2.5.3 Implicit Association Test – Words

The Implicit Association Test (IAT) (Greenwald et al., 1998) is one of the most frequently used tasks to assess implicit cognitions. It seeks to "measure implicit attitudes by measuring their underlying automatic evaluation" (Greenwald et al., 1998). Participants are asked to sort "target" words and "attribute" words using different keys on the computer keyboard. The theory is that reaction times will be quicker if the words which are sorted using the same key are "evaluatively compatible" or "congruent" than if they are non-compatible (or incongruent).

For the task to work, it is necessary that the words be as similar as possible to each other in terms of length, number of syllables, and familiarity. For this study, it was also important that the words were viewed as being positive or negative

with regard to weight and shape. There is no database indicating the age at which people generally learn words associated with weight and shape, so a brief online questionnaire was designed in order to identify words: this is described in Appendix 3.

Following the results from the study described in Appendix 3, the "Attribute" words chosen relating to Slim were: trim, slim, fit, slender; the "Attribute" words related to Not Slim were: stout, podgy, obese, fat. Participants were asked to sort words belonging to the "target" categories of Self (me, mine, I, myself) and Other (she, they, their, her).

After a brief practice trial using words relating to trees, flowers, cities and countries, participants completed a practice trial for Self/Slim pairings, and then a longer trial of this pairing, followed by a practice and then longer trial of "Self/Not Slim" pairings. The hypothesis was that participants with more positive views of their own bodies and who viewed themselves as smaller would be faster to respond to "Self-Slim" pairings, and those with more negative views of their bodies and who viewed themselves as larger would be faster to respond to "Self-Not slim" pairings. The task was conducted using EPrime.

2.5.4 Implicit Association Task – Images

For this task, the same methodology as the IAT word task described above was used, but words relating to body size were replaced with images of the same women modified to appear of normal weight or slightly overweight (selected from the images in the *Morphed Photographic Figure Scale*, whose development is described in Chapter 3). The hypothesis was that participants with more positive

views of their own bodies and who viewed themselves as smaller would be faster to respond to "Self-normal sized image" pairings, and those with more negative views would be faster to respond to "Self-overweight image" pairings. The task was conducted using EPrime.

2.6 Statistical Analysis

Means, standard deviations and median and interquartile ranges for both existing and novel measures at baseline were calculated, as well as descriptive data on age and educational background. Linear regression analysis was used to assess the correlations between new and existing measures of body size and satisfaction with own size. Linear regression rather than correlation was used in order to generate 95% confidence intervals around the relationships as well as point estimates and p values.

Cronbach's alpha was used to investigate the test-retest reliability of the avatar measures of "own BMI" and "ideal BMI".

Results from the Implicit Association Test were analysed using the improved scoring algorithm described by Greenwald (Greenwald, Nosek, & Banaji, 2003), including following their guidance as to which responses should be discarded as outliers. This generates "Greenwald's D", analogous to an effect size calculation for differences between "congruent" and "incongruent" word pairings.

All analyses were completed using Stata 15.

2.7 Ethics

The main study described here was approved by the Oxford University Central University Research Ethics Committee (MS-IDREC-C1-2015-083).

3. Results

3.1 Descriptive data

Ninety three participants were recruited between 27th Aug 2015 and 22nd April 2016. A total of 499 people were assessed for eligibility; 406 were excluded as they either did not meet the criteria, or did not respond to invitations to participate. Descriptive data on the participants and the percentage who completed each task are presented in Table 13; a CONSORT diagram for the study is in Chapter 5.

Measures		N (%) with data for this measure
Age (mean, SD)	22.7 (3.15)	93 (100)
Education level (N, % degree or above)	65 (69.9%)	93 (100)
N (%) right handed	77 (82.8%)	93 (100)
Baseline EDE-Q Weight (mean, SD)	1.03 (1.11)	93 (100)
Baseline EDE-Q Shape (mean, SD)	1.45 (1.24)	93 (100)
Baseline EDE-Q Restraint (mean, SD)	0.83 (1.02)	93 (100)
Baseline EDE-Q Eating Concern (mean, SD)	0.34 (0.53)	93 (100)
Total EDE-Q score	0.91 (0.88)	93 (100)
BSSS (mean, SD) (range 1-5)	3.42 (0.70)	93 (100)
Measured BMI (mean, SD)	21.9 (2.11)	92 (98.9)
Baseline rating of own size from 0 (too thin) to 10 (too fat) (mean, SD)	5.72 (0.94)	93 (100)
Baseline satisfaction with own size from 0 (very dissatisfied) to 10 (very satisfied) (mean, SD)	6.14 (2.02)	93 (100)
Own body Avatar BMI (1) (mean, SD)	22.7 (4.43)	92 (98.9)
Own body Avatar BMI (2) (mean, SD)	21.5 (3.74)	92 (98.9)
Own body Avatar BMI (mean, SD)	22.1 (3.84)	92 (98.9)
Difference between actual & mean avatar BMI (avatar BMI – measured BMI) (mean, SD)	0.25 (2.76)	91 (97.9)
Ideal body Avatar BMI (1) (mean, SD)	19.6 (3.00)	92 (98.9)
Ideal body Avatar BMI (2) (mean, SD)	18.6 (2.78)	92 (98.9)
Ideal body Avatar BMI (mean, SD)	19.1 (2.73)	92 (98.9)
Difference between perceived and ideal avatar BMI (ideal – perceived) (mean, SD)	-2.99 (3.77)	92 (98.9)
Greenwald's D for incongruent – congruent words (mean, SD)	0.34 (0.43)	93 (100)
Greenwald's D for incongruent – congruent images (mean, SD)	0.05 (0.03)	93 (100)
Beauty words RT (mean, SD)	590 (60.1)	91 (97.9)
Ugly words RT (mean, SD)	617 (69.1)	91 (97.9)
Neutral words RT (mean, SD)	632 (74.3)	91 (97.9)
LDT words recalled (N, SD)	5.16 (2.15)	91 (97.9)

Table 13: Descriptive data

3.2 VAS measures of size and satisfaction with size

VAS measures of size and satisfaction correlated strongly with longer questionnaire measures of satisfaction with size, and also with broader measures of eating disorder pathology (Table 14).

3.3 Avatar Measures

Despite participants being presented with images of two different starting sizes (one their body weight plus 10%, one their body weight minus 10%), test-retest reliability for "own BMI" measured using Cronbach's alpha was 0.88. Test-retest reliability for "ideal BMI" was 0.90.

Avatar "own BMI" correlated very strongly with measured BMI, although participants tended to slightly overestimate their own BMI (by 0.25 kg/m^2 (SD 2.8)). Avatar "own BMI" also correlated with the VAS measure of size, and a range of measures of satisfaction with size (Table 14). Both "ideal BMI", and to a greater extent the difference between the "own BMI" and "ideal BMI" avatars correlated strongly with measures of body dissatisfaction and other measures of disordered eating (Table 14).

3.4 Outfit Choice Questionnaire

Results from the novel *Outfit Choice Questionnaire* correlated strongly with existing measures of body satisfaction, and with broader EDE-Q measures of eating disorder pathology.

3.5 Biscuit Consumption

There was no evidence of an association between body size or satisfaction with body size and quantity of biscuits consumed, except some evidence for an association between greater body satisfaction (BSSS) and eating larger quantities of biscuits (Table 14).

	Measured BMI	VAS Size	VAS Satisfaction	BSSS	EDE-Q Total	EDE-Q Shape	EDE-Q Weight	EDE-Q Eating	EDE-Q Restraint
VAS Size	**0.22 (0.15, 0.30), p<0.001**	-	-	**-0.71 (-0.95, -0.48), p<0.001**	**0.63 (0.45, 0.81), p<0.001**	**0.43 (0.30, 0.56), p<0.001**	**0.52 (0.38, 0.66), p<0.001**	**0.68 (0.33, 1.02), p<0.001**	**0.43 (0.26, 0.60), p<0.001**
VAS Satisfaction	**-0.26 (-0.45, -0.07), p=0.009**	**-1.08 (-1.46, -0.69), p<0.001**	**-**	**1.96 (1.53, 2.40), p<0.001**	**-1.28 (-1.68, -0.88), p<0.001**	**-0.91 (-1.19, -0.63), p<0.001**	**-0.96 (-1.29, -0.64), p<0.001**	**-1.46 (-2.20, -0.72), p<0.001**	**-0.92 (-1.29, -0.56), p<0.001**
Avatar: "own BMI"	**1.21 (0.94, 1.48), p<0.001**	**2.32 (1.61, 3.05), p<0.001**	**-0.62 (-0.99, -0.24), p=0.001**	**-1.68 (-2.76, -0.59), p=0.003**	**1.17 (0.28, 2.05), p=0.010**	**0.87 (0.25, 1.49), p=0.007**	**1.04 (0.35, 1.73), p=0.004**	1.32 (-0.18, 2.82), p=0.084	0.59 (-0.19, 1.37), p=0.135
Avatar "ideal BMI"	**0.57 (0.33, 0.82), p<0.001**	0.05 (-0.57, 0.67), p=0.866	0.27 (-0.002, 0.55), p=0.052	**1.38 (0.63, 2.14), p<0.001**	**-0.88 (-1.50, -0.26), p=0.006**	**-0.60 (-1.04, -0.16), p=0.008**	**-0.63 (-1.13, -0.13), p=0.014**	-0.89 (-1.96, 0.17), p=0.099	**-0.74 (-1.28, -0.20), p=0.008**
Difference between ideal & own BMI	**-0.64 (-0.98, -0.29), p<0.001**	**-2.27 (-2.98, -1.57), p<0.001**	**0.9 (0.55, 1.23), p<0.001**	**3.06 (2.14, 3.98), p<0.001**	**-2.05 (-2.83, -1.26), p<0.001**	**-1.47 (-2.02, -0.92), p<0.001**	**-1.67 (-2.28, -1.05), p<0.001**	**-2.21 (-3.64, -0.79), p=0.003**	**-1.33 (-2.05, -0.61), p<0.001**
Outfits task	**-0.10 (-0.19, -0.02), p=0.017**	**-0.32 (-0.51, -0.14), p=0.001**	**0.17 (0.09, 0.26), p<0.001**	**0.71 (0.49, 0.92), p<0.001**	**-0.48 (-0.66, -0.29), p<0.001**	**-0.37 (-0.50, -0.25), p<0.001**	**-0.39 (-0.53, -0.24), p<0.001**	**-0.27 (-0.44, -0.09), p=0.003**	**-0.50 (-0.84, -0.17), p=0.003**
Biscuits	0.54 (-0.92, 1.99) p=0.465	-2.32 (-5.53, 0.88) p=0.152	1.41 (-0.06 2.87), p=0.060	**4.45 (0.26, 8.64), p=0.038**	-2.77 (-6.16, 0.62), p=0.108	-1.52 (-3.94, 0.89), p=0.213	-2.32 (-5.00, 0.35), p=0.088	-2.22 (-7.95, 3.51), p=0.443	-2.61 (-5.53, 0.30), p=0.078

Table 14: Comparing new with existing Explicit Measures of Body Size & Satisfaction (Beta coefficients with 95% Confidence Intervals & p values)

3.6 Lexical Decision Task

The LDT was not associated with any existing or novel measures of body satisfaction or size, except for an association between faster responses to ugly words and the Word IAT (Table 16).

3.7 Lexical Decision Task – Recall

Recalling a greater number of "beauty" than "ugly" category words was associated with existing measures of body dissatisfaction (BSSS and EDE-Q) (Table 15). Greater recall of beauty than ugly words was also associated with smaller VAS size, higher VAS satisfaction, and with a bigger difference between "own BMI" and "ideal BMI" on the Avatar measure, number of biscuits eaten, and the Word IAT (Table 16)

3.8 Implicit Association Task – Words

There was evidence for a strong association between the Word IAT and both the BSSS and the EDE-Q (Table 15); responding faster to "self" words paired with "slim" words than to "self" words paired with "not slim" words (that is, an implicit belief that one is slim) was associated with higher scores (indicating greater satisfaction) on the BSSS, and lower scores (indicating less eating disorder pathology) on the EDE-Q.

Faster responses to "self" words paired with "slim" words than to "self" words paired with "not slim" words was also associated with lower VAS size, higher VAS satisfaction, and with a smaller difference between "own BMI" and "ideal BMI" avatars Table 16

3.9 Implicit Association Task – Images

The Image IAT was associated with measured BMI, such that participants with higher BMI responded more quickly to "self" words paired with slightly overweight images than to "self" words paired with normal weight images. However, the Image IAT was not associated with any existing questionnaire measures of body satisfaction (Table 15) or with any novel measures of body satisfaction, with the exception of the Word IAT.

	Measured BMI	BSSS	EDE-Q Total	EDE-Q Shape	EDE-Q Weight	EDE-Q Eating	EDE-Q Restraint
Word IAT (Greenwald D)	-0.02 (-0.06, 0.02), p=0.482	**0.21 (0.09, 0.33), p=0.001**	**-0.16 (-0.26, -0.06), p=0.001**	**-0.11 (-0.18, -0.04), p=0.003**	**-0.11 (-0.18, -0.03), p=0.007**	-0.15 (-0.33, 0.01), p=0.064	**-0.15 (-0.23, -0.06), p=0.001**
Images IAT (Greenwald D)	**-0.03 (-0.06, -0.01), p=0.019**	0.15 (-0.02, 0.14), p=0.145	-0.02 (-0.09, 0.04), p=0.455	-0.03 (-0.08, 0.01), p=0.174	-0.02 (-0.08, 0.03), p=0.351	-0.06 (-0.17, 0.05), p=0.298	0.01 (-0.04, 0.07), p=0.530
LDT: beauty – ugly RT	-2.11 (-5.90, 1.67), p=0.270	6.58 (-4.73, 17.9), p=0.251	1.43 (-8.01, 10.9), p=0.764	2.88 (-3.78, 9.54), p=0.392	0.45 (-7.02, 7.91), p=0.906	2.98 (-12.74, 18.71), p=0.707	-1.35 (-9.48, 6.78), p=0.742
LDT: positive – neutral RT	-0.13 (-3.93, 3.68), p=0.948	1.11 (-10.7, 12.9), p=0.852	-0.08 (-0.34, 0.17), p=-0.507	-3.15 (-9.59, 3.27), p=0.332	-3.83 (-11.03, 3.36), p=0.293	-1.21 (-16.5, 14.0), p=0.875	2.89 (-10.8, 4.98), p=0.467
LDT recall beauty – ugly	-0.05 (-0.20, 0.10), p=0.527	**0.74 (0.33, 1.16), p=0.001**	**-0.46 (-0.80, -0.12), p=0.009**	**-0.34 (-0.58,-0.10), p=0.006**	**-0.35 (-0.62,-0.08), p=0.011**	**-0.78 (-1.34,-0.21), p=0.008**	-0.24 (-0.54, 0.06), p=0.118
LDT recall positive – neutral	-0.09 (-0.19, 0.01), p=0.089	0.16 (-0.15, 0.47), p=0.32	-0.08 (-0.34, 0.17), p=0.51	-0.09 (-0.27, 0.09), p=0.306	-0.05 (-0.25, 0.15), p=0.632	-0.04 (-0.46, 0.39), p=0.87	**-0.05 (-0.27, 0.17), p=0.662**

Table 15: Comparing new Implicit Measures with existing Explicit Measures (Beta coefficients with 95% Confidence Intervals & p values)

	VAS Size	VAS Satisfaction	Avatar "Own BMI"	Avatar "Ideal BMI"	Difference between ideal & own avatar size	Outfits Task	Biscuits eaten	Word IAT	Images IAT
Word IAT (Greenwald D)	**-0.16 (-0.25, -0.07), p=0.001**	**0.08 (0.04, 0.12), p<0.001**	-0.02 (-0.04, 0.004), p=0.112	0.03 (-0.002, 0.06), p=0.064	**0.04 (0.01, 0.06) p=0.003**	0.04 (-0.06, 0.14), p=0.434	0.002 (-0.004, 0.009), p=0.459	-	-
Images IAT (Greenwald D)	-0.03 (-0.09, 0.03), p=0.333	0.03 (-0.00, 0.06), p=0.055	**-0.02 (-0.03, -0.003), p=0.019**	-0.01 (-0.03, 0.01), p=0.241	0.01 (-0.003, 0.03), p=0.127	0.001 (-0.07, 0.07), p=0.987	-0.001 (-0.005, 0.003), p=0.572	**0.20 (0.07, 0.33), p=0.002**	-
LDT: beauty – ugly RT	-3.50 (-12.3, 5.28), p=0.430	-1.71 (-5.79, 2.38), p=0.408	-1.63 (-3.78, 0.52), p=0.136	-0.35 (-3.41, 2.70), p=0.818	1.50 (-0.69, 3.69), p=0.178	0.26 (-9.09, 9.09), p=0.956	-0.27 (-0.84, 0.29), p=0.336	-0.14 (-19.3, 19.1), p=0.989	-20.4 (-49.8, 9.06), p=0.172
LDT: positive – neutral RT	3.89 (-12.6, 4.77), p=0.374	-0.53 (-4.50, 3.44), p=0.792	-1.62 (-3.78, 0.52), p=0.136	1.47 (-1.46, 4.40), p=0.322	-0.71 (-2.84, 1.42), p=0.509	-0.73 (-9.82, 8.35), p=0.873	0.36 (-0.27, 0.98), p=0.257	**19.2 (1.12, 37.3), p=0.038**	27.2 (-1.15, 55.6), p=0.060
LDT recall beauty – ugly	**-0.37 (-0.69, -0.04), p=0.030**	**0.18 (0.03, 0.32), p=0.019**	-0.04 (-0.12, 0.03), p=0.311	0.09 (-0.02, 0.20), p=0.123	**0.09 (0.01, 0.17), p=0.030**	0.25 (-0.10, 0.60), p=0.153	**0.02 (0.00, 0.05), p=0.048**	**0.77 (0.07, 1.47), p=0.032**	0.96 (-0.14, 2.07), p=0.087
LDT recall positive – neutral	-0.10 (-0.34, 0.14), p=0.396	0.06 (-0.05, 0.17), p=0.306	-0.05 (-0.11, 0.01), p=0.092	-0.03(-0.11, 0.05), p=0.461	0.04 (-0.02, 0.09), p=0.239	0.13 (-0.12, 0.38), p=0.299	-0.01 (-0.03, 0.00), p=0.090	0.08 (-0.43, 0.59), p=0.763	-0.09 (-0.89, 0.71), p=0.822

Table 16: Comparing new Implicit Measures with new Explicit Measures of Body Size & Satisfaction (Beta coefficients with 95% Confidence Intervals & p values)

4. Discussion

4.1 Main Findings

The Avatar method for measuring perceived own body size has good test-retest reliability and correlated very well with measured BMI. It was quick and easy to administer, and could be completed online. Discrepancies between Avatar measures of "own BMI" and "ideal BMI" also correlated well with existing measures of body satisfaction and again were quick and easy to administer.

The *Outfit Choice Questionnaire* also correlated well with existing measures, replicating results from Study 3 (Chapter 3), and may capture a behavioural outcome. However, weight of biscuits eaten did not relate to measures of size or satisfaction with size.

The novel implicit measure using an IAT Word task correlated well with existing measures of both body size and satisfaction with size. The IAT image task correlated with measured and perceived BMI, suggesting that objective and explicit self-report of BMI is highly related to implicit perception of own body size. The LDT measure did not correlate with any existing measures of body satisfaction; however recall of "beauty" versus "ugly" words was strongly associated with existing measures of body dissatisfaction.

4.2 Strengths and Limitations

Several novel measures of perception of, and satisfaction with, own body size were tested in a large group of healthy women and compared these new measures with existing, validated measures of body satisfaction. The novel measures have

various advantages over existing methods. The Avatar method of measuring perceived and ideal body size is simple and quick, and compared to other Avatar measures has greater ecological validity. The *Outfit Choice Questionnaire* is the first questionnaire to attempt to capture a behavioural measure of body satisfaction.

The implicit measures are the first in the field to attempt to capture implicit beliefs about own body size (Word and Image IAT) and satisfaction with own size (Word IAT, recall of LDT words). Implicit measures have the potential advantage over explicit measures that they may enable assessment of beliefs that people might attempt to conceal through a wish to conform to social norms, or into which they may have a lack of insight (Konstantakopoulos et al., 2011; Nisbett & Bellows, 1977). Such tools could potentially be used to predict people at risk of eating disorders, or evaluate the success of treatment and risk of relapse. It is also possible that such measures may be useful in predicting response to treatment early on, as has been the case in implicit measures of mood enabling very early prediction of response to treatment in people suffering from depression (Warren et al., 2015).

However there are limitations to this work. The measures in this study have not been repeated, so (apart from the two-stage body avatar assessment), there are no data on test-retest reliability. This work is with healthy women, so although there is a wide range of body satisfaction scores, there are no results for patients with current or previous eating disorders. The results are only applicable to those with a BMI in the healthy range, and the sample has a disproportionate number

of University of Oxford students: these factors may limit the extent to which results are generalizable to the population as a whole. However, broadening recruitment methods to include social media adverts meant that the sample did not solely consist of university students.

The Avatar modeling of body size used two starting points, one larger and one smaller than participant's own size, as participants may be anchored by the starting size of the model (Furnham & Boo, 2011); however it is not known how accurate participants would have been using the avatar if they had been given starting points at more extreme values. The experimenter did not watch the participant manipulating the Avatar, but did remain in the room, and this may have affected participant representation of their own size.

4.3 Future Directions

Having described in detail the new measures designed for use in capturing a range of explicit, implicit and behavioral beliefs about body size and satisfaction with own body, Chapter 5 will go on to describe the results of the intervention in Study 4.

CHAPTER 5: EFFECTS OF REPEATED EXPOSURE TO BODIES OF DIFFERENT SIZES ON PERCEPTION OF AND SATISFACTION WITH OWN SIZE

1. Introduction

As discussed in Chapter 2, Studies 1 and 2 had various limitations in terms of developing an intervention with potential for use as an adjunctive treatment for body dissatisfaction in women with eating disorders. One limitation was that the images used were not ecologically valid, as they were simply stretched photographs. A second limitation was that there was a limited range of outcome measures, as baseline and follow up data were all collected in the space of one session, and the outcome measures may have been affected by the demand characteristics of the study. The third limitation is that any clinical intervention would be likely to be given repeatedly, rather than as a one-off.

Study 4 was designed with the aim of overcoming these limitations. The newly-developed ecologically valid images of women's bodies described in Chapter 3 were used. A range of outcome measures was collected (including implicit measures which may be less affected by demand characteristics) at baseline (Day 1) (described in Chapter 4) and at follow up (Day 8). Finally, participants were asked to complete the training intervention twice a day for a week.

Participants were randomised to training using images of the same women modified to appear "underweight", "overweight" or "neither over nor underweight". The hypotheses, based on the results from Studies 1 and 2, were

that following exposure to images of "underweight", as opposed to "neither over nor underweight" or "overweight" bodies, participants would (1) judge others' bodies as larger, (2) judge their own bodies as larger, and (3) be less satisfied with their body size. The further hypothesis was that increase in satisfaction with body size would lead to (4) increased biscuit consumption following the task, and that (5) results would be consistent across the ranges of measures described in Chapter 4. Hypothesis 2, that exposure to bodies of different sizes would change perception of own body size, as measured using a VAS scale, was the pre-registered primary outcome measure

2. Methods

2.1 Participants

The recruitment of participants for Study 4 is described in Chapter 4. The study protocol, including defined primary outcome measures, was pre-registered on the Open Science Framework (https://osf.io/pwd36/). In the first session, demographic information on participants, including nationality, education level, menstrual cycle and media use (both traditional media and social media) was collected.

2.2 Measures

2.2.1 Morphed Photographic Figure Scale

In order to investigate whether their perception of the size of *others'* bodies was altered by a week of exposure to different-sized bodies, at both baseline (Day 1) and follow-up (Day 8), participants were asked to rate the 90 images in the

Morphed Photographic Figure Scale (described in Chapter 3) as over, under or neither over- nor under- weight. Participants were shown the 90 images of women's bodies in a randomised order, using EPrime software. The accompanying text (as in the online Study 3) asked: "Is this woman over-weight (o), under-weight (u) or neither over- nor under-weight (n)?", and participants were asked to select the appropriate key on the keyboard to identify the weight category.

2.2.2 Measures described in Chapter 4

At baseline (Day 1) and follow up (Day 8), participants completed the novel and existing measures of perception of own body size, and satisfaction with body size described in detail in Chapter 4. Briefly, these comprised: VAS measures of own size (from too thin (0) to too fat (10); this was the pre-registered primary outcome measure) satisfaction with own size (from very dissatisfied (0) to very satisfied (10)), and satiety (from very hungry (0) to very full (10)), Avatar "own BMI" and Avatar "ideal BMI", biscuit consumption, the Outfit Choice Questionnaire, the Lexical Decision Task, Lexical Decision Task - Recall, Implicit Association Test – Words, Implicit Association Test – Images, the EDE-Q (C. G. Fairburn & Beglin, 1994), and the BSSS (Pingitore et al., 1997).

2.2.3 Questionnaire measures of mood, anxiety and self-esteem

At both baseline and follow-up, participants also completed questionnaire measures of their mood, anxiety and self-esteem, which may affect the way they perceive their body, and may be altered as a result of changes in perception of

own body. They completed the following questionnaires: the Positive and Negative Affect Scale (PANAS) (Watson, Clark, & Tellegen, 1988), the State-Trait Anxiety Inventory (STAI) (State) (Spielberger, 1970), the Beck Depression Inventory (BDI) (A. T. Beck, Ward, Mendelson, Mock, & Erbaugh, 1961; A. T. S. Beck, R.A.; Brown, G.K., 1996), and the Rosenberg Self-Esteem Scale (RSES)(Rosenberg, 1965)(Rosenberg 1965).

2.2.4 Handedness questionnaire

At baseline only, participants also completed a questionnaire measure of handedness (Annett 1970), to check that this was randomly distributed between groups, since some tasks required responding with left or right hand.

2.3 Training Task Stimuli

Using the *Morphed Photographic Figure Scale* (described in Chapter 3), training images were selected for this study based on whether the images were categorised as underweight, overweight or neither over- nor under-weight by the 173 subjects in online Study 3. Results from the equivalent population to those recruited to the current study (i.e. female, BMI 18-25 kg/m^2, no current or previous eating disorder, aged 18-30 years) were used. Images for the "neither over nor underweight" group were chosen by selecting those versions of the images which the majority of participants rated as neither over nor underweight (all images selected were categorised as being neither over- nor underweight by over 90% of participants). "Underweight" images were chosen by selecting images that around half the online participants categorised as underweight, and "overweight" images where around half the online participants categorised them

as overweight. Images were selected in this way in order not to engender feelings of shock or revulsion in participants at being shown extremely emaciated or overweight women, and with the aim of having quite subtle differences between groups, such that subjects were unaware of the purpose of the experiment. The images were matched, in that the same women, manipulated to appear at different weights, was used in the three randomisation groups. Each group of stimuli contained 8 images of the same women, at weights categorised as either overweight, underweight or neither over- nor underweight, according to the randomisation group. More images would have been used, but only 8 women had morphed versions that were categorised into all three weight groups. Example stimuli are shown in Figure 7.

Figure 17: Examples of stimuli used (left: "underweight"; middle "neither over- nor underweight"; right "overweight")

2.4 Procedure

Participants attended two experimental sessions each lasting between 1 and 1.5 hours at the Department of Psychiatry, University of Oxford, on Day 1 and Day 8. Participants were asked to eat a light meal 2 hours before attending, and to avoid caffeine, alcohol and cigarettes in the 2 hours before the session. At the start of Session 1, participants gave informed consent for the tasks they would be undertaking; however they remained blind to the hypothesis and purpose of the study. Participants were randomised to one of the three groups using a computer generated random sequence (group assignment determined whether each

individual would be trained using images of overweight, underweight or neither over- nor underweight women). An individual not involved in any testing or participant contact (JD (see Statement of Contributions)) assigned one of the three training conditions to each participant number, so both experimenters and participants were blind to training condition.

In both Session 1 and Session 2, participants completed the tasks described above in "Measures". In Session 1, they were then briefly trained in how to use the laptop and the training task. They were asked to complete the five minute training task twice a day for a week. They were told that the laptop would record each trial, and they were invited to choose what times of the day would be convenient for them to complete the task. All participants agreed to receive automated text messages to their mobile phones at the times they had selected, to remind them to complete the task.

At the end of Session 2, participants were also weighed and their height was measured to allow their BMI to be calculated. They were thanked for their participation and given a gift of £60 to thank them for their time. They were also debriefed as to the purpose of the study.

2.5 The Training Task

The training task consisted of a one-back working memory task. The stimuli (photographs of 8 women, altered as described above to appear either overweight, underweight, or neither overweight nor underweight according to the randomisation group) were presented sequentially to participants, who were instructed to respond as to whether the image had the same identity as the one

presented immediately before, or a different identity. This was to ensure that participants paid attention to the images displayed, which has been shown to enhance adaptation effects (Rhodes et al., 2011). In the underweight condition, the images were all of women modified to appear underweight; in the overweight condition, the images were all of the same women, modified to appear overweight; in the 'neither over- nor underweight' condition, the images were all of the same women, morphed to appear neither over nor underweight. A one-back rather than a two-back task was used in this experiment as participants had reported that the two-back task was too difficult in Studies 1 and 2.

The training task was programmed by a programmer, JD (see Statement of Contributions), using HTML, javascript and stylesheets. The program was designed to run in a web browser without an internet connection, with the results stored offline within the browser. Each image was displayed for 2000ms with a 20ms inter-trial interval. To prevent low level visual adaptation, the images were spatially jittered by an amount varying randomly from the centre of the screen by up to +/- 51 pixels in the x coordinate and +/- 38 pixels in the y coordinate (approximately 10% in each direction of the 1024 by 768 pixel display). Image size was 562 by 765 pixels, with a viewing distance of approximately 1m. Failure to respond within the 2000ms presentation resulted in no response being recorded for that trial. Each training block consisted of 48 trials in which 8 images were the same as the preceding image. A total of 3 blocks, giving 144 individual trials, were presented, taking approximately 5 minutes to complete. Participants were given feedback on their accuracy half way

through and at the end of the session. If participants did not make any responses during the session, they were prompted that they needed to respond.

2.6 Statistical Analysis

Means, standard deviations, medians and interquartile ranges were calculated for all measures at baseline, as well as descriptive data on age, educational background and social media use. All measures, including those described at greater length in Chapter 4, are presented in Table 17 at baseline for the entire sample, and by randomization group.

The primary outcome measure was VAS judgement of own body size on Day 8, adjusted for Day 1 score. The effects of training on judgement of own size were calculated using linear regression to compare post-intervention scores, adjusted for pre-intervention scores. Randomisation groups were treated as a categorical measure. All secondary outcome measures were assessed using regression analysis to compare Session 2 scores, each adjusted for relevant Session 1 scores. Randomisation group was treated as a categorical measure throughout. Analyses were repeated after adjusting for the following potentially confounding variables: number of images seen during training, time since last training (in tertiles as this variable was not normally distributed), measured BMI, and time spent using media (in tertiles as this variable was not normally distributed). Power calculation indicated that three groups of 30 would give 90% power at an alpha level of 5% to detect an effect size of 0.38.

Analyses were completed using Stata 15.

3. Results

3.1 Descriptive Data

Study recruitment is shown in the CONSORT diagram (Figure 8), and baseline descriptive data overall and by randomisation group are presented in Table 17.

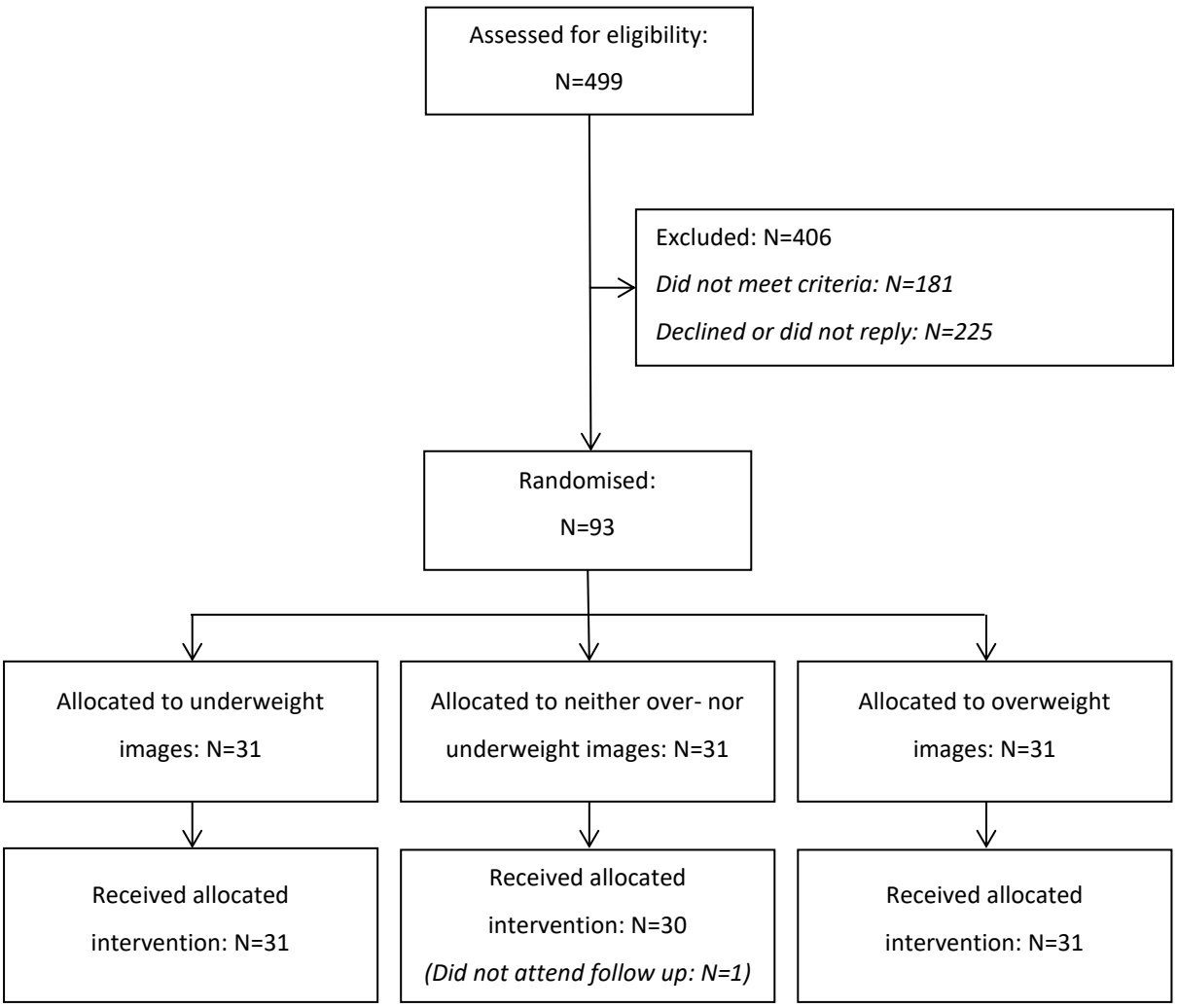


Figure 18: CONSORT diagram for Study 4

Participant characteristics	All Groups	Adapted to underweight	Adapted to neither over nor under-weight	Adapted to overweight
Age (mean, SD)	22.7 (3.15)	22.9 (3.57)	23.4 (2.94)	21.7 (2.73)
Measured BMI (kg/m^2) (mean, SD)	21.9 (2.11)	21.9 (2.16)	21.8 (2.20)	22.1 (2.02)
Spend >4 hours a day using media (N, %)	28 (30.1)	10 (32.3)	9 (29.0)	9 (29.0)
Number of Facebook friends (mean, SD)	527 (295)	525 (282)	508 (305)	547 (305)
Right handed (N,%)	77 (82.8)	23 (74.2)	26 (83.9)	28 (90.3)
Country of Origin UK (N,%)	70 (75.2)	23 (74.2)	23 (74.2)	24 (77.4)
Degree level education (N, %)	45 (48.4)	12 (38.7)	18 (58.1)	15 (48.4)
Taking Oral Contraceptive Pill (N, %)	58 (62.3)	19 (61.3)	17 (54.8)	22 (71.0)
Baseline VAS satiety (0-10) (mean, SD)	5.92 (1.41)	6.27 (1.44)	5.84 (1.37)	5.64 (1.37)
Baseline VAS own size (0-10) (mean, SD)	5.72 (0.94)	5.50 (0.73)	5.94 (1.07)	5.71 (0.96)
Baseline VAS satisfaction with own size (0-10) (mean, SD)	6.41 (2.02)	6.47 (1.85)	5.80 (2.04)	6.14 (2.16)
PANAS (positive) (mean, SD)	30.8 (6.68)	30.7 (6.63)	30.6 (6.29)	31.1 (7.27)
PANAS (negative) (mean, SD)	12.3 (2.82)	12.1 (2.11)	12.0 (2.61)	12.7 (3.59)
STAI (mean, SD)	28.9 (7.05)	28.2 (7.39)	28.2 (6.06)	30.2 (7.63)
BDI (mean, SD)	3.25 (3.86)	3.00 (2.89)	3.29 (3.91)	3.45 (4.70)
RSES (mean, SD)	22.4 (4.57)	22.4 (3.95)	22.3 (4.98)	22.4 (4.86)
N images seen during training (mean, SD)	1934 (259)	1935 (198)	1926 (273)	1941 (302)
Last training >6 hours before follow up (N, %)	30 (33.3)	9 (30.0)	11 (36.7)	10 (33.3)
Proportion correct responses in N-back (mean, SD)	0.62 (0.02)	0.62 (0.02)	0.61 (0.02)	0.62 (0.02)

Table 17: Descriptive data on groups at baseline

The majority of participants gave their country of origin as the UK; other participants were from Australia (3), Austria (2), Canada (3), Germany (2), India (2), Ireland (1), Jersey (1), Latvia (1), Malaysia (1), Malta (1), New Zealand (1), Pakistan (1), Russia (1) and the USA (3). The majority of participants described their highest level of education as having a University Degree; other participants had Level 3 NVQ (1), access to Higher Education (1), A levels (26), or a post-graduate qualification (20)

3.2 Change in perception of the size of others' bodies – Morphed Photographic Figure Scale

There was evidence for differences between training groups in the number of images of other women in the Morphed Photographic Figure Scale perceived as being overweight or underweight ($p<0.001$). Following training with underweight compared to neither over- nor underweight images, women categorised fewer images as underweight (6.43 fewer images (95% confidence interval [CI] -9.27, -3.59, $p<0.001$)) (Table 18; Figure 19), and 6.41 more images as overweight (95% CI 3.58, 9.24, $p<0.001$) (Table 18; Figure 20). In contrast, following training with overweight images, women viewed 7.42 more images as being underweight (95% CI 4.62, 10.2, $p<0.001$) (Figure 19; Table 18) and 5.92 fewer images as being overweight (95% CI-8.75, -3.09, $p<0.001$) (Figure 20; Table 18). These differences persisted following adjustment for potential confounding variables (number of images seen during training, time since last training session [in tertiles], measured BMI, time spent using media [in tertiles]) (Table 18).

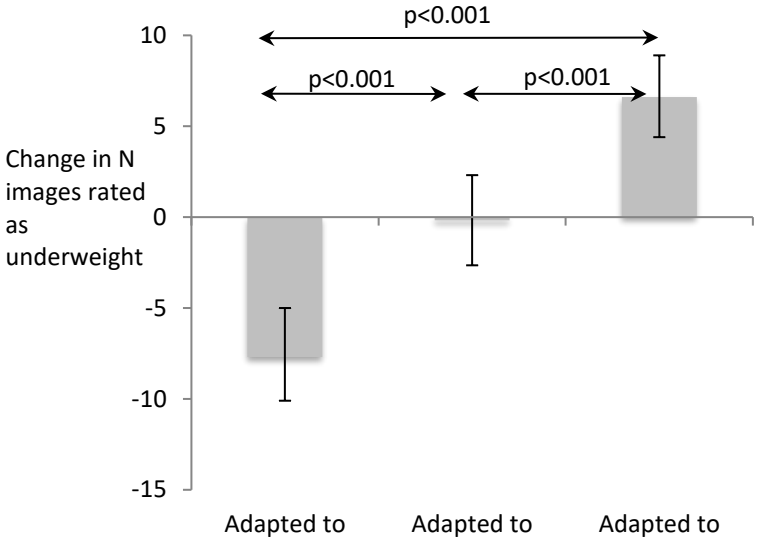


Figure 19: Change in N images categorised as "underweight", pre to post task, by group (error bars represent 95% confidence intervals)

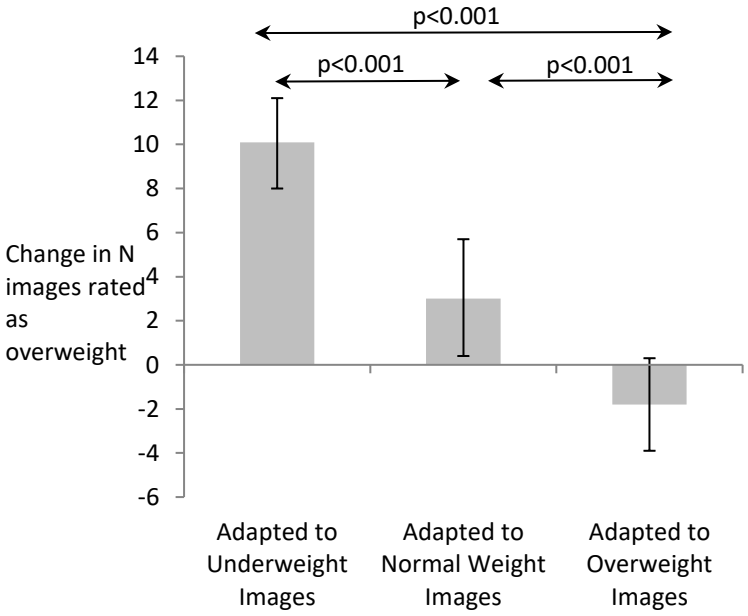


Figure 20: Change in N images categorised as "overweight" pre to post task, by group (error bars represent 95% confidence intervals)

3.3 VAS measures of size and satisfaction with size

There was no evidence of differences between training groups on the primary outcome measure of VAS own size (test of trend across groups: p=0.778); nor was there any evidence of a difference between training groups in VAS satisfaction with own size (test of trend across groups: p=0.179) (Table 18). This lack of between-group difference persisted following adjustment for the possible confounding variables described above (Table 18).

3.4 Avatar measures

There were differences between training groups in the "own BMI" avatars (test of trend p=0.007) (Figure 21; Table 18) and "ideal BMI" avatars (test of trend p=0.034) avatars (Figure 22; Table 18). Training with underweight images, in comparison with training with normal or overweight images, led to both a smaller "ideal BMI", and also to a smaller perceived "own BMI". Differences in "own BMI" persisted following adjustment for confounders; strength of evidence for differences in "ideal BMI" was reduced following adjustment (test of trend p=0.129) (Table 18).

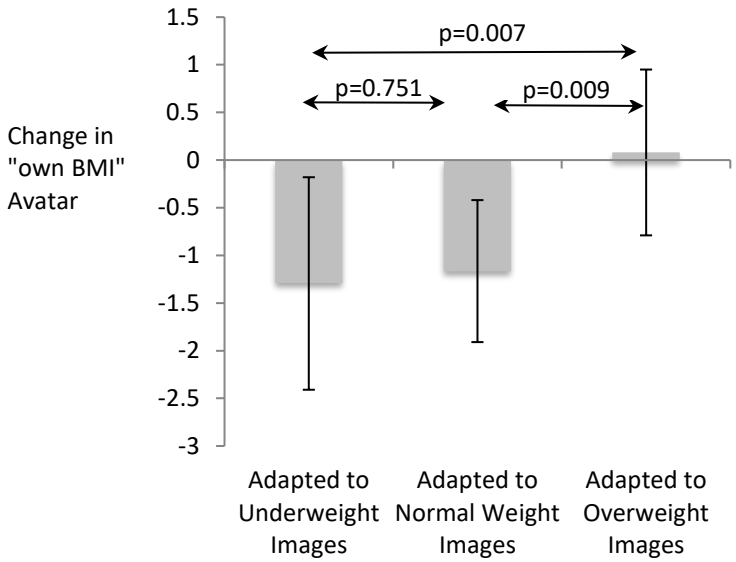


Figure 21: Change in BMI of "own BMI" Avatar, pre to post-task, by group (error bars represent 95% confidence intervals)

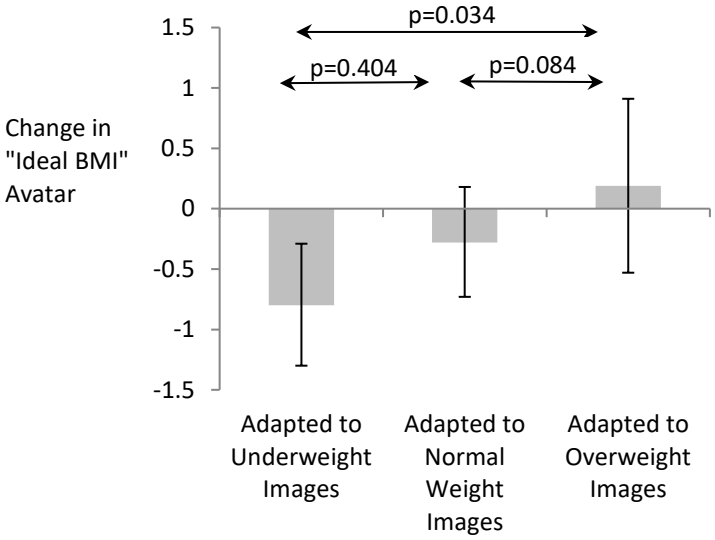


Figure 22: Change in BMI of "Ideal BMI" Avatar, pre to post-task, by group (error bars represent 95% confidence intervals)

	Comparison with those adapted to neither over- nor underweight images*			Adjusted** comparison with those adapted to neither over- nor underweight images (N=89)*		
	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups
VAS own size (N=92) (Day 8 adjusted for Day 1)	0.11 (-0.29, 0.51), p=0.577	-0.02 (-0.41, 0.38), p=0.937	p=0.778	0.01 (-0.40, 0.43), p=0.953	-0.07 (-0.48, 0.34), p=0.747	p=0.911
VAS satisfaction with own size (N=92) (Day 8 adjusted for Day 1)	0.16 (-0.56, 0.88), p=0.657	-0.48 (-1.20, 0.24), p=0.186	p=0.179	0.20 (-0.59, 0.98), p=0.622	-0.40 (-1.21, 0.41), p=0.324	p=0.303
N computer images viewed as underweight (N=91) (Day 8 adjusted for Day 1)	**-6.43 (-9.27, -3.59), p<0.001**	**7.42 (4.62, 10.2), p<0.001**	**p<0.001**	**-6.44 (-9.36, -3.51), p<0.001**	**7.43 (4.41, 10.5), p<0.001**	**p<0.001**
N computer images viewed as overweight (N=91) (Day 8 adjusted for Day 1)	**6.41 (3.58, 9.24), p<0.001**	**-5.92 (-8.75, -3.09), p<0.001**	**p<0. 001**	**6.41 (3.58, 9.24), p<0. 001**	**-5.91 (-8.75, -3.09), p<0.001**	**p<0.001**
"Own BMI" Avatar (N=90) (Day 8 adjusted for Day 1)	**-0.17 (-1.21, 0.88), p=0.751**	**1.40 (0.35, 2.44), p=0.009**	**p=0.007**	**-0.04 (-1.03, 0.94), p=0.928 (N=88)**	**1.49 (0.47, 2.52), p=0.005 (N=88)**	**p=0.004**
"Ideal BMI" Avatar (N=90) (Day 8 adjusted for Day 1)	**-0.31 (-1.03, 0.42), p=0.404**	**0.64 (-0.09, 1.36), p=0.084**	**p=0.034**	-0.23 (-0.96, 0.51), p=0.544 (N=88)	0.52 (-0.25, 1.28), p=0.181 (N=88)	p=0.129

***Linear Regression Analyses with Randomisation Group as the categorical exposure variable; **adjusted for N images seen, age, time since last training, measured BMI, time spent using media*

Table 18: Changes in perception of size of others, own size and satisfaction with own size (Study 4)

3.5 Questionnaire Measures

The only between-group difference in the questionnaire measures (the BSSS, EDEQ, STAI, PANAS, BDI and RSES) was weak evidence for higher PANAS positive scores following adaptation to smaller images (test for trend p=0.058); this was very slightly strengthened following adjustment for possible confounders (test for trend p=0.054) (Table 19). There was also a suggestion of higher self-esteem scores following adaptation to underweight images (test for trend p=0.044); the strength of this decreased following adjustment for possible confounders (p=0.105) (Table 19).

	Comparison with those adapted to neither over- nor underweight images* (N=92)			Adjusted** comparison with those adapted to neither over- nor underweight images* (N=89)		
	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups
BSSS	0.04 (-0.13, 0.20), p=0.649	-0.07 (-0.23, 0.09), p=0.400	p=0.411	0.04 (-0.13, 0.21), p=0.641	-0.06 (-0.24, 0.21), p=0.516	p=0.516
EDEQ	0.05 (-0.11, 0.20), p=0.559	0.03 (-0.11, 0.18), p=0.675	p=0.833	0.02 (-0.14, 0.18), p=0.83	0.001 (-0.16, 0.16), p=0.989	p=0.970
STAI	0.13 (-2.62, 2.88), p=0.926	1.33 (-1.44, 4.10), p=0.344	p=0.577	-0.08 (-2.95, 2.79), p=0.957	0.56 (-2.44, 3.56), p=0.711	p=0.894
PANAS Positive	1.57 (-0.97, 4.11), p=0.223	-1.50 (-4.04, 1.04), p=0.243	p=0.058	1.22 (-1.29, 3.73), p=0.335	-1.85 (-4.47, 0.76), p=0.162	p=0.054
PANAS Negative	-0.40 (-2.15, 1.35), p=0.651	0.83 (-0.92, 2.59), p=0.349	p=0.362	-0.44 (-2.23, 1.33), p=0.617	0.46 (-1.39, 2.32), p=0.622	p=0.594
RSES	**0.53 (-0.81, 1.87), p=0.436**	**-1.14 (-2.48, 0.20), p=0.095**	**p=0.044**	0.46 (-0.90, 1.83), p=0.500	-0.99 (-2.40, 0.43), p=0.170	p=0.105
BDI	0.19 (-1.21, 1.60), p=0.787	0.16 (-1.25, 1.56), p=0.823	p=0.959	0.09 (-1.19, 1.38), p=0.887	-0.72 (-2.06, 0.62), p=0.289	p=0.409

***Linear Regression Analyses with Randomisation Group as the categorical exposure variable; **adjusted for N images seen, time since last training, age, measured BMI, time spent using media*

Table 19: Changes in questionnaire measures of body satisfaction, mood, anxiety and self-esteem (Study 4)

3.6 Outfit Choice Questionnaire

There was no statistical evidence for between group differences using the Outfits Task (test for trend p=0.239) (Table 20).

3.7 Biscuit consumption

There was no statistical evidence for between group differences in terms of the number of biscuits consumed (test for trend p=0.519) (Table 20).

3.8 Lexical Decision Task

There was no evidence for between group differences in speed of reaction for "Beauty" minus "Ugly" words (Table 20), or for "Positive" minus "Neutral" words.

3.9 Lexical Decision Task Recall

There was no evidence for between group differences in LDT word recall (Table 20).

3.10 Implicit Association Test – Words

There was no evidence for between group differences in the Word IAT (test for trend p=0.453) (Table 20).

3.11 Implicit Association Test – Images

Following adjustment for possible confounding variables in the Images IAT, there was some evidence that participants were faster to respond to slightly overweight images paired with "self" words following adaptation to overweight images compared to other images (test for trend in adjusted analysis p=0.043) (Table 20).

	Comparison with those adapted to neither over- nor underweight images*			Adjusted** comparison with those adapted to neither over- nor underweight images*		
	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups	Adapted to underweight images	Adapted to overweight images	Test of Trend across groups
Outfit Choice Questionnaire (N=92)	0.06 (-0.14, 0.26), p=0.539	-0.11 (-0.30, 0.09), p=0.288	p=0.239	0.05 (-0.16, 0.26), p=0.665 (N=89)	-0.12 (-0.34, 0.10), p=0.268 (N=89)	p=0.274
Weight biscuits eaten (N=90)	-3.30 (-9.35, 2.75), p=0.281	-2.68 (-8.80, 3.43), p=0.385	p=0.519	-4.56 (-10.8, 1.72), p=0.152 (N=87)	-3.81 (-10.5, 2.84), p=0.258 (N=87)	p=0.321
IAT words (N=91)	-0.08 (-0.29, 0.12), p=0.423	-0.13 (-0.33, 0.08), p=0.216	p=0.453	-0.09 (-0.31, 0.13), p=0.425	-0.19 (-0.42, 0.03), p=0.091	p=0.238
IAT images (N=91)	0.01 (-0.14, 0.15), p=0.943	-0.12 (-0.26, 0.02), p=0.094	p=0.140	-0.002 (-0.15, 0.15), p=0.974	-0.17 (-0.33, -0.02), p=0.031	**p=0.043**
LDT reaction time to beauty – ugly words	-13.9 (-33.2, 5,42),p=0.156	-4.10 (-23.2, 14.9), p=0.670	p=0.156	-13.5 (-33.6, 6,68), p=0.187	-2.22 (-22.8, 18.4), p=0.830	p=0.361
LDT recall of beauty – ugly words (N=90)	0.21 (-0.45, 0.87), p=0.535	-0.02 (-0.64, 0.68), p=0.962	p=0.784	0.24 (-0.45, 0.93), p=0.49	0.12 (-0.59, 0.84), p=0.740	p=0.782

**Linear Regression Analyses with Randomisation Group as the categorical exposure variable; **adjusted for N images seen, age, time since last training, measured BMI, time spent using media*

Table 20: Changes in novel measures of perceived body size and satisfaction with size (Study 4)

4. Discussion

4.1 Main Findings

Perception of the size of others' bodies changed according to the images shown during training, with those shown "underweight" images subsequently viewing fewer images as underweight and more images as overweight, and vice versa. This was consistent with the findings from Studies 1 and 2, described in Chapter 2, as well as the wider evidence base. Avatar "ideal BMI" also changed following training, with those shown underweight images creating smaller "ideal BMI" avatars than those shown overweight images. These results were again in the expected direction.

However, contrary to the hypotheses, and in contrast to the results from Studies 1 and 2, there was no evidence of an effect of repeated exposure to images of different sizes on the primary outcome measure of VAS of perceived own size, or VAS satisfaction with own size.

There was no evidence for between training group differences in explicit questionnaire measures of body dissatisfaction, the *Outfit Choice Questionnaire*, biscuit consumption, or the LDT or word IAT tasks.

Participants created smaller "own size" avatars after training with underweight as opposed to overweight images, suggesting they viewed themselves as smaller following training with underweight images. Consistent with this, they had higher self-esteem scores following adaptation to the "underweight" images, and showed a trend to higher PANAS positive scores. They also showed faster

reaction times for overweight images and self-referent words after training with overweight images, and had a trend towards recalling fewer "beauty" related words after training with overweight images. This will be discussed at greater length in the Discussion (Chapter 7), but it suggests that some different processes may be in action when the training is repeated over the course of a week, and that participants may come to identify themselves in relation to the training images.

4.2 Strengths and limitations

This study builds on previous work on understanding the effect of exposure to images of women of different sizes on perception of and satisfaction with own size. The study was appropriately powered to find plausible between group differences, and both subjects and experimenters were blind to randomisation group. The stimuli had been carefully developed to be more ecologically valid than those used previously, and the images used had been rated by over 100 participants with similar characteristics as participants in this study, to ensure they were categorised as being over, under or neither over- nor under-weight. The use of a range of outcome measures meant that the study was not reliant purely on explicit measures which could be affected by demand characteristics.

Although attempts were made to recruit a wider range of participants, many respondents were students at either the University of Oxford or Oxford Brookes University, the sample was highly educated, and this may limit the generalisability of the study to other populations. Recruitment of women from the 18 to 30 age group also limits generalisability outside this age group.

Although the study was not solely reliant on measures that could be biased by demand characteristics, some measures may be influenced in this way by subjects who had identified the purpose of the study.

4.3 Future Directions

The nature of the results from this study, in particular the fact that they were inconsistent with the results found in Studies 1 and 2, meant that it did not seem appropriate to trial the intervention in people with a current eating disorder, as this may have led to negative consequences in terms of their perception of and satisfaction with their own size. However, the results described in Chapter 4 indicate that the new behavioural and implicit methods of measuring perception of and satisfaction with own body size do correlate well with existing questionnaire measures of perception of and satisfaction with own size. Therefore, Study 5 tests these new measures in a case control study, comparing results in those with a current diagnosis of an eating disorder, with those in age matched controls.

CHAPTER 6: HOW WOMEN WITH AND WITHOUT EATING DISORDERS PERCEIVE BODIES: A CASE-CONTROL STUDY

1. Introduction

Study 5 was designed in order to investigate the differences between women with and without eating disorders in their perception of their own and others' bodies. The aim was to investigate whether the new tasks tested in healthy women in Chapter 4 are able to differentiate between women with and without an eating disorder. Novel tasks where the results are strongly associated with presence or absence of an eating disorder are likely to be useful in furthering our understanding of body perception in eating disorders, which may help both predict response to treatment, and design new interventions.

The hypotheses were that there would be differences between Case and Control participants across the range of novel measures described in Chapter 4. Specifically, the hypotheses were that (1) the images in the *Morphed Photographic Figure Scale* would be more likely to be seen as overweight, and less likely to be seen as underweight by those with an eating disorder; (2) VAS measures would show that those with an eating disorder viewed themselves as bigger and were less satisfied with their bodies; (3) the difference between Avatar "own BMI" and measured BMI would be bigger in those with an eating disorder (i.e. Cases would inaccurately view their body as larger than it really was to a greater extent than Controls); (4) Avatar "Ideal BMI" would be smaller in Cases; (5) *Outfit Choice Questionnaire* scores would be lower in Cases; (6) Cases would be faster to respond to "ugly" than "beauty" words in the Lexical Decision Task, and would

recall more “ugly” than “beauty” words; (7) Cases would respond faster to “self” words paired with negative weight words, and with larger images in the Implicit Association Tasks.

2. Methods

2.1 Participants

Both Case and Control participants were recruited by email advertisement to undergraduate and graduate students and staff at the University of Oxford via their College Junior and Middle/Graduate Common Rooms and University departments; word-of-mouth; posters in Departments and Colleges; and advertisements on Facebook and Twitter. The study advert was also put up on the websites “Patients Active in Research” and on the Department of Psychiatry, University of Oxford.

In order to increase recruitment of participants with a current eating disorder, email advertisements were sent to a list of research-interested people with eating disorders held by RP (see Statement of Contributions).

Inclusion criteria required participants to be female and aged 16 to 45 years old, to speak English as a first language and to have a Body Mass Index (BMI) of less than 30kg/m^2; Case participants were required to have a current diagnosis of any eating disorder, including Anorexia Nervosa, Bulimia Nervosa, Other Specified Feeding or Eating Disorder or Binge Eating Disorder (confirmed using the EDE) (C. G. W. Fairburn, G.T. , 1993). Participants were excluded if they had a previous but no current eating disorder; had any current serious neurological illness (e.g.

MS, epilepsy); had a former brain injury with lasting consequences (e.g. amnesia); were currently pregnant; smoked more than 10 cigarettes per day; had dyslexia; or had taken part in the study of women's perceptions of their own and others bodies described in Chapter 4.

Control participants were also excluded if they were currently taking psychoactive medication that may affect cognitive function (e.g. street/illicit drugs, methylphenidate, St John's Wort), had a current diagnosis of another mental illness requiring treatment, or had a history of a previous eating disorder.

Potentially interested participants were directed to a brief online screening questionnaire which they completed to ensure that they met inclusion criteria. People meeting screening criteria were emailed to invite them to sign up for a time slot to participate.

2.2 Measures

2.2.1 Eating Disorder Examination

The Eating Disorder Examination (EDE) (C. G. W. Fairburn, G.T. , 1993) was completed with all participants to determine whether they had a current eating disorder. Where there was any diagnostic uncertainty, this was discussed with MO'C (see Statement of Contributions). Participants also completed the Clinical Impairment Assessment (CIA) (C. G. Fairburn et al., 2009).

2.2.2 Novel measures of body size and satisfaction

Participants completed the novel measures of perception of own and others' body size, and satisfaction with body size described in detail in Chapters 4 and 5.

Briefly, these comprised: the *Morphed Photographic Figure Scale*, VAS measures of own size (from too thin (0) to too fat (10)), satisfaction with own size (from very dissatisfied (0) to very satisfied (10)), and satiety (from very hungry (0) to very full (10)), Avatar "own BMI" and Avatar "ideal BMI", the *Outfit Choice Questionnaire*, the Lexical Decision Task, Lexical Decision Task - Recall, Implicit Association Test – Words, Implicit Association Test – Images.

2.2.3 SCID Screen

The Structured Clinical Interview for DSM-IV (SCID) (First, 1996) was completed with all participants to screen for other current and previous mental illness.

2.2.4 Questionnaire measures

In addition, participants completed the National Adult Reading Test (NART) (Nelson, 1982) to estimate verbal IQ in order to check that this was similar in both groups. They also completed the Body Shape Satisfaction Scale (BSSS) (Pingitore et al., 1997), along with questionnaire measures of their mood, anxiety and self-esteem: the Positive and Negative Affect Scale (PANAS) (Watson et al., 1988), the State-Trait Anxiety Inventory (STAI) (State) (Spielberger, 1970), the Beck Depression Inventory (BDI) (A. T. S. Beck, R.A.; Brown, G.K., 1996) and the Rosenberg Self-Esteem Scale (RSES) (Rosenberg, 1965). Participants also gave some basic demographic information including answering questions about their use of social media, and completed a questionnaire measure of handedness (Annett, 1970), to check that this was similar between groups, since some tasks required responding with left or right hand.

2.2.5 Height and Weight

Participants were asked whether they wanted to be told their weight and height or not, and were told they could step onto the scales backwards if they would prefer. They were reminded that weight varies during the day and between sets of scales.

2.3 Procedure

Participants attended a single session, lasting between 1.5 and 2 hours at the Department of Psychiatry, University of Oxford. Participants were asked to eat a light meal 2 hours before attending, and to avoid caffeine, alcohol and cigarettes in the 2 hours before the session. At the start of the session, participants gave informed consent for the tasks they would be undertaking; however they remained blind to the hypotheses and purpose of the study. Before attending, participants aged 16 or 17 were also sent a Patient Information Sheet (Parent/Carer version) and a Parent/Carer Consent form. They were asked to either bring their parent/carer with them to the session, or to bring the signed Parent/Carer Consent form with them, as well as going through the consent process themselves.

Participants completed the tasks described above in the following order:

1. NART
2. SCID
3. VAS Satiety, VAS Size, VAS Satisfaction
4. Avatar "own BMI"
5. Lexical Decision Task

6. Implicit Association Task (Words)
7. Implicit Association Task (Images)
8. Five-minute break
9. Lexical Decision Task (Recall)
10. Avatar "ideal BMI"
11. *Morphed Photographic Figure Scale*
12. *Outfit Choice Questionnaire*
13. Handedness Questionnaire
14. State Trait Anxiety Inventory
15. Positive and Negative Affect Scale
16. Beck Depression Inventory
17. Body Shape Satisfaction Scale
18. Rosenberg Self-Esteem Scale
19. Eating Disorder Examination
20. Clinical Impairment Questionnaire
21. Height
22. Weight

After completing these tasks, participants were debriefed. All participants were given a debriefing sheet which included details of organisations where they could seek help with any mental health problems, including eating disorders. Participants aged 16 or 17 were encouraged to share this information with their parent or guardian, and where there was felt to be an immediate risk to a 16 or 17 year olds health, this information would be shared with a parent or guardian.

2.4 Statistical Analysis

Means, standard deviations, medians and interquartile ranges were calculated for each group for all measures.

Differences between the Case and Control groups on each measure were tested for, using logistic regression analysis to calculate Odds Ratios (OR), 95% confidence intervals and P values for the differences between Cases and Controls. Cronbach's alpha was calculated to investigate the test-retest reliability of the avatar measures of "own BMI" and "ideal BMI".

The results from the LDT were analysed using the method described in the paper describing the previous use of the task (Smith et al., 2014), that is, excluding incorrect responses and responses with a reaction time of >1000msecs. Differences between Cases and Controls on their reaction times for Beauty minus Ugly words, and Neutral minus Positive words were tested. Results from the Implicit Association Tests were generated using the improved scoring algorithm described by Greenwald (Greenwald et al., 2003), including following their guidance as to which responses should be discarded as outliers. This method generates "Greenwald's D", analogous to an effect size calculation for differences between "congruent" and "incongruent" word/image pairings.

Data entry accuracy was checked by AP (see Statement of Contributions).

Analyses were completed using Stata 15.

2.5 Ethics

The study was approved by the Oxford University Medical Sciences Inter-Divisional Research Ethics Committee (Reference Number R45241/RE002).

3. Results

3.1 Recruitment and demographics

Sixty participants were recruited between 22nd June 2016 and 29th March 2018 (the recruitment period was extended by 8 months due to a period of maternity leave). A total of 392 women completed the eligibility screening questionnaire; 237 were excluded as they were not eligible; the remainder did not respond to an invitation to participate. Study recruitment is shown in the CONSORT diagram (Figure 23), and participant demographics are presented in Table 21.

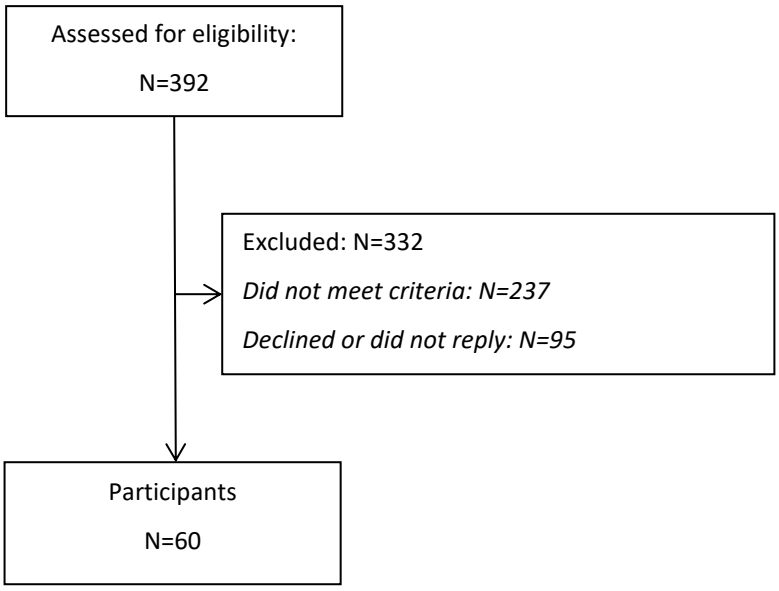


Figure 23: CONSORT diagram for Study 5

Participant Demographics	Cases (N=30)	Controls (N=30)
Age (mean, SD)	25.6 (7.25)	25.5 (7.55)
Verbal IQ (mean, SD)	114 (4.85)	116 (5.76)
Country of origin (N, %)		
UK/Ireland	27 (90)	22 (73.3)
USA	1 (3.33)	4 (13.3)
Canada	0	2 (6.67)
Gibralter	1 (3.33)	0
Nigeria	0	1 (3.33)
Iraq	1 (3.33)	0
Smoker (<5/day) (N, %)	2 (6.67)	2 (6.67)
Highest level of education (N, %)		
GCSE	2 (6.67)	1 (3.33)
A level	5 (16.7)	3 (10.0)
Degree	7 (23.3)	9 (30.0)
Post-graduate	3 (10.0)	4 (13.3)
Missing	13 (43.3)	13 (43.3)
Hours since last ate (mean, SD)	2.63 (2.44)	2.03 (0.55)
Measured BMI (mean, SD)	19.3 (4.13)	23.0 (2.19)
Handedness (N, % right handed)	27 (90)	26 (86.7)

Table 21: Participant Demographics

The Case and Control groups were well matched in terms of age, IQ score, smoking status and handedness. A few more of the Control participants came

from the US and Canada. Many participants did not answer the question about their highest level of educational attainment, but those who did were comparable.

The Control group were better at responding to our request that they ate a light meal around 2 hours before attending; two Case participants had eaten more than 10 hours before attending their session, in both individuals because due to their illness they struggled to eat anything until late in the day and hadn't eaten since the previous evening. As expected, mean BMI was lower in the Case group, but there was still a spread of BMI as participants with Bulimia Nervosa were included.

3.2 SCID diagnoses

In the Case group, three participants had current social anxiety, four had current General Anxiety Disorder; four had current panic disorder and two had current OCD. Eight were currently depressed, and three had current PTSD. Two had a diagnosis of bipolar affective disorder. One was alcohol dependent and three were occasionally using illegal/recreational drugs; none were intoxicated or in withdrawal during the testing session. In the Control group, four participants reported occasionally using illegal/recreational drugs; none were intoxicated or in withdrawal during the testing session.

Of the Cases, three had a previous diagnosis of panic disorder; three of OCD; three of PTSD; 15 of depression; four of recreational drug use and one of harmful use of alcohol). In the Control group, three had a previous diagnosis of panic disorder; five of PTSD; nine of depression and three of recreational drug use.

In the Case group, 10 participants were taking antidepressant medication. Two of these were also taking antipsychotic medication, and one was also taking a benzodiazepine. None of the Control group participants were taking psychoactive medication.

3.3 Eating Disorder Diagnoses

Table 22 shows the numbers of patients meeting diagnostic criteria for each eating disorder, as assessed using the EDE and DSM 5 diagnostic criteria. The duration of illness was somewhat skewed and so is given in terms of median and interquartile range. Table 23 shows the differences between groups on the continuous measures related to eating disorders, disordered eating behaviours and body dissatisfaction.

Eating Disorder	Cases (N=30)	Controls (N=30)
Current self-reported AN (N, %)	14 (46.7)	0
Current self-reported BN (N, %)	8 (26.7)	0
Current self-reported EDNOS (N, %)	6 (20.0)	0
Current EDE-diagnosed AN (N, %)	15 (50.0)	0
Current EDE-diagnosed BN (N, %)	6 (20.0)	0
Current EDE-diagnosed OSFED (AN) (N, %)	5 (16.7)	0
Current EDE-diagnosed OSFED (BN) (N, %)	3 (10.0)	0
Current EDE-diagnosed OSFED (BED) (N, %)	1 (3.33)	0
Duration of ED (years) (median, IQ range)	5.5 (3, 15)	0

Table 22: Description of eating disorders

Questionnaire Measure	Cases (N=30)	Controls (N=30)
EDE – Weight Concern scale (mean, SD)	3.18 (1.44)	0.7 (0.83)
EDE – Shape Concern scale (mean, SD)	3.68 (1.27)	1.13 (1.29)
EDE –Restraint scale (mean, SD)	3.33 (1.43)	0.65 (0.78)
EDE – Eating scale (mean, SD)	2.10 (1.16)	0.10 (0.16)
EDE Total (mean, SD)	3.07 (1.12)	0.58 (0.72)
CIA score (mean, SD)	27.5 (11.0)	3.97 (3.93)
BSSS (mean, SD)	2.13 (0.60)	3.30 (0.73)

Table 23: Measures of disordered eating behaviour and body dissatisfaction

3.4 Use of Media

Media Use	Cases (N=30)	Controls (N=30)
Social media (hours use per day) (mean, SD)	1.80 (1.09)	2.32 (2.57)
Use Facebook N (%)	28 (93.3)	30 (100)
Facebook friends (mean, SD)	471 (348)	595 (436)
Use Twitter N (%)	15 (50)	15 (50)
Twitter: N following (median, IQ range)	10 (0, 120)	0 (0, 100)
Twitter: N followers (median, IQ range)	0 (0, 70)	1.5 (0, 70)
Use Instagram N (%)	18 (58.6)	21 (70.0)
Instagram: N following (median, IQ range)	55 (0, 220)	60 (0, 230)
Instagram: N followers (median, IQ range)	50 (0, 500)	95 (0, 200)
Read magazines (N (%) who read magazines)	8 (26.6)	11 (36.6)
Magazines (hours reading per day) (median, IQ range)	0 (0, 0.25)	0 (0, 0.25)
Watch television (N (%) who watch television)	20 (66.6)	24 (80.0)
TV (hours watching per day) (median, IQ range)	1.00 (0, 3)	1.25 (0.5, 2)

Table 24: Media use

Social media use was fairly normally distributed within each group. The apparently higher mean hours use per day in the Control group was driven by one participant who reported using social media for 15 hours per day. Excluding this extreme outlier (all other Control participants reported use between 0.45 and 4 hours per day), gave a Control group mean of 1.88 (SD 0.9), very similar to that seen in the Case group.

The majority of participants were members of Facebook and their number of "Friends" on Facebook was normally distributed and similar between groups. Half of each group did not use Twitter, and the distribution of Twitter followers and numbers participants were following on Twitter was skewed by this; similarly many participants did not use Instagram, and the distributions of Instagram followers and followings was skewed by this. Hence descriptive statistics on Twitter and Instagram use are given in terms of median and interquartile range.

The majority of participants spent no time reading magazines, and a significant minority of participants watched no television – these results are therefore also described in terms of medians and interquartile ranges.

3.5 Other Questionnaire Measures

Questionnaires	Cases (N=30)	Controls (N=30)	Odds Ratio (OR) with 95% Confidence Intervals and p values
STAI – state	44.3 (12.4)	31.4 (6.17)	1.16 (1.07, 1.26), p<0.001
PANAS positive	24.2 (6.60)	30.4 (7.68)	0.89 (0.82, 0.96), p=0.003
PANAS negative	21.4 (8.40)	13.4 (4.10)	1.23 (1.10, 1.39), p=0.001
BDI	23.2 (12.5)	5.70 (5.58)	1.22 (1.10, 1.36), p<0.001
BSSS	2.13 (0.60)	3.30 (0.73)	0.06 (0.01, 0.25), p<0.001
RSES	12.6 (6.15)	21.3 (4.86)	0.78 (0.69, 0.88), p<0.001

Table 25: Questionnaire Measures

Table 25 shows the differences between groups in terms of their anxiety, depression, self-esteem and positive and negative affect at the time of the study.

STAI was normally distributed, and mean anxiety was higher in Cases than in Controls. PANAS scores were normally distributed and negative scores were higher in Cases, whilst positive scores were higher in Controls. Self-esteem (RSES) scores were normally distributed and higher in Controls. Depression (BDI) scores were negatively skewed with low scores in many participants, but were also lower in Controls than Cases.

3.6 Perception of weight of others' bodies (*Morphed Photographic Figure Scale*)

Table 26 shows the percentage of the *Morphed Photographic Figure Scale* bodies that were viewed by Cases and Controls as being underweight, and as being overweight. Both the number viewed as underweight, and the number viewed as overweight, were normally distributed. There were differences between Cases and Controls, with Cases believing that around 8% fewer bodies were underweight and 13% more bodies were overweight. These results persisted after adjusting for BMI.

MPFS: Others' bodies	Cases (N=30)	Controls (N=30)	Odds Ratio (OR) with 95% Confidence Intervals and p values
Percentage of MPFS bodies perceived as underweight (mean, SD)	20.3 (10.1)	29.2 (9.90)	0.92 (0.86, 0.97), p=0.003
Percentage of MPFS bodies perceived as overweight (mean, SD)	45.8 (12.7)	31.0 (9.53)	1.13 (1.06, 1.22), p<0.001

Table 26: Percentages of MPFS bodies perceived as over and underweight

3.7 VAS measures of size, satisfaction with size and satiety

VAS measures of satiety, size and satisfaction with size were all normally distributed. There were no differences in satiety between Cases and Controls, but there was evidence to suggest that Cases viewed themselves as being larger than Controls, and Cases being less satisfied with their body size than Controls (Table 27).

VAS scales	Cases (N=30)	Controls (N=30)	Odds Ratio (OR) with 95% Confidence Intervals and p values
VAS satiety (0-10) (mean, SD)	5.89 (1.94)	5.78 (1.75)	1.03 (0.78, 1.37), p=0.81
VAS size (0-10) (mean, SD)	7.15 (2.13)	6.04 (0.73)	1.57 (1.10, 2.25), p=0.014
VAS satisfaction (0-10) (mean, SD)	2.55 (1.96)	5.68 (2.02)	0.40 (0.24, 0.66), p<0.001

Table 27: VAS measures of size, satiety and satisfaction

3.8 Avatar measures

Cronbach's alpha for avatar demonstrated "own weight" was 0.92, and for "ideal weight" was 0.93. Avatar measures of "own BMI" and "ideal BMI" were all normally distributed. As shown in Table 28, and contrary to the hypothesis, Cases were not less accurate at estimating their size than Controls. The main difference between Cases and Controls was in their ideal size. Cases had an "ideal BMI" that was considerably lower than the "ideal BMI" for Control participants. Controls had an "ideal BMI" of 21.6, in the healthy range, whilst Cases had an "ideal BMI" of 15.5, which is substantially underweight.

In view of the decision to recruit participants with any eating disorder, which may have influenced the results in terms of body size perception, these analyses were also run including only those participants with a diagnosis of AN or OSFED (AN). No differences were found between this subset of cases and controls on accuracy of size representation; participants with AN/OSFED (AN) had an ideal BMI of 14.5 (SD 3.3).

Avatar Measures	Cases (N=30)	Controls (N=30)	Odds Ratio (OR) with 95% Confidence Intervals and p values
"Own BMI" (avatar) BMI	23.1 (8.48)	26.2 (4.09)	0.93 (0.86, 1.01), p=0.08
"Ideal BMI" (avatar) BMI	15.5 (3.67)	21.6 (3.19)	0.61 (0.48, 0.77), p<0.001
Actual measured BMI	19.3 (4.13)	23.0 (2.19)	
Accuracy of perception of own BMI (Actual BMI – Avatar BMI)	-3.80 (6.81)	-3.21 (3.40)	0.98 (0.89, 1.08), p=0.67
Difference between Avatar "own BMI" and "ideal BMI" (Ideal – Real)	-7.66 (8.46)	-4.67 (4.28)	0.93 (0.86, 1.01), p=0.10

Table 28: Avatar measures of "own" and "ideal" BMI

3.9 Outfit Choice Questionnaire

There was a difference between the mean outfit score in the Cases (mean 4.22, SD 0.82) compared to that in the Controls (mean 5.34, SD 0.80). Cases had a lower mean score than Controls (Odds Ratio 0.20 (95% CI 0.09, 0.46), p<0.001), indicating that they would feel less comfortable wearing a range of outfits. Results persisted following adjustment for measured BMI.

3.10 Lexical Decision Task and Recall

There was a trend towards a slight difference between Case and Control participants in their reaction times to beauty words minus reaction time to ugly words (Table 29), with Controls being faster to respond to Beauty words, whilst Cases were equally fast with both Beauty and Ugly words. There was no difference between groups in their reaction time to neutral words minus their reaction time to positive words. In the Recall task, there were no differences between groups in the total numbers of words recalled, or in the categories of words recalled, in either accurate or inaccurate word recall. Repeating analyses using AN/OSFED cases only did not change the results.

LDT	Cases (N=30)	Controls (N=29)	Odds Ratio (OR) with 95% Confidence Intervals and p values
RT for beauty - ugly words	0.74 (64.1)	26.9 (37.2)	1.01 (0.99, 1.03), p=0.06
RT for neutral - positive words	41.1 (48.8)	51.8 (41.8)	0.99 (0.98, 1.01), p=0.36
Total number words recalled accurately	5.33 (2.76)	4.59 (2.49)	
Total words written (including incorrect recall)	6.66 (2.47)	5.83 (2.28)	
Positive – Neutral Recall	0.33 (1.21)	0.48 (1.18)	0.90 (0.58, 1.39), p=0.63
Beauty – Ugly Recall	0.27 (2.23)	0.86 (1.64)	0.85 (0.65, 1.12), p=0.25
Positive – Neutral False Recall	-0.53 (1.25)	-0.23 (0.82)	0.75 (0.44, 1.27), p=0.28
Beauty – Ugly False Recall	-0.13 (0.78)	-0.27 (0.49)	1.27 (0.64, 2.54), p=0.50

Table 29: Lexical Decision Task and Recall

3.11 Implicit Association Test – Words & Images

IAT	Cases (N=30)	Controls (N=30)	Odds Ratio (OR) with 95% Confidence Intervals and p values
Greenwald D - Words	0.16 (0.54)	0.23 (0.42)	0.74 (0.25, 2.16), p=0.58
Greenwald D - Images	0.05 (0.31)	0.07 (0.35)	0.83 (0.17, 3.98), p=0.82

Table 30: Implicit Association Tasks

There were no differences between Cases and Controls in the Greenwald D score representing differences in response time to self-words paired with positive or negative weight words, or under or overweight images, as shown in Table 30. Repeating analyses using AN/OSFED cases only did not change the results.

4. Discussion

4.1 Main Findings

As hypothesised, there were considerable differences between those with and without an eating disorder in terms of how they viewed the weight status of others' bodies, with Cases assigning a greater number of bodies to the "overweight" category, and a smaller number of bodies to the "underweight" category.

VAS measures of size showed that Cases reported feeling that their bodies were bigger than Controls, and VAS satisfaction measures showed Cases felt less satisfied with their bodies.

However, using the Avatar method, Case "own BMI" avatars were as accurate as Control "own BMI" avatars – both groups overestimated their own BMI, but to the same extent as each other. Avatar "ideal BMI" sizes were however considerably smaller in Cases than Controls, with Case "ideal BMI" being dangerously underweight, whilst Control "ideal BMI" was in the healthy BMI range.

Using the Outfit Choice Questionnaire, there were considerable differences between Cases and Controls in their mean score, which was lower in Cases indicating that on average they would feel less comfortable wearing a range of outfits.

However, contrary to the hypotheses, there were no differences between Cases and Controls in their scores on any of the novel implicit measures, although there

was a trend suggesting that Controls respond faster to Beauty than Ugly words, whilst Cases respond equally fast to both.

4.2 Strengths and limitations

The Cases and Controls were well matched with each other in terms of age, verbal IQ and handedness. The EDE was used to determine whether or not participants had a current eating disorder, and this is the gold standard measure.

However, women with any eating disorder were recruited and this may explain some of the unexpected findings. It may be that there are differences between body perception in those with Anorexia Nervosa/OSFED – Anorexia Nervosa, compared with those with Bulimia Nervosa/OSFED – Bulimia Nervosa or Binge Eating Disorder (BED). Including participants with BED may have been particularly unhelpful from this point of view, since body dissatisfaction is not a key part of the diagnosis as it is in AN and BN. However, there is evidence that patients with BED also have high levels of body dissatisfaction (Grilo & Masheb, 2005), and binge eating is associated with higher levels of body dissatisfaction that those who are similarly overweight but do not binge eat (Grilo, Masheb, Brody, Burke-Martindale, & Rothschild, 2005). Excluding the single participant with BED, and indeed rerunning the analyses including only the twenty participants with either AN or OSFED (AN) in the Case group, did not substantially alter the results.

It is also possible that some tasks were influenced by demand characteristics of the study. For example, the percentage of women in the MPFS viewed by Control participants as being underweight was 29.2% in this study, whereas in the

Internet study described in Chapter 3, participants rated just 18.5% of the images as underweight, more in line with the view of the Case participants in this study. This might be explained by a wish to fit in with the expectations of the study. Recruitment of women from the 16 to 45 age group also limits generalisability outside this age group.

Another limitation is in the choice of Control group. Whilst they were well-matched with the Cases, they were also recruited via many of the same methods. The participants will therefore all have had some interest in how women view their bodies, and this may have led to a sample which included more people with higher levels of body dissatisfaction than might be found in the general population.

A further possible limitation was the decision to exclude Control participants with significant psychiatric comorbidity. Rates of comorbidity were high in the Case group, and so it may be that between group differences were due to differences in comorbidities rather than to differences in eating disorder pathology

4.3 Future Directions

The finding that the main difference between Case and Control participants appears to be their perception of what constitutes a "normal" weight in others, and a very low "ideal BMI" raises the possibility that the intervention tested in healthy women in Chapter 5 may in fact be useful in the treatment of women with a current eating disorder. The intervention tested in Chapter 5 does seem to be able to shift the perception of what constitutes a normal body weight in

others, and to move own ideal BMI in the expected direction, and if this effect could be achieved in women with a current eating disorder it may lead to improvements in their core pathology.

Other interesting avenues to pursue include tracking individuals with eating disorders over the course of their "treatment as usual", in order to see whether perception of "normal" body size, and own "ideal BMI" resolves with weight restoration, and to study these characteristics in a group of women who have recovered from an eating disorder. It would also be informative to track body perception in normal adolescence, in order to better understand how eating disorders develop – does a change in what is viewed as a normal body weight lead to developing an eating disorder, or does it come about as a result of body changes?

CHAPTER 7: DISCUSSION

1. Main findings

The main findings of the thesis are reiterated here, according to the hypotheses laid out at the end of the Introduction.

1. Following exposure to images of "underweight", as opposed to "normal" or "overweight" bodies, women

   a. judged others' bodies as larger, in both a single 15-minute intervention (Studies 1 & 2, Chapter 2) and in a week-long, repeated intervention (Study 4b, Chapter 5).
   b. judged their own bodies as larger in a single 15-minute intervention (Studies 1 & 2, Chapter 2); however, in a week-long, repeated intervention (Study 4b, Chapter 5) there were no differences in VAS measures of own size, and Avatar depiction of own size was smaller.
   c. were less satisfied with their body size, in a single 15-minute intervention (Studies 1 & 2, Chapter 2); however in a week-long, repeated intervention (Study 4b, Chapter 5) there were no changes in satisfaction, although Avatar "ideal BMI" was smaller.
   d. ate equivalent numbers of snacks following the tasks, in both a single 15-minute intervention (Studies 1 & 2, Chapter 2) and a week-long, repeated intervention (Study 4b, Chapter 5).

2. Changes in perception of others' body size, and satisfaction with own body size persisted to the following day, but change in perception of own size did not (Study 2, Chapter 2).

3. Women with lower BMI, and those who were less satisfied with their own weight, judged more photographs of other women as being overweight (Study 3, Chapter 3); those with a lower BMI also judged fewer women as being underweight.

4. In healthy women (Study 4a, Chapter 4):

   a. the Avatar method of demonstrating body size perception correlated well with measured BMI.
   b. The *Outfit Choice Questionnaire* correlated well with existing measures of body satisfaction; quantity of snacks eaten did not.
   c. The IAT Word task correlated well with existing measures of body size and satisfaction with size.
   d. The IAT Image task correlated with measured BMI.
   e. The LDT did not correlate with any existing measures of body satisfaction; however LDT recall of more "beauty" than "ugly" words was strongly associated with existing measures of body dissatisfaction

5. Women with a current eating disorder (Cases) in comparison with Controls (Study 5, Chapter 5):

   a. judged more bodies as being overweight, and fewer as underweight.

b. were no different in terms of misperception of own body size: both groups overestimated their body size to the same extent.
c. had a smaller ideal body size.
d. felt less satisfied with their bodies using the VAS measure and the *Outfit Choice Questionnaire*, as well as on existing measures of body satisfaction.
e. did not differ in terms of their scores on the IAT Word, IAT Images or LDT Recall, although there was a trend suggesting that Controls respond faster to Beauty than Ugly words in the LDT, whilst Cases respond equally fast to both.

2. Interpreting findings in relation to current literature

2.1 Perception of body size

2.1.1 Perception of others' body size

Studies 1, 2 and 4 showed very clearly that exposing women to images of women's bodies of different sizes alters their subsequent perception of the size of others' bodies. Exposure to larger bodies leads to subsequent perception of other bodies as smaller, and vice versa (Studies 1 and 2), and those shown "overweight" images subsequent view fewer images as "overweight" and more images as "underweight", and vice versa (Study 4). This is a strong effect, and one which has now been found in multiple studies (Brooks et al., 2016; Glauert et al., 2009; Mohr et al., 2016; Robinson & Christiansen, 2014; Robinson & Kirkham, 2014; Sturman et al., 2017; Winkler & Rhodes, 2005), and with multiple types of images of bodies (shifting aspect ratio (Winkler & Rhodes, 2005), computer-generated images (Glauert et al., 2009), computer-manipulated photographs ((Brooks et al.,

2016; Mohr et al., 2016; Sturman et al., 2017) and photographs of human bodies (Robinson & Christiansen, 2014; Robinson & Kirkham, 2014). Explicit feedback, such as that given in the Gledhill 2017 study (Gledhill et al., 2017), is not necessary for these effects to occur, and the effects are seen separately for muscle and fat (Sturman et al., 2017). Study 2 further confirms that these effects are also found in those with high baseline levels of body dissatisfaction.

Whilst the study methodology does not allow definitive specification of the mechanism of action of change in perception of the size of others' bodies, many authors have argued that this is an "adaptation effect", (Brooks et al., 2016; Glauert et al., 2009; Sturman et al., 2017; Winkler & Rhodes, 2005) as described at greater length in the Introduction. Adaptation effects describe repeated or extended exposure to smaller images leading to participants experiencing the perceptual aftereffect of viewing subsequent images as larger, and vice versa, and could plausibly be the underlying mechanism here.

A related interpretation follows from the "norm-based" coding of bodies described by Rhodes (Rhodes et al., 2013) in relation to body identity. Described at greater length in the introduction, the central idea is that representations of bodies, like those of faces, are "norm-based". That is to say that each dimension is coded by pairs of neural channels tuned to opposite extremes, with "normal" being represented by equal activation in both channels. Repeated exposure to bodies of different sizes leads to the body weight "norm" shifting, so that what is perceived as "normal" becomes more underweight following exposure to multiple

underweight images, and more overweight following exposure to multiple overweight images.

Study 3 showed that perception of body size in others varies according to both participant BMI, and the Weight Concern subscale of the EDE-Q. Participants with higher BMI judged more images of others' bodies as underweight, and fewer of them as overweight, even after adjusting for other variables, including EDE-Q Weight Concern. This finding is consistent with that from other studies: subjects with a lower BMI are more likely to overestimate the BMI of others' bodies (Tovee et al., 2000). Following the Rhodes (Rhodes et al., 2013) "norm-based" model, this suggests that own body size may be influential in determining the "norm", which is then used as a reference for individuals when they are deciding what constitutes "normal weight", as opposed to over or underweight.

Study 3 also showed that participants with EDE-Q Weight Concern scores in the clinical range judged more images of others' bodies as being overweight, a difference which persisted even after adjusting for participant BMI. Study 5 extended the results of Study 3, finding that participants with a current eating disorder judged a much greater number of bodies as being overweight, and fewer as being underweight, in comparison with a healthy control group. Once again, these results persisted after adjustment for participant BMI. These results suggest that participants with an eating disorder, or high levels of concern about their weight, have a lower set point for what they view as being "normal weight" than control groups, and this appears to be independent of their own BMI. This lower

threshold for what constitutes "normal" weight may contribute to dissatisfaction with their own weight.

The findings from Studies 3 and 5 contrast with a previous study (Tovee et al., 2000), which found no relationship between EDE-Q Weight Concern or diagnosed eating disorders, and accuracy of estimation of the BMI of others' bodies. However, this difference in findings may be due to the fact that Studies 3 and 5 investigated participant *opinion* on weight category, whereas the Tovee study required participants to estimate the actual BMI of the other bodies, and did not request an opinion as to whether the participants viewed bodies with these BMI's as being over or underweight.

2.1.2 Perception of own body size

Study 2, and the mega-analysis of Studies 1 and 2, found evidence that exposing participants to images of "overweight" as opposed to "underweight" or "normal weight" bodies led to them perceiving their own size as smaller. However, this was not replicated in Study 4b: repeated exposure to images morphed using Psychomorph software to appear of different weights did not alter the primary outcome measure of VAS perceived own size. Indeed, the opposite was the case: participants created smaller "own BMI" avatars after a week of exposure to "underweight" as opposed to "overweight" or "neither under- nor overweight" images, suggesting they viewed themselves as smaller following exposure to "underweight" images. They also showed faster reaction times for overweight images and self-referent words after training with overweight images. This result was also seen as a trend in the results from the Images IAT in Study 4b, where

participants shown underweight images subsequently responded more quickly to "Self" words in relation to "neither under- nor overweight" images.

The results from Study 2 are consistent with results found in a previous study in which participants were shown images of self and others (Brooks et al., 2016). The researchers found that exposure to either thin versions of self or thin versions of others led to subsequent images of both self and others being viewed as larger (and vice versa). However, the effects were larger when the exposure and testing were with the same body type (e.g. both self, or both other). These results can again be plausibly explained by the mechanism of adaptation.

However, "adaptation effects" cannot explain the results seen in Study 4, in which women repeatedly exposed to "underweight" as opposed to "overweight" or "neither under- nor overweight" images over the course of a week subsequently rated more images of other women as overweight, and yet created Avatars depicting themselves as smaller. However, these results may fit the Rhodes "norm-based" updating model (Rhodes et al., 2013), if one imagines that women have an internal representation of own size in relation to overall "population-norm". For example, imagine an individual woman who believes herself to be slightly bigger than the "population-norm". If her definition of "population-norm" is shifted to become smaller following exposure to "underweight" images (as it appears to be in Study 4 with regard to perception of others' size), her perception of her own size, if it is defined in relation to the "population-norm", would move with it. Thus, she would still perceive herself as slightly bigger than the

"population-norm", but since that "population-norm" would be smaller, she would also perceive her own body as smaller than it was before.

It is possible that different mechanisms may be involved due to the different time scales of Studies 1 and 2 and Study 4, with simple visual adaptation effects leading to the immediate changes in perception of own size seen in Studies 1 and 2, and a norm-updating model leading to the different effects seen in Study 4.

Another possible explanation for this difference relates to a slight change in methodology between Studies 1 and 2, and Study 4. In Studies 1 and 2, participants were instructed to look in the mirror following the task, with the aim that they would then relate the bodies they had seen in the task to their own body. This was not part of the instructions for Study 4, as it was felt that participants would be likely to come across their own reflection over the course of the week of training, and would apply the task images to their own bodies at these times. However, it may be that for images of other women to be related to own body, explicit self-reference with use of a mirror is necessary. On the other hand, since there were shifts in perception of own body size in Study 4 (in the opposite direction to that hypothesised) it seems likely that participants did relate the images shown to their own bodies.

2.2 Satisfaction with own body size

Studies 1 and 2 showed that satisfaction with own size changes following exposure to images of women's bodies of different sizes, with exposure to larger bodies resulting in increased satisfaction. This finding was stronger in Study 2, in which participants were more dissatisfied with their bodies at baseline.

This finding was not replicated in Study 4, in which there was no evidence of any difference between groups on the primary outcome measure of change in VAS-rated satisfaction with body size following the intervention. There was also no evidence for between-group differences in existing questionnaire measures of body dissatisfaction (EDEQ or BSSS scores) or in the LDT or Word IAT.

Some of the measures of satisfaction went in the opposite direction from that hypothesised, although in line with the finding that perceived own size decreased following exposure to "underweight" images. Following adaptation to the "underweight" images, participants had higher self-esteem scores and showed a trend to higher PANAS positive scores, suggesting that participants felt generally more positive after training with underweight bodies.

However, avatar "ideal BMI" changed following training in the hypothesised direction, with those shown underweight images subsequently creating smaller "ideal BMI" avatars than those shown overweight images. This would again fit with a norm-updating model, assuming that "ideal" is fixed in relation to "norm", so if "norm" gets smaller then so does "ideal" and vice versa.

In a review described at length in the Introduction (Grabe et al., 2008), multiple studies have shown that seeing images of underweight models leads to women feeling less satisfied with their bodies. The flaws of these studies, which largely concern a lack of appropriate control condition, or a failure to measure satisfaction at baseline, were extensively described in the Introduction, and limit the usefulness of these studies. It may be that the conflicting results found in Study 4 were due to the images used being of normal rather than idealised

women. This may lead to women associating their own body more strongly with these images than they would with glamorous images of models. However, normal women rather than glamorous models were also used in Studies 1 and 2, in which nonetheless satisfaction with own size *was* altered. The best explanation of these conflicting results seems to be that on a short time scale, visual adaptation effects lead to own body being perceived as larger following exposure to smaller bodies, and thus to increased satisfaction; over a week, repeated exposure to smaller bodies shifts the "norm", but also shifts perception of own and "ideal" body in the same direction, so the difference between actual and ideal size is not altered and therefore satisfaction with own size is unchanged.

2.3 Behaviours related to body satisfaction

None of the Studies 1, 2 and 4 found any effect of exposure to images of different sizes on chocolate or biscuit consumption. Following the null results in Studies 1 and 2, the snack was changed from Maltesers to small biscuits, because chocolate is a food which is often craved, and such cravings often involve both approach and avoidance behaviours, as well as guilt, making it difficult to interpret the meaning of eating larger or smaller amounts (Cartwright & Stritzke, 2008). However, no consistent effects were found when using small biscuits either. This may be due to individual variation in whether a change in satisfaction with own size leads to the consumption of more or less food. It appears that variations in the amount of food eaten, at least when using this brief measure of quantity of snacks consumed, is not a useful behavioural outcome measure.

Study 4 found no evidence for between-group differences in the behavioural measure of satisfaction, the *Outfit Choice Questionnaire*, which is consistent with the lack of change found in other measures of body satisfaction.

2.4 Duration of effects

Differences in perception of body size of others were still present at 24-hour follow up in Study 2. This was consistent with Study 4 findings: one third of participants had not done any training for more than 6 hours before the follow-up measurements, and yet the changes in perception of others' body size were not affected by this delay in measurement. This suggests that these effects on how others' bodies are perceived are quite long-lasting.

Changes in satisfaction with own size did persist in Study 2, although changes in perception of own size did not.

2.5 Measuring body size perception and satisfaction

2.5.1 Measuring perception of own body size

The Avatar method for measuring perceived own body size had good test-retest reliability and correlated very well with measured BMI. It was quick and easy to administer, and could be completed online. Discrepancies between Avatar measures of "own BMI" and "ideal BMI" also correlated well with existing measures of body satisfaction. The ecological validity of the bodies created using this measure meant that they may represent a better measure of how women perceive their bodies than avatars without such constraints.

2.5.2 Measuring satisfaction with body size

The *Outfit Choice Questionnaire* was validated against existing measures of body dissatisfaction and correlated well with these existing measures across three settings: Study 3 (online), Study 4 (in person) and Study 5 (clinical population). However, weight of biscuits eaten did not relate to measures of size or satisfaction with size.

The novel implicit measure using an IAT Word task correlated well with existing measures of both body size and satisfaction with size. The IAT image task correlated with measured and perceived (using Avatar measures) BMI, suggesting that objective BMI and self-report of BMI is highly related to implicit perception of own body size. The LDT measure did not correlate with any existing measures of body satisfaction; however recall of "beauty" compared to "ugly" words was strongly associated with existing measures of body dissatisfaction.

2.6 Differences between those with and without eating disorders

VAS measures of size showed that Cases reported feeling that their bodies were bigger than Controls, and VAS satisfaction measures showed Cases felt less satisfied with their bodies.

However, using the Avatar method, Case "own BMI" avatars were as accurate as Control "own BMI" avatars – both groups overestimated their own body size, but to the same extent as each other. This is interesting in the context of other work on body size estimation in eating disorders, the majority of which (Gardner & Brown, 2014) have found that those with an eating disorder showed greater

overestimation of their body size. It is notable that of the studies summarised here (Gardner & Brown, 2014), those not finding a difference between groups also used some version of modifying whole body size on screen, either using photographs of self (Urdapilleta, Cheneau, Masse, & Blanchet, 2007) or adjusting a drawing of a whole male or female body (Benninghoven, Tetsch, & Jantschek, 2008), although some studies using this type of technique did find between group differences (Cornelissen, Johns, & Tovee, 2013; Roy & Forest, 2007). Some of the lack of apparent difference in Study 5 may have been due to the decision to include Cases with Bulimia Nervosa as well as Anorexia Nervosa; however repeating analyses excluding those with Bulimia Nervosa did not alter results. It is also possible that differences between groups were not found here as a smaller number of participants were included that in some other studies. However, it should also be noted that a range of methodologies was used in the studies comprising this review (Gardner & Brown, 2014), and there has been no attempt at meta-analysis of previous results, or to investigate whether there is any apparent publication bias.

Avatar "ideal BMI" sizes were however considerably smaller in Cases than Controls, with Case "ideal BMI" being dangerously underweight, whilst Control "ideal BMI" was in the healthy BMI range.

Using the *Outfit Choice Questionnaire*, there were considerable differences between Cases and Controls in their mean score, which was lower in Cases, indicating that on average they would feel less comfortable wearing a range of outfits.

However, contrary to hypotheses, there were no differences between Cases and Controls in their scores on any of the novel implicit measures, although there was a trend suggesting that Controls respond faster to Beauty than Ugly words, whilst Cases respond equally fast to both.

3. Strengths and limitations

3.1 Studies aiming to modify body size and satisfaction

Studies 1, 2 and 4b build on previous research and used larger samples than previously used, to test whether exposure to bodies of different sizes can alter perception of own and others' body size, and satisfaction with own body. The studies were tightly controlled, in that all groups saw images of the same women, manipulated to appear in different sizes. Experimenters and participants were blinded to randomisation group and participants were unaware of the purpose of the study. Study 2 precisely replicated the methods of Study 1, and Studies 2 and 4 were both preregistered with defined hypotheses and primary outcome measures.

However, Studies 1 and 2 had some limitations. Firstly, the images created were not very sophisticated, merely varying the overall width of the bodies rather than altering their proportions in any other way; the images were not piloted to see whether they were viewed as realistic, and they were all of women of White ethnicity. However, the images were more sophisticated than those used in some previous studies, in that they were photographs of real women, rather than 2D or 3D sketches. Secondly, the wording of the task in which participants judged their own size included the phrases 'too fat' and 'too thin', which may have drawn on

their feelings of satisfaction with their body as well as their perception of their actual size. Thirdly, participants may have guessed the purpose of the study, and results may have been affected by demand characteristics; participant awareness of the study hypotheses was not assessed. Fourthly, the follow-up time in Study 2 was only 24 hours, and we therefore do not know whether effects may last longer than this. Fifth, participants were almost all students at the University of Bristol, and no data was collected on their ethnicity or socio-economic status; they therefore may not be representative of the national population of 18–25 year olds.

Another important limitation to consider concerns the size of the images used in the different categories. The "normal" images in Studies 1 and 2 comprised images of women with a known BMI of 22 to 23 kg/m^2, but it was not possible to calculate the BMI of the groups created by altering the widths of the images. The mean BMI of the participants (21.8 kg/m^2 in Study 1; 22.1 kg/m^2 in Study 2) was at the lower end of the range of the BMI of the 'normal' weight group of images (22 to 23 kg/m^2). At baseline, participants rated other images with a BMI of 22 to 23kg/m^2 as being towards the overweight end of the scale (mean of 4.78 on a 7-point scale (Study 1); 5.09 (Study 2)), despite this BMI being in the middle of the healthy range of BMI. Therefore, those completing the "normal" weight version of the task may have been exposed to images they considered to be slightly overweight. This does appear to be the case: exposure to "normal" weight images led to viewing subsequent computer images as smaller, perceiving own size as smaller and being more satisfied with own size.

In Studies 1 and 2, changes in perception of others' size appear to have been predominantly driven by decreases in perception of others' size following exposure to "normal" or "overweight" images. Exposure to the "underweight" images had less effect on subsequent perception of own size. This may be because the "underweight" images used in Studies 1 and 2 were in fact very similar in size to the media images surrounding us, and may be even be larger than average media images. For example, a study of Internet models found that over 50% had a BMI of ≤18 (Owen & Laurel-Seller, 2000).

Study 4b attempted to address the limitations of Studies 1 and 2. The stimuli used were carefully developed to be more ecologically valid than those used previously, and were also rated by over 100 participants with similar characteristics as participants, to ensure they were perceived as being over, under or neither over- nor under-weight. This appears to have corrected the problem seen in Studies 1 and 2, where the "normal" weight images had the effects expected with "overweight" images, and exposure to "underweight" images had no effect. In Study 4, the numbers of other's bodies perceived as being "over" and "underweight" altered in the predicted directions after exposure to both the "overweight" and "underweight" groups.

Furthermore, the use of a range of outcome measures, including implicit measures, meant that Study 4b results were not purely reliant on explicit measures which could be affected by demand characteristics.

However, limitations remained. Despite attempts to recruit a wide range of participants by broadening recruitment methods to include adverts on social

media, many Study 4b respondents were students at either the University of Oxford or Oxford Brookes University, and highly educated: therefore they may not be representative of the general population of 18-30 year-old women. Relatedly, limiting inclusion criteria to those aged 18-30, female, with no current or previous eating disorder, and with a BMI within the normal range mean that the findings may not be true of the population as a whole, and may not apply to those with an eating disorder. Finally, although the study did not solely use measures that could be biased by demand characteristics, some measures may have been influenced in this way by subjects who had identified the purpose of the study.

3.2 Design and testing of new measures

Study 3 was an online study to test how the *Morphed Photographic Figure Scale* images were perceived in terms of weight status, and how the *Outfit Choice Questionnaire* related to existing measures of body satisfaction and eating disorders. Strengths of Study 3 included its large size, and broad range of participants, as well as enabling new measures to be compared to existing validated measures. Limitations of Study 3 included again a limited group of responders, mostly UK-based and highly educated. Since the study was conducted online, there was no way of verifying that participants were who they said they were, and no incentive for them to respond accurately rather than, for example, continually pressing the same response button for each question in order to finish the study more quickly.

Study 4a tested several novel measures of perception of, and satisfaction with, own body size in a large group of healthy women and compared these new measures with existing, validated measures of body satisfaction. The novel measures have various advantages over existing methods. The Avatar method of measuring perceived and ideal body size is simple and quick, and compared to other Avatar measures has greater ecological validity. The *Outfit Choice Questionnaire* is the first questionnaire to attempt to capture a behavioural outcome of body satisfaction. The implicit measures are the first in the field to attempt to capture implicit beliefs about own body size (Word and Image IAT) and satisfaction with own size (Word IAT, recall of LDT words). Implicit measures have the potential advantage over explicit measures that they may enable assessment of beliefs that people might attempt to conceal through a wish to conform to social norms, or into which lack insight (Konstantakopoulos et al., 2011).

However, Study 4a did have limitations. The Avatar modeling of body size used two starting points, one larger and one smaller than participant's own size, as participants may be anchored by the starting size of the model; it is not known whether their representations of their bodies would have been different had they been given starting points at more extreme values. Secondly, although the experimenter did not watch the participant manipulating the Avatar, they did remain in the room, and this may have affected participant representation of their own size. Thirdly, with the exception of the avatar assessments, the measures in this study were not repeated, so there is no data on test-retest reliability.

Study 5 extended the testing of these new measures in Cases and Controls. Strengths of the study included that the two groups were well matched in terms of age, NART score and handedness. The EDE was used to determine whether or not participants had a current eating disorder, and this is the gold standard measure.

However, there were also limitations to Study 5. The inclusion of women with any eating disorder may explain some of the unexpected findings, although it is defensible in the context of the transdiagnostic approach to understanding eating disorders (C. G. Fairburn & Harrison, 2003). It may be that there are differences between body perception in those with Anorexia Nervosa/OSFED – Anorexia Nervosa, compared with those with Bulimia Nervosa/OSFED – Bulimia Nervosa. Secondly, it is possible that some tasks were influenced by demand characteristics of the study: for example, the percentage of women in the MPFS viewed by Control participants as being underweight was 29.2% in this study, whereas in the online study described in Chapter 3, participants rated just 18.5% of the images as underweight, more in line with the view of the Case participants in Study 5. This might be explained by the demand characteristics of the study, whereby Control participants were aware that they had been recruited because they did not have an eating disorder, and they may have believed that it would be a better fit with the expectations of the study to report more images as being underweight.

Another limitation is in the choice of Control group. Whilst they were well-matched with the Cases, they were also recruited via many of the same methods.

The participants therefore all had some interest in how women view their bodies, and this may have led to a sample which included more people with higher levels of body dissatisfaction than might be found in the general population. However, the rates of body dissatisfaction or eating disorder pathology in Controls do not seem to have been unusually high.

Once again, restricting recruitment to women from the 16 to 45 age group also limits the extent to which the results would generalise to other age groups and to men.

4. Implications

These findings are potentially important at a public health level, in view of the current media environment, where women are constantly exposed to images of underweight women. The studies have repeatedly demonstrated that even relatively short exposure (considerably less than the average amount of time that Study 4 participants reported spending using social media, reading magazines or watching TV) to images of normal women's bodies of different sizes changes perception of the size others' bodies, and that these effects last for at least 24 hours.

Whilst results were inconsistent with regard to whether exposure to different-sized bodies can also alter perception of own size and satisfaction with own size, such exposure does appear to shift what women consider to be an "ideal" body size. The main difference between women with and without an eating disorder in Study 5 was that women with an eating disorder classified more bodies as overweight and fewer as underweight. Although they didn't misperceive their

own body size to any greater extent than controls, they had a much smaller "ideal" body size and lower levels of body satisfaction.

In view of this, encouraging media outlets to choose models of a healthy body size would be likely to lead to women believing a more healthy body size to be both "normal" and "ideal". This may be helpful for those currently suffering from an eating disorder, whose perception of what constitutes a normal and ideal size is much smaller. Since the direction of causality in terms of whether smaller ideal body size leads to eating disorders, or vice versa, is not known, it is not possible to know whether such an intervention might be helpful in preventing eating disorders. However, there is evidence that body dissatisfaction is related to subsequent eating disorders (Jacobi et al., 2004; E. Stice et al., 2011), suggesting that smaller ideal body size may occur before eating disorders. Evidence suggests that choosing larger models would not diminish the effectiveness of advertising (Halliwell & Dittmar, 2004), so such changes would not have a negative impact in this area.

For those with a current eating disorder, the results when taken together suggest that exposure to more overweight bodies may be useful in redefining what they consider to be a normal and an ideal weight. Such an intervention may be a useful adjunct in eating disorder treatment, and may not require explicit feedback regarding what is "normal", as was used successfully in the Gledhill study (Gledhill et al., 2017) to shift perception of which bodies were perceived as "fat" or "thin", with resulting decreases in eating disorder symptomatology.

The hope in designing implicit measures of body perception and body satisfaction was that they may have had a use in predicting people at risk of eating disorders, or in predicting their likely response to treatment early on, as has been the case with implicit measures of mood enabling very early prediction of response to treatment in people suffering from depression (Warren et al., 2015). However, although in Study 4 the Word and Image IAT tasks did relate to existing measures of body satisfaction and body size, in the Case Control study (Study 5), the measures did not reliably differentiate between the groups, so it seems unlikely that these measures would be useful for these purposes.

5. Directions for future research

5.1 Trialling interventions in different groups

Women in the 18-30 age group have higher levels of eating disorders than males, (Smink et al., 2014), but the work described here could be usefully repeated in men, and indeed extended to a wider age group,.

Future research could also usefully investigate whether interventions such as those delivered in Studies 1, 2 and 4 are effective in shifting what patients with a diagnosed eating disorder view as being a normal and ideal weight, and testing whether any such shift results in a change in body satisfaction. Such work would be a necessary part of the pathway to developing a novel adjunctive treatment for eating disorders.

5.2 Understanding mechanisms

It is difficult to explain why despite all being exposed to a large number of underweight images in the media, not everyone becomes dissatisfied with their body size, and an even smaller number develop an eating disorder. One possible explanation is that this relates to attention. This thesis has shown that focusing attention on women's bodies of different sizes can alter the perception of what is perceived as being normal and ideal in terms of body size. Other research has found that women with high levels of body dissatisfaction attend preferentially to images of thin bodies (Cho & Lee, 2013). Participants with higher levels of body dissatisfaction have longer gaze fixation, and greater frequency of gaze to thinner bodies compared to participants with less body dissatisfaction, and this is associated with subsequently viewing other images as larger (Stephen, Sturman, Stevenson, Mond, & Brooks, 2018). It is possible that smaller normal and ideal body sizes in those with an eating disorder relate to them spending more time attending to images of, and indeed to real, women, who are of lower weight. Such an attentional bias could plausibly be on the causal pathway: women with this attentional bias may be more likely to develop an eating disorder. Alternatively, it could be that such biases might be the result of having an eating disorder or of being underweight. Research could usefully test whether such attentional biases exist in real-world settings, whether they are causally related to, or occur as a result of, body dissatisfaction and eating disorders, and whether they are amenable to attentional retraining, which would make them a possible treatment target.

Another mechanistic question arising from this work relates to whether glamorous models might have a different effect on participant view of and satisfaction with own size. Repeating these experiments using manipulated images of glamorous models would be useful in developing our understanding of the effect of images in the media on perception of "normal" and "ideal" size. It is possible that glamorous images may have a greater effect, as they are more attractive. Alternatively, they may have less effect if they seem less relevant to observers than images of other "normal" women.

5.3 Developing novel treatments

Although exposure to images of different sizes can shift perception of "normal" and "ideal" body size, it is possible that this is not sufficient to alter perception of own size. More intense experiences may be necessary to bring this about, such as experiencing one's own body as being of a different size using virtual reality. Early research suggests that such virtual reality experiences can change perceived own body size in both control participants and those with anorexia nervosa (Keizer, van Elburg, Helms, & Dijkerman, 2016).

5.4 The *Outfit Choice Questionnaire*

The *Outfit Choice Questionnaire* warrants further exploration as a novel way of capturing how women feel about their bodies. The next step would be to test the measure twice in a larger group of women both with and without an eating disorder, to obtain test-retest reliability data. It would also be valuable to measure it before, during and after treatment, alongside existing measures of body dissatisfaction, to confirm its validity in relation to current scales, and

establish its predictive validity in terms of outcomes. Key outfits from each factor which are better at discriminating between groups could be picked out in order to make the scale shorter to implement.

6. Conclusions

The thesis describes that exposure to images of bodies of different sizes reliably changes perception of other's body size. Exposure to overweight images of women's bodies leads to more subsequent images being perceived as underweight and vice versa. Effects on perception of own size are more complex and appear to vary with the time course of exposure. Attempts to develop implicit measures of body size and satisfaction with size have had mixed success: although in healthy women the new IAT measures showed good correlation with existing measures of body satisfaction, they did not successfully discriminate between those with and without eating disorders. However a novel behavioural measure of body satisfaction, the *Outfit Choice Questionnaire*, does show promise in its ability to distinguish between groups with and without a current eating disorder. The work provides a starting point for a range of further studies with potential future implications for the prevention and treatment of both body dissatisfaction and eating disorders.

REFERENCES

Addams, J. (1834). An account of a peculiar optical phaenomenon seen after having looked at a moving body, etc. *London and Edinburgh Philosophical Magazine and Journal of Science, 5*, 373-374.

Ahern, A. L., Bennett, K. M., & Hetherington, M. M. (2008). Internalization of the ultra-thin ideal: positive implicit associations with underweight fashion models are associated with drive for thinness in young women. *Eat Disord, 16*(4), 294-307. doi:10.1080/10640260802115852

Alleva, J. M., Sheeran, P., Webb, T. L., Martijn, C., & Miles, E. (2015). A Meta-Analytic Review of Stand-Alone Interventions to Improve Body Image. *PLoS One, 10*(9), e0139177. doi:10.1371/journal.pone.0139177

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th edition ed.). Arlington, VA: American Psychiatric Publishing.

Annett, M. (1970). A classification of hand preference by association analysis. *Br J Psychol, 61*(3), 303-321.

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders. A meta-analysis of 36 studies. *Arch Gen Psychiatry, 68*(7), 724-731. doi:10.1001/archgenpsychiatry.2011.74

Bacaltchuk, J., & Hay, P. (2003). Antidepressants versus placebo for people with bulimia nervosa. *Cochrane Database Syst Rev*(4), CD003391. doi:10.1002/14651858.CD003391

Bacaltchuk, J., Hay, P., & Trefiglio, R. (2001). Antidepressants versus psychological treatments and their combination for bulimia nervosa. *Cochrane Database Syst Rev*(4), CD003385. doi:10.1002/14651858.CD003385

Baker, J. D., Williamson, D. A., & Sylve, C. (1995). Body-Image Disturbance, Memory Bias, and Body Dysphoria - Effects of Negative Mood Induction. *Behavior Therapy, 26*(4), 747-759. doi:Doi 10.1016/S0005-7894(05)80043-7

Bearman, S. K., Martinez, E., Stice, E., & Presnell, K. (2006). The Skinny on Body Dissatisfaction: A Longitudinal Study of Adolescent Girls and Boys. *J Youth Adolesc, 35*(2), 217-229. doi:10.1007/s10964-005-9010-9

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Arch Gen Psychiatry, 4*, 561-571.

Beck, A. T. S., R.A.; Brown, G.K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Benninghoven, D., Tetsch, N., & Jantschek, G. (2008). Patients with eating disorders and their siblings. An investigation of body image perceptions. *Eur Child Adolesc Psychiatry, 17*(2), 118-126. doi:10.1007/s00787-007-0645-9

Boyce, J. A., Kuijer, R. G., & Gleaves, D. H. (2013). Positive fantasies or negative contrasts: the effect of media body ideals on restrained eaters' mood, weight

satisfaction, and food intake. *Body Image, 10*(4), 535-543. doi:10.1016/j.bodyim.2013.07.002

Brooks, K. R., Mond, J. M., Stevenson, R. J., & Stephen, I. D. (2016). Body Image Distortion and Exposure to Extreme Body Types: Contingent Adaptation and Cross Adaptation for Self and Other. *Front Neurosci, 10*, 334. doi:10.3389/fnins.2016.00334

Byrne, S., Wade, T., Hay, P., Touyz, S., Fairburn, C. G., Treasure, J., . . . Crosby, R. D. (2017). A randomised controlled trial of three psychological treatments for anorexia nervosa. *Psychol Med*, 1-11. doi:10.1017/S0033291717001349

Cartwright, F., & Stritzke, W. G. (2008). A multidimensional ambivalence model of chocolate craving: construct validity and associations with chocolate consumption and disordered eating. *Eat Behav, 9*(1), 1-12. doi:10.1016/j.eatbeh.2007.01.006

Cash, T. F., Cash, D. W., & Butters, J. W. (1983). Mirror, Mirror, on the Wall ... Contrast Effects and Self-Evaluations of Physical Attractiveness. *Personality and Social Psychology Bulletin, 9*(3), 351-358. doi:Doi 10.1177/0146167283093004

Cash, T. F., & Deagle, E. A., 3rd. (1997). The nature and extent of body-image disturbances in anorexia nervosa and bulimia nervosa: a meta-analysis. *Int J Eat Disord, 22*(2), 107-125.

Cattarin, J. A., Thompson, J. K., Thomas, C., & Williams, R. (2000). Body image, mood, and televised images of attractiveness: The role of social comparison.

Journal of Social and Clinical Psychology, 19(2), 220-239. doi:DOI 10.1521/jscp.2000.19.2.220

Cho, A., & Lee, J. H. (2013). Body dissatisfaction levels and gender differences in attentional biases toward idealized bodies. *Body Image, 10*(1), 95-102. doi:10.1016/j.bodyim.2012.09.005

Cornelissen, P. L., Johns, A., & Tovee, M. J. (2013). Body size over-estimation in women with anorexia nervosa is not qualitatively different from female controls. *Body Image, 10*(1), 103-111. doi:10.1016/j.bodyim.2012.09.003

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

Crouch, A., & Degelman, D. (1998). Influence of female body images in printed advertising on self-ratings of physical attractiveness by adolescent girls. *Perceptual and Motor Skills, 87*(2), 585-586. doi:DOI 10.2466/pms.1998.87.2.585

Cserjesi, R., Vermeulen, N., Luminet, O., Marechal, C., Nef, F., Simon, Y., & Lenard, L. (2010). Explicit vs. implicit body image evaluation in restrictive anorexia nervosa. *Psychiatry Res, 175*(1-2), 148-153. doi:10.1016/j.psychres.2009.07.002

de Vos, J., Houtzager, L., Katsaragaki, G., van de Berg, E., Cuijpers, P., & Dekker, J. (2014). Meta analysis on the efficacy of pharmacotherapy versus placebo on anorexia nervosa. *J Eat Disord, 2*(1), 27. doi:10.1186/s40337-014-0027-x

Dold, M., Aigner, M., Klabunde, M., Treasure, J., & Kasper, S. (2015). Second-Generation Antipsychotic Drugs in Anorexia Nervosa: A Meta-Analysis of Randomized Controlled Trials. *Psychother Psychosom, 84*(2), 110-116. doi:10.1159/000369978

Exterkate, C. C., Vriesendorp, P. F., & de Jong, C. A. J. (2009). Body attitudes in patients with eating disorders at presentation and completion of intensive outpatient day treatment. *Eat Behav, 10*(1), 16-21. doi:https://doi.org/10.1016/j.eatbeh.2008.10.002

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: interview or self-report questionnaire? *Int J Eat Disord, 16*(4), 363-370.

Fairburn, C. G., Cooper, Z., Doll, H. A., O'Connor, M. E., Bohn, K., Hawker, D. M., Palmer, R. L. (2009). Transdiagnostic cognitive-behavioral therapy for patients with eating disorders: a two-site trial with 60-week follow-up. *Am J Psychiatry, 166*(3), 311-319. doi:10.1176/appi.ajp.2008.08040608

Fairburn, C. G., & Harrison, P. J. (2003). Eating disorders. *Lancet, 361*(9355), 407-416. doi:10.1016/S0140-6736(03)12378-1

Fairburn, C. G. W., G.T. . (1993). *Binge Eating: Nature, Assessment and Treatment*. London, UK: Guildford Press.

Farrell, C., Lee, M., & Shafran, R. (2005). Assessment of body size estimation: A review. *European Eating Disorders Review, 13*(2), 75-88. doi:10.1002/erv.622

First, M. B. S., R.L.; Gibbon, M.; Williams, J.B.W. . (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders, Clinician Version (SCID-CV)*. Washington, DC: American Psychiatry Press.

Frederick, D. A., Peplau, L. A., & Lever, J. (2006). The swimsuit issue: Correlates of body image in a sample of 52,677 heterosexual adults. *Body Image, 3*(4), 413-419. doi:10.1016/j.bodyim.2006.08.002

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics, 40*(1), 35-42. doi:https://doi.org/10.1016/j.socec.2010.10.008

Gardner, R. M., & Brown, D. L. (2014). Body size estimation in anorexia nervosa: a brief review of findings from 2003 through 2013. *Psychiatry Res, 219*(3), 407-410. doi:10.1016/j.psychres.2014.06.029

Garner, D. M. (1991). *Eating Disorder Inventory-2: Professional Manual*: Odessa, FL: Psychological Assessment Resources.

Glauert, R., Rhodes, G., Byrne, S., Fink, B., & Grammer, K. (2009). Body dissatisfaction and the effects of perceptual exposure on body norms and ideals. *Int J Eat Disord, 42*(5), 443-452. doi:10.1002/eat.20640

Gledhill, L. J., Cornelissen, K. K., Cornelissen, P. L., Penton-Voak, I. S., Munafo, M. R., & Tovee, M. J. (2017). An interactive training programme to treat body image disturbance. *Br J Health Psychol, 22*(1), 60-76. doi:10.1111/bjhp.12217

Grabe, S., Ward, L. M., & Hyde, J. S. (2008). Role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. *Psychol Bull, 134*(3), 460-476. doi:10.1037/0033-2909.134.3.460

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *J Pers Soc Psychol, 74*(6), 1464-1480. doi:Doi 10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *J Pers Soc Psychol, 85*(2), 197-216.

Grilo, C. M., & Masheb, R. M. (2005). Correlates of body image dissatisfaction in treatment-seeking men and women with binge eating disorder. *Int J Eat Disord, 38*(2), 162-166. doi:10.1002/eat.20162

Grilo, C. M., Masheb, R. M., Brody, M., Burke-Martindale, C. H., & Rothschild, B. S. (2005). Binge eating and self-esteem predict body image dissatisfaction among obese men and women seeking bariatric surgery. *Int J Eat Disord, 37*(4), 347-351. doi:10.1002/eat.20130

Grogan, S., Gill, S., Brownbridge, K., Kilgariff, S., & Whalley, A. (2013). Dress fit and body image: a thematic analysis of women's accounts during and after trying on dresses. *Body Image, 10*(3), 380-388. doi:10.1016/j.bodyim.2013.03.003

Grover, V. P., Keel, P. K., & Mitchell, J. P. (2003). Gender differences in implicit weight identity. *Int J Eat Disord, 34*(1), 125-135. doi:10.1002/eat.10167

Haines, J., Neumark-Sztainer, D., Wall, M., & Story, M. (2007). Personal, behavioral, and environmental risk and protective factors for adolescent overweight. *Obesity (Silver Spring), 15*(11), 2748-2760. doi:Doi 10.1038/Oby.2007.327

Halliwell, E., & Dittmar, H. (2004). Does size matter? The impact of model's body size on women's body-focused anxiety and advertising effectiveness. *Journal of Social and Clinical Psychology, 23*(1), 104-122. doi:DOI 10.1521/jscp.23.1.104.26989

Hummel, D., Rudolf, A. K., Brandi, M. L., Untch, K. H., Grabhorn, R., Hampel, H., & Mohr, H. M. (2013). Neural adaptation to thin and fat bodies in the fusiform body area and middle occipital gyrus: An fMRI adaptation study. *Hum Brain Mapp, 34*(12), 3233-3246. doi:10.1002/hbm.22135

Hummel, D., Rudolf, A. K., Untch, K. H., Grabhorn, R., & Mohr, H. M. (2012). Visual adaptation to thin and fat bodies transfers across identity. *PLoS One, 7*(8), e43195. doi:10.1371/journal.pone.0043195

Image, A. P. P. G. o. B. (2012). *Reflections on Body Image*. All Party Parliamentary Group on Body Image and Central YMCA Retrieved from http://issuu.com/bodyimage/docs/reflections_on_body_image.

Jacobi, C., Hayward, C., de Zwaan, M., Kraemer, H. C., & Agras, W. S. (2004). Coming to terms with risk factors for eating disorders: application of risk terminology and suggestions for a general taxonomy. *Psychol Bull, 130*(1), 19-65. doi:10.1037/0033-2909.130.1.19

Kalodner, C. R. (1997). Media influences on male and female non-eating-disordered college students: A significant issue. *Eating Disorders: The Journal of Treatment & Prevention*(5), 47–57.

Keel, P. K., & Brown, T. A. (2010). Update on course and outcome in eating disorders. *Int J Eat Disord, 43*(3), 195-204. doi:10.1002/eat.20810

Keizer, A., van Elburg, A., Helms, R., & Dijkerman, H. C. (2016). A Virtual Reality Full Body Illusion Improves Body Image Disturbance in Anorexia Nervosa. *PLoS One, 11*(10), e0163921. doi:10.1371/journal.pone.0163921

Khalsa, S. S., Portnoff, L. C., McCurdy-McKinnon, D., & Feusner, J. D. (2017). What happens after treatment? A systematic review of relapse, remission, and recovery in anorexia nervosa. *J Eat Disord, 5*, 20. doi:10.1186/s40337-017-0145-3

Khan, S., & Petroczi, A. (2015). Stimulus-response compatibility tests of implicit preference for food and body image to identify people at risk for disordered eating: a validation study. *Eat Behav, 16*, 54-63. doi:10.1016/j.eatbeh.2014.10.015

Konstantakopoulos, G., Tchanturia, K., Surguladze, S. A., & David, A. S. (2011). Insight in eating disorders: clinical and cognitive correlates. *Psychological Medicine, 41*(9), 1951-1961. doi:10.1017/S0033291710002539

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med, 16*(9), 606-613.

Lin, L. F., & Kulik, J. A. (2002). Social comparison and women's body satisfaction. *Basic and Applied Social Psychology, 24*(2), 115-123. doi:Doi 10.1207/153248302753674622

Luce, K. H., & Crowther, J. H. (1999). The reliability of the Eating Disorder Examination-Self-Report Questionnaire Version (EDE-Q). *Int J Eat Disord, 25*(3), 349-351.

Micali, N., Hagberg, K. W., Petersen, I., & Treasure, J. L. (2013). The incidence of eating disorders in the UK in 2000-2009: findings from the General Practice Research Database. *BMJ Open, 3*(5). doi:10.1136/bmjopen-2013-002646

Mitchell, J. E., Roerig, J., & Steffen, K. (2013). Biological therapies for eating disorders. *Int J Eat Disord, 46*(5), 470-477. doi:10.1002/eat.22104

Mohr, H. M., Rickmeyer, C., Hummel, D., Ernst, M., & Grabhorn, R. (2016). Altered Visual Adaptation to Body Shape in Eating Disorders: Implications for Body Image Distortion. *Perception, 45*(7), 725-738. doi:10.1177/0301006616633385

Muennig, P., Jia, H., Lee, R., & Lubetkin, E. (2008). I think therefore I am: perceived ideal weight as a determinant of health. *Am J Public Health, 98*(3), 501-506. doi:10.2105/AJPH.2007.114769

Nelson, H. E. (1982). *National Adult Reading Test (NART): For the Assessment of Premorbid Intelligence in Patients with Dementia: Test Manual*. Windsor, UK National Foundation for Educational Research.

Neumark-Sztainer, D., Eisenberg, M. E., Fulkerson, J. A., Story, M., & Larson, N. I. (2008). Family meals and disordered eating in adolescents: longitudinal findings from project EAT. *Arch Pediatr Adolesc Med, 162*(1), 17-22. doi:10.1001/archpediatrics.2007.9

Neumark-Sztainer, D., Paxton, S. J., Hannan, P. J., Haines, J., & Story, M. (2006). Does body satisfaction matter? Five-year longitudinal associations between body satisfaction and health behaviors in adolescent females and males. *J Adolesc Health, 39*(2), 244-251. doi:10.1016/j.jadohealth.2005.12.001

NICE. (2017). *Eating disorders: recognition and treatment (NG69)*.

Nisbett, R. E., & Bellows, N. (1977). Verbal Reports About Causal Influences on Social Judgments - Private Access Versus Public Theories. *J Pers Soc Psychol, 35*(9), 613-624. doi:Doi 10.1037//0022-3514.35.9.613

Nisbett, R. E., & Wilson, T. D. (1977). Telling More Than We Can Know - Verbal Reports on Mental Processes. *Psychological Review, 84*(3), 231-259. doi:Doi 10.1037/0033-295x.84.3.231

Oakes, M. E., Slotterback, C. S., & Mecca, E. K. (2003). Gender differences in perceptions of normal body weights. *Current Psychology, 22*(2), 93-99. doi:DOI 10.1007/s12144-003-1000-z

Ogden, J., & Mundray, K. (1996). The effect of the media on body satisfaction: The role of gender and size. *European Eating Disorders Review, 4*(3), 171-182.

Owen, P. R., & Laurel-Seller, E. (2000). Weight and shape ideals: Thin is dangerously in. *Journal of Applied Social Psychology, 30*(5), 979-990. doi:DOI 10.1111/j.1559-1816.2000.tb02506.x

Paxton, S. J., Neumark-Sztainer, D., Hannan, P. J., & Eisenberg, M. E. (2006). Body dissatisfaction prospectively predicts depressive mood and low self-esteem

in adolescent girls and boys. *Journal of Clinical Child and Adolescent Psychology, 35*(4), 539-549. doi:DOI 10.1207/s15374424jccp3504_5

Penton-Voak, I. S., Thomas, J., Gage, S. H., McMurran, M., McDonald, S., & Munafo, M. R. (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychol Sci, 24*(5), 688-697. doi:10.1177/0956797612459657

Pingitore, R., Spring, B., & Garfield, D. (1997). Gender differences in body satisfaction. *Obes Res, 5*(5), 402-409.

Quick, V., Eisenberg, M. E., Bucchianeri, M. M., & Neumark-Sztainer, D. (2013). Prospective Predictors of Body Dissatisfaction in Young Adults: 10-year Longitudinal Findings. *Emerg Adulthood, 1*(4), 271-282. doi:10.1177/2167696813485738

Quick, V., Wall, M., Larson, N., Haines, J., & Neumark-Sztainer, D. (2013a). Personal, behavioral and socio-environmental predictors of overweight incidence in young adults: 10-yr longitudinal findings. *Int J Behav Nutr Phys Act, 10*, 37. doi:10.1186/1479-5868-10-37

Quick, V., Wall, M., Larson, N., Haines, J., & Neumark-Sztainer, D. (2013b). Personal, behavioral and socio-environmental predictors of overweight incidence in young adults: 10-yr longitudinal findings. *International Journal of Behavioral Nutrition and Physical Activity, 10*. doi:Artn 37 Doi 10.1186/1479-5868-10-37

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Res, 46*(18), 2977-2987. doi:10.1016/j.visres.2006.03.002

Rhodes, G., Jeffery, L., Boeing, A., & Calder, A. J. (2013). Visual coding of human bodies: perceptual aftereffects reveal norm-based, opponent coding of body identity. *J Exp Psychol Hum Percept Perform, 39*(2), 313-317. doi:10.1037/a0031568

Rhodes, G., Jeffery, L., Evangelista, E., Ewing, L., Peters, M., & Taylor, L. (2011). Enhanced attention amplifies face adaptation. *Vision Res, 51*(16), 1811-1819. doi:10.1016/j.visres.2011.06.008

Ricciardelli, L. A., & McCabe, M. P. (2001). Dietary restraint and negative affect as mediators of body dissatisfaction and bulimic behavior in adolescent girls and boys. *Behav Res Ther, 39*(11), 1317-1328.

Robinson, E., & Christiansen, P. (2014). The Changing Face of Obesity: Exposure to and Acceptance of Obesity. *Obesity (Silver Spring), 22*(5), 1380-1386. doi:10.1002/oby.20699

Robinson, E., & Kirkham, T. C. (2014). Is he a healthy weight? Exposure to obesity changes perception of the weight status of others. *Int J Obes (Lond), 38*(5), 663-667. doi:10.1038/ijo.2013.154

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

Roy, M., & Forest, F. (2007). Assessment of body image distortion in eating and weight disorders: the validation of a computer-based tool (Q-BID). *Eat Weight Disord, 12*(1), 1-11.

Schoemaker, C., van Strien, T., & van der Staak, C. (1994). Validation of the eating disorders inventory in a nonclinical population using transformed and untransformed responses. *Int J Eat Disord, 15*(4), 387-393.

Skinner, A. B., Helen; Woods, Andy; Etchells, Peter; Penton-Voak, Ian; Clark, Andrew; Harmer, Catherine J; Park, Rebecca; Broome, Matthew; Munafò, Marcus. (2017). The Morphed Photographic Figure Scale: Creation and Validation of a Novel Set of Realistic Female Body Stimuli. doi:10.31234/osf.io/9h73c

Smink, F. R., van Hoeken, D., Oldehinkel, A. J., & Hoek, H. W. (2014). Prevalence and severity of DSM-5 eating disorders in a community cohort of adolescents. *Int J Eat Disord, 47*(6), 610-619. doi:10.1002/eat.22316

Smith, A. R., Joiner, T. E., Jr., & Dodd, D. R. (2014). Examining implicit attitudes toward emaciation and thinness in anorexia nervosa. *Int J Eat Disord, 47*(2), 138-147. doi:10.1002/eat.22210

Sonneville, K. R., Calzo, J. P., Horton, N. J., Haines, J., Austin, S. B., & Field, A. E. (2012). Body satisfaction, weight gain and binge eating among overweight adolescent girls. *Int J Obes (Lond), 36*(7), 944-949. doi:10.1038/ijo.2012.68

Spielberger, C. D., Richard L. Gorsuch, and Robert Edward Lushene. . (1970). *State-trait anxiety inventory (self-evaluation questionnaire).* : Consulting Psycholologists Press.

Spring, V. L., & Bulik, C. M. (2014). Implicit and explicit affect toward food and weight stimuli in anorexia nervosa. *Eat Behav, 15*(1), 91-94. doi:10.1016/j.eatbeh.2013.10.017

Stephen, I. D., Sturman, D., Stevenson, R. J., Mond, J., & Brooks, K. R. (2018). Visual attention mediates the relationship between body satisfaction and susceptibility to the body size adaptation effect. *PLoS One, 13*(1), e0189855. doi:10.1371/journal.pone.0189855

Stice, E. (2001). A prospective test of the dual-pathway model of bulimic pathology: mediating effects of dieting and negative affect. *J Abnorm Psychol, 110*(1), 124-135.

Stice, E., Marti, C. N., & Durant, S. (2011). Risk factors for onset of eating disorders: evidence of multiple risk pathways from an 8-year prospective study. *Behav Res Ther, 49*(10), 622-627. doi:10.1016/j.brat.2011.06.009

Stice, E., Schupak-Neuberg, E., Shaw, H. E., & Stein, R. I. (1994). Relation of media exposure to eating disorder symptomatology: An examination of mediating mechanisms. *Journal of Abnormal Psychology, 103*(4), 836-840. doi:10.1037/0021-843x.103.4.836

Stunkard, A. J., Sorensen, T., & Schulsinger, F. (1983). Use of the Danish Adoption Register for the study of obesity and thinness. *Res Publ Assoc Res Nerv Ment Dis, 60*, 115-120.

Sturman, D., Stephen, I. D., Mond, J., Stevenson, R. J., & Brooks, K. R. (2017). Independent Aftereffects of Fat and Muscle: Implications for neural encoding, body space representation, and body image disturbance. *Sci Rep, 7*, 40392. doi:10.1038/srep40392

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., Stieger, S., Harris, A. S., Nader, I. W., Pietschnig, J., Voracek, M., & Tovee, M. J. (2012). Further Investigation of the Validity and Reliability of the Photographic Figure Rating Scale for Body Image Assessment. *J Pers Assess*. doi:10.1080/00223891.2012.660293

Swanson, S. A., Crow, S. J., Le Grange, D., Swendsen, J., & Merikangas, K. R. (2011). Prevalence and correlates of eating disorders in adolescents. Results from the national comorbidity survey replication adolescent supplement. *Arch Gen Psychiatry, 68*(7), 714-723. doi:10.1001/archgenpsychiatry.2011.22

Taylor, M. J., & Cooper, P. J. (1992). An Experimental-Study of the Effect of Mood on Body Size Perception. *Behav Res Ther, 30*(1), 53-58. doi:Doi 10.1016/0005-7967(92)90096-Y

Thompson, J. K. H., Leslie J.; Altabe, Madeline; Tantleff-Dunn, Stacey. (1999). *Exacting beauty: Theory, assessment, and treatment of body image disturbance*. Washington, DC, US: American Psychological Association.

Tiddeman, B., M. Burt, and D. Perrett,. (2001). Prototyping and transforming facial textures for perception research. . *Computer Graphics and Applications, 21*(5), 42-50.

Tiggemann, M., & Andrew, R. (2012). Clothing choices, weight, and trait self-objectification. *Body Image, 9*(3), 409-412. doi:10.1016/j.bodyim.2012.02.003

Tiggemann, M., & Lacey, C. (2009). Shopping for clothes: Body satisfaction, appearance investment, and functions of clothing among female shoppers. *Body Image, 6*(4), 285-291. doi:10.1016/j.bodyim.2009.07.002

Tovee, M. J., Emery, J. L., & Cohen-Tovee, E. M. (2000). The estimation of body mass index and physical attractiveness is dependent on the observer's own body mass index. *Proc Biol Sci, 267*(1456), 1987-1997. doi:10.1098/rspb.2000.1240

Urdapilleta, I., Cheneau, C., Masse, L., & Blanchet, A. (2007). Comparative study of body image among dancers and anorexic girls. *Eat Weight Disord, 12*(3), 140-146.

van den Berg, P., & Neumark-Sztainer, D. (2007). Fat 'n happy 5 years later: Is it bad for overweight girls to like their bodies? *Journal of Adolescent Health, 41*(4), 415-417. doi:DOI 10.1016/j.jadohealth.2007.06.001

Warren, M. B., Pringle, A., & Harmer, C. J. (2015). A neurocognitive model for understanding treatment action in depression. *Philos Trans R Soc Lond B Biol Sci, 370*(1677), 20140213. doi:10.1098/rstb.2014.0213

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol, 54*(6), 1063-1070.

Webster, M. A., & MacLeod, D. I. (2011). Visual adaptation and face perception. *Philos Trans R Soc Lond B Biol Sci, 366*(1571), 1702-1725. doi:10.1098/rstb.2010.0360

Wegner, B. S., Hartmann, A. M., & Geist, C. R. (2000). Effect of exposure to photographs of thin models on self-consciousness in female college students. *Psychol Rep, 86*(3 Pt 2), 1149-1154. doi:10.2466/pr0.2000.86.3c.1149

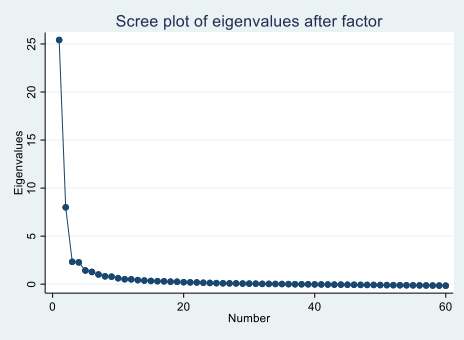
Wilcox, K., & Laird, J. D. (2000). The impact of media images of super-slender women on women's self-esteem: Identification, social comparison, and self-perception. *Journal of Research in Personality, 34*(2), 278-286. doi:DOI 10.1006/jrpe.1999.2281

Winkler, C., & Rhodes, G. (2005). Perceptual adaptation affects attractiveness of female bodies. *Br J Psychol, 96*(Pt 2), 141-154. doi:10.1348/000712605X36343

World Health Organisation (1992). *ICD-10 Classifications of Mental and Behavioural Disorder: Clinical Descriptions and Disgnostic Guidelines.* Geneva: World Health Organisation.

Appendix 1: Factor Analysis of Outfit Choice Questionnaire

Factor analysis was conducted to establish the structure of the *Outfit Choice Questionnaire*. Unforced principal factors analysis gave the scree plot of Eigenvalues shown in Appendix Figure 1 and the rotated version gave the results shown in Appendix Table 1.



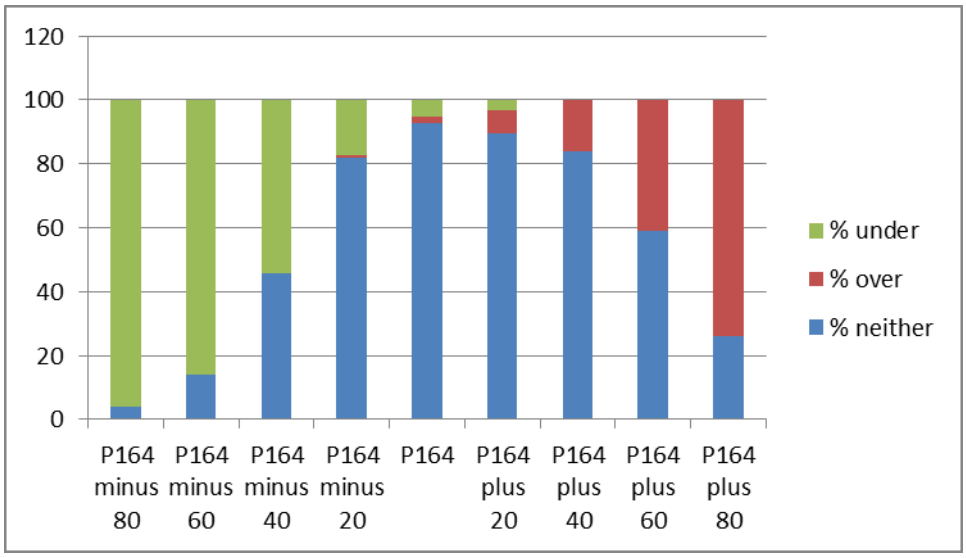
Appendix Figure 1: Scree plot of Eigenvalues after unforced principal factors analysis

Factor	Variance	Difference	Proportion	Cumulative Proportion
Factor 1	11.27	2.94	0.24	0.24
Factor 2	8.32	3.47	0.18	0.42
Factor 3	4.85	0.37	0.10	0.52
Factor 4	4.48	0.45	0.10	0.62
Factor 5	4.04	0.16	0.09	0.70
Factor 6	3.88	1.86	0.08	0.79
Factor 7	2.01	0.01	0.04	0.83
Factor 8	2.00	1.03	0.04	0.87
Factor 9	0.98	0.12	0.02	0.89
Factor 10	0.85	0.05	0.02	0.91

Appendix Table 1: Proportion of variance explained by Factors 1-10 in unforced principal factors analysis

Appendix 2

Appendix Figure 2: Example of weight categorisation of one group of Study 3 images by participants meeting Study 4 criteria



Appendix 3: Words About Bodies

A brief online study was conducted to generate appropriate words for use in the word IAT.

Participants were presented with 19 words chosen because they were related to weight and shape (toned, fit, slender, slim, trim, lean, lithe, curvaceous, shapely, thin, skinny, obese, fat, flabby, chunky, tubby, podgy, stout, chubby and plump). For each word, they were asked "How negative or positive is this word when used to describe a woman's body?" with scoring options from 1 (very negative) to 6 (very positive). For each word, they were then asked how old they thought they were when they first learnt it.

The questionnaire ran online from 3rd July 2015 to 24th July 2015, and 164 responses were received from 18-30 year old women with English as their first language. Using their responses to these two questions, and the characteristics of the words in terms of length and number of syllables, positive and negative body words were matched on length, estimated age of acquisition, and median score (positive words scored a median of 5 or 6; negative words scored a median of 1 or 2).

The study was approved by the Oxford University Central University Research Ethics Committee (MS-IDREC-C1-2015-123).

ROYAL SOCIETY OPEN SCIENCE

rsos.royalsocietypublishing.org

Research

Cite this article: Bould H *et al*. 2018 Effects of exposure to bodies of different sizes on perception of and satisfaction with own body size: two randomized studies. *R. Soc. open sci.* **5**: 171387.
http://dx.doi.org/10.1098/rsos.171387

Received: 14 September 2017
Accepted: 2 April 2018

Subject Category:
Psychology and cognitive neuroscience

Subject Areas:
psychology/cognition/behaviour

Keywords:
body, weight, body dissatisfaction, eating disorders, perception, body size

Author for correspondence:
Helen Bould
e-mail: helen.bould@psych.ox.ac.uk

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4075349.

THE ROYAL SOCIETY PUBLISHING

Effects of exposure to bodies of different sizes on perception of and satisfaction with own body size: two randomized studies

Helen Bould[1,2], Rebecca Carnegie[2], Heather Allward[3], Emily Bacon[3], Emily Lambe[3], Megan Sapseid[3], Katherine S. Button[4], Glyn Lewis[5], Andy Skinner[3], Matthew R. Broome[1,6], Rebecca Park[1], Catherine J. Harmer[1], Ian S. Penton-Voak[3] and Marcus R. Munafò[3,7]

[1]Department of Psychiatry, University of Oxford, Oxford, UK
[2]Centre for Academic Mental Health, and [3]School of Experimental Psychology, University of Bristol, Bristol, UK
[4]Department of Psychology, University of Bath, Bath, UK
[5]Division of Psychiatry, University College London, London, UK
[6]Institute for Mental Health, University of Birmingham, Birmingham, UK
[7]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

HB, 0000-0001-8163-3210; CJH, 0000-0002-1609-8335; MRM, 0000-0002-4049-993X

Body dissatisfaction is prevalent among women and associated with subsequent obesity and eating disorders. Exposure to images of bodies of different sizes has been suggested to change the perception of 'normal' body size in others. We tested whether exposure to different-sized (otherwise identical) bodies changes perception of own and others' body size, satisfaction with body size and amount of chocolate consumed. In Study 1, 90 18–25-year-old women with normal BMI were randomized into one of three groups to complete a 15 min two-back task using photographs of women either of 'normal weight' (Body Mass Index (BMI) 22–23 $kg\,m^{-2}$), or altered to appear either under- or over-weight. Study 2 was identical except the 96 participants had high baseline body dissatisfaction and were followed up after 24 h. We also conducted a mega-analysis combining both studies. Participants rated size of others' bodies, own

size, and satisfaction with size pre- and post-task. Post-task ratings were compared between groups, adjusting for pre-task ratings. Participants exposed to over- or normal-weight images subsequently perceived others' bodies as smaller, in comparison to those shown underweight bodies ($p < 0.001$). They also perceived their own bodies as smaller (Study 1, $p = 0.073$; Study 2, $p = 0.018$; mega-analysis, $p = 0.001$), and felt more satisfied with their size (Study 1, $p = 0.046$; Study 2, $p = 0.004$; mega-analysis, $p = 0.006$). There were no differences in chocolate consumption. This study suggests that a move towards using images of women with a BMI in the healthy range in the media may help to reduce body dissatisfaction, and the associated risk of eating disorders.

1. Background

Body dissatisfaction is a risk factor for eating disorders [1], low mood, obesity and weight gain in adolescents and young adults of both genders [2–5]. It is common, with prevalence studies finding 35% of adolescent girls [6], and two-thirds of adult women [7], are dissatisfied with their bodies. One main component of body dissatisfaction relates to body size, with over 75% of adolescent girls [8], and over 95% of adult women [9] reporting a wish to be thinner, and 74% of adult women reporting a wish to lose weight [10]. This is often ascribed to 'internalization of the thin ideal' (the idea that individuals, to differing amounts, try to conform to media-driven, socially defined ideas of attractiveness, including body size) [11,12]. If it were possible to alter perception of own body size, this could potentially be used as a way of reducing body dissatisfaction and its negative consequences, including obesity and weight gain [5], in the general population.

There is some evidence that exposing people to images of bodies of different sizes alters their perception of *others*' body size, in that it changes their idea of what is considered to be a 'normal' size for others' bodies. The female body size viewed as being most 'normal' becomes thinner after exposure to thinner bodies, and larger following exposure to larger bodies [13,14]. Viewing photographs of males whose body mass index (BMI) is in the obese, as opposed to normal, category leads to both male and female participants judging a man with a BMI in the overweight range to have a healthy, normal weight [15,16]. Although two of these studies included only images of women, and two only images of men, the results across both studies were consistent in terms of a change in perception of what constitutes a 'normal' body size in others. The body size perceived as most attractive also became smaller following exposure to thinner as opposed to larger bodies [13,14].

However, studies looking at the impact of exposure to bodies of different sizes on perception of *own* body size have been small (*N*s of 14 and 16) [17,18]. Participants in these two studies were shown 'thin' or 'fat' modified photographs of self or others, and were then asked to pick the most accurate photograph of themselves: those shown thinner-than-actual photographs of self or others judged a thinner-than-actual picture of self as most accurate; those shown fatter-than-actual photographs of self or others subsequently judged a fatter-than-actual photograph as most accurate. This suggests that photographs of others may change perception of own size, but the studies were too small to be confident that these results did not arise by chance.

In addition to whether exposure to bodies of different sizes changes perception of own body *size*, another important question is whether it changes *satisfaction* with own body size. Meta-analysis [19] concludes that in general women feel less satisfied with their bodies after viewing photographs of thin models compared to neutral images or normal-sized models. However, of the 80 studies included in this meta-analysis, 75 either had no control group, or compared viewing thin models with neutral images of some sort, including two which used images of older men and women, and infants and children playing [20,21], which does not give insight into what aspects of the models' bodies might be having this effect. Of those using women's bodies in the control group, one study used images which were chosen to vary according to attractiveness rather than weight [22], and one used television footage of 'thin and attractive' women in the experimental group and compared this with 'not thin or attractive' women in the control group [23], again leaving it uncertain whether the differences were due to factors relating to attractiveness or to weight. Two studies used clothing adverts or fashion models from women's magazines in the experimental group, but compared them with 'pictures from specialized women's clothing catalogues' for 'more amply furnished women' in one case [24], and 'overweight bordering on obese women' from other magazines in the other [25]. It seems likely that such images would vary in terms of the attractiveness and glamour of the images as well as their size and weight. One study attempted to look at the effects of body satisfaction of photographs which varied on weight alone, using

the face of one woman pasted onto either a relatively slim or an overweight body [26]. This study found lower levels of body satisfaction in the group shown a single overweight image. However, the study did not measure baseline levels of body satisfaction, so it is possible that these differences between groups were pre-existing. Interpreting these results is further complicated by the study scenario, in which the women were told that they were about to meet a male participant who would choose whether he would rather go on a date with them, or the other woman, whose photograph they were purportedly being shown. One further small study ($N = 40$, randomized into four groups) used a single black and white print advert for Kraft cheese, featuring an 'attractive female model' in two conditions, and a version using the model altered to appear overweight using Adobe Photoshop (no further details of the manipulation are reported) in a further two conditions; participants rated self-attractiveness after viewing the image for 15 s. Self-attractiveness ratings were lower following exposure to the 'overweight' models, but again no baseline measurements were taken [27].

One recent paper has examined whether giving feedback on participants' judgement as to the size of others' bodies can reduce clinical eating disorder symptom scores. Forty participants with high baseline body dissatisfaction were shown a series of CGI images of women's bodies and asked to rate each of them as 'fat' or 'thin'. They were then randomized to either complete a 'training' task in which they were given feedback regarding whether their response was correct or not, with the aim of shifting the threshold where they perceived a body as fat, or a control task where the feedback was consistent with their baseline judgements. The training task successfully shifted their perception of the cut-off between 'fat' and 'thin' bodies within the task, and resulted in decreased levels of eating disorder symptoms (shape, weight and eating concerns), which were sustained for two weeks [28]. Results were replicated in 20 patients with atypical anorexia nervosa, raising the possibility that such interventions may be a useful adjunct in eating disorder treatment.

To our knowledge, no studies to date have investigated the effect of shifting perception of body size on any behavioural outcome, such as food consumption. Whether food consumption increases or decreases might enhance our understanding of how body dissatisfaction can influence subsequent eating disorders or weight gain.

The studies presented here build on the previous literature to investigate whether the change in perception of body size of others extends to perception of and satisfaction with own body size, even without giving explicit 'training' instructions about which bodies are 'fat' or 'thin', and how long such effects last. We used a tightly controlled set of images, comprising photographs of the same women, whose BMI was in the middle of the normal range (22–23 $kg\,m^{-2}$) ('normal weight'), either unmodified, or modified to appear either 'overweight' or 'underweight' (figure 1). Participants were randomized to see one weight category of images. After seeking to confirm the previous findings that exposure to such images changes perception of the weight of others' bodies, we went on to test, in a sample six times larger than that previously used, the hypothesis that such exposure changes subsequent perception of the size of the participant's own body. We then tested whether such exposure to different sized bodies can change satisfaction with own body size. We also investigated whether changes in perception of or satisfaction with size resulted in behavioural differences in terms of amount of chocolate consumed. Study 2 aimed to replicate the results from Study 1 in a sample of participants with high baseline levels of body dissatisfaction, which may be considered to be more similar to an eating disorder patient group, and also investigated whether the effects would persist for 24 h. Given the higher prevalence of body dissatisfaction in women than men, we decided that these studies would focus on women only.

In both studies, we hypothesized that following exposure to images of 'underweight', as opposed to 'normal' or 'overweight' bodies, participants would (1) judge others' bodies as larger, (2) judge their own bodies as larger, and (3) be less satisfied with their body size. We further hypothesized that increase in satisfaction with body size would lead to (4) increased chocolate consumption following the task, possibly due to decreased worry about own weight. In Study 2, we further hypothesized (5) that changes would persist to the following day.

2. Methods: Study 1

2.1. Participants

We recruited female 18–25 year olds from the general population, by email invitation to a database of volunteers willing to be contacted about psychology research at the University of Bristol (consisting of University of Bristol students), and through advertisements on the University of Bristol Precinct.

Figure 1. Example of stimuli used (left: 'underweight'; middle: 'normal weight'; right: 'overweight').

Potential participants completed a brief electronic screening questionnaire giving their age, sex, fluency in English, height, weight, and whether they had a previous diagnosis of an eating disorder; they were not asked for details of their ethnicity or socio-economic status.

Participants were eligible to take part if they were female, English (or fluent in English), had no history of a diagnosed eating disorder, and had a body mass index (BMI) of 19–25 kg m^{-2} (normal range). Participants were offered £5 of 'Love to Shop' vouchers for their time.

2.2. Measures

Participants completed the Body Dissatisfaction (BD) subscale of the Eating Disorder Inventory (EDI) [29] online before attending the session. The Body Dissatisfaction subscale comprises nine items, and we scored them as suggested for a non-clinical population [30]: four from 1 (Never) to 6 (Always) (e.g. 'I think that my thighs are too large'), and five from 1 (Always) to 6 (Never) (e.g. 'I think that my hips are just the right size'). The Body Dissatisfaction subscale of the EDI has an internal reliability of 0.90 in a clinical, and 0.94 in a non-clinical, sample [30]. Participants completed the PHQ-9 (Patient Health Questionnaire) [31] during the session, before completing the task, as mood may impact on body dissatisfaction [32,33]. The PHQ-9 is a 9 item depression measure consisting of questions which score the DSM-IV depression criteria as '0' (not at all) to '3' (nearly every day). The scale has an internal reliability of 0.89. Before and after completing the task, participants estimated the size of four normal-weight stimuli on the computer screen, on a Visual Analogue Scale (VAS) from slightly underweight (1) to slightly overweight (7). Two of these stimuli were from the same set of images as the task stimuli (although not images which had been used in the tasks), and were dressed like the task stimuli in a sports bra and shorts, and two were fully clothed. Participants also completed VAS of own size, from 0 (too thin) to 10 (too fat), and satisfaction with own size, from 0 (not at all satisfied) to 10 (extremely satisfied), both before and after the task.

2.3. Stimuli

Stimuli used for the task were created from full body photographs of 10 different women aged between 18 and 25 years with a BMI of 22 to 23 kg m^{-2} and of White ethnicity, selected from a set of images of women with known BMI held in the School of Experimental Psychology, University of Bristol. These stimuli were used for the control group, 'normal weight'. The stimuli were modified by increasing the width of the images of women's bodies (initially 880 pixels) by 150 pixels to create an 'overweight' group, and decreasing the width of the women's bodies by 150 pixels to create an 'underweight' group. Images were cropped so the width of all photographs remained 880 pixels. Faces were blurred to preserve the

privacy of the women in the photographs, and to prevent subjects using faces to recognize the stimuli. An example of the stimuli used is given in figure 1 in black and white; full colour images were used in the task.

2.4. Procedure

Participants attended the University of Bristol, dressed in close-fitting clothing, as requested in the Participant Information Sheet and in the email inviting them. They provided informed consent, consenting to the tasks they would undertake, but remaining blind to the hypothesis and purpose of the study. They were tested individually, by one of two experimenters (H.B. and R.C.). They were randomized to one of the three groups by a computer-generated random sequence to ensure balance across groups. An author (K.S.B.) not involved in any testing or participant contact assigned letters to the three training conditions; experimenters and participants were blind to which training condition was which.

Participants completed the pre-task measures described above. They then completed a single session of the training. Following this, they were asked to remove outer layers of clothing (e.g. cardigans, jumpers, jackets, coats), look at themselves in a full length mirror, and then complete the post-task measures. The purpose of using the mirror was to ensure that the participant saw their own body in relation to the bodies to which they had been exposed. The experimenter then weighed them and measured their height without shoes. They were then asked to take a seat outside the room, the experimenter explaining 'I need to check the computer has registered the tasks correctly', and they were offered a bowl containing 300 g chocolate (Maltesers) while they waited. Maltesers were chosen as they are a palatable snack food that most people like, and because since they are quite small, people might choose to eat a variable quantity. After 2–3 min, they were invited back into the room, the remaining chocolate was weighed and they were debriefed.

The training session consisted of a two-back working memory task. The stimuli (photographs of the 10 women adjusted to be either 'over', 'under' or 'normal' (BMI 22–23 $kg\,m^{-2}$) weight) were presented sequentially (and repeatedly) to participants, who were instructed to respond as to whether the image was a 'target' (i.e. had the same identity as the individual presented two trials previously) or a 'non-target' (i.e. had a different identity from the individual presented two trials previously). This was to ensure that participants paid attention to the images displayed, which has been shown to enhance training effects [34]. In the 'underweight' condition, the 10 images modified to appear 'underweight' were displayed. In the overweight condition, the 10 images modified to appear 'overweight' were displayed. In the control condition, the 10 images of 'normal weight' women were displayed.

The tasks were programmed using E-Prime version 2.0 software (Psychology Software Tools Inc., Pittsburgh, PA, USA). Each image was displayed for 2000 ms, with a 20 ms inter-trial interval. To prevent low-level visual adaptation to the images, they were spatially jittered by an amount varying randomly from the centre of the screen by up to ±51 pixels in the x coordinate and ±38 pixels in the y coordinate (approximately 10% in each direction of the 1024 by 768 pixel display). Image size was 562 by 765 pixels, with a viewing distance of approximately 1 m. Failure to respond within the 2000 ms presentation resulted in no response being recorded for that trial. Each task block consisted of 48 trials, in which 8 'targets' were presented. A total of 7 blocks, giving a total of 336 trials, were presented, taking approximately 15 min for each participant.

2.5. Statistical analysis

Data were checked to ensure that no data were missing. The primary outcome was judgement of own body size (VAS from too thin (0) to too fat (10)) post-task. The secondary outcome was participant satisfaction with their body size after the intervention (VAS from not at all satisfied (0) to extremely satisfied (10)). Other outcome measures included perception of size of the women with BMI 22–23 $kg\,m^{-2}$ shown after the task, and weight of chocolate taken by participants after finishing the task.

For each outcome measure, we performed a linear regression analysis, treating randomization group as the categorical exposure variable and adjusting for the participants' baseline on each measure, by including in the regression analysis. All analyses were repeated adding adjustments for age, BMI and PHQ-9 score. The data that form the basis of the results presented here are available with restricted access from the data.bris Research Data Repository (https://data.bris.ac.uk/data) doi:10.5523/bris.rydgtj6qorrf1ndeh2e3c74g8, with author permission.

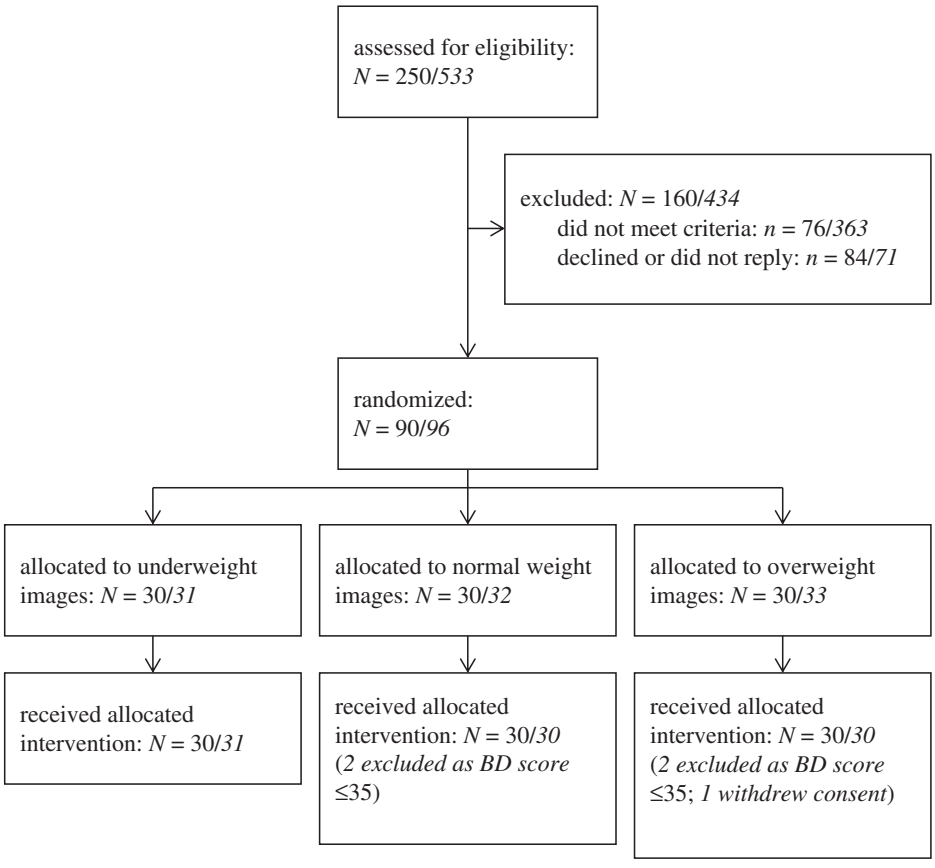


Figure 2. CONSORT diagram for Studies 1 and 2 *(Study 2 italicized)*.

Scores for individual measurements that were more than three standard deviations from the mean were judged to be outliers, and sensitivity analyses were repeated excluding such outliers.

Power calculation shows that three groups of 30 gave us 90% power at a 5% alpha level to detect an effect size of $d = 0.38$.

3. Results: Study 1

3.1. Descriptive data

We recruited 90 women between 11th November 2012 and 16th December 2013. Study recruitment is shown in the CONSORT diagram (figure 2) and descriptive data in table 1.

3.2. Perception of size of computer images (Hypothesis 1)

There was evidence of a difference between the means for the three groups in their judgement of the size of the computer images shown following exposure to different sized images, with those exposed to 'overweight' images during the task perceiving subsequent images as smaller than those exposed to normal or 'underweight' images ($p < 0.001$) (figure 3*a*). Adjustment for potential confounders (baseline PHQ-9, age, BMI) did not alter these results substantially. Full adjusted and unadjusted results for all measures are given in a supplementary online table (electronic supplementary material, table S2).

3.3. Perception of own body size (Hypothesis 2)

We did not find evidence to support a difference between the three groups in their mean judgement of their own size following exposure to different sized images ($p = 0.073$) (figure 4*a*). Adjusting for potential confounders did not alter these results substantially.

Table 1. Descriptive data on groups at baseline and post-training.

	Study 1				Study 2			
	all groups mean (s.d.)	adapted to 'underweight' mean (s.d.)	adapted to 'normal weight' mean (s.d.)	adapted to 'overweight' mean (s.d.)	all groups mean (s.d.)	adapted to 'underweight' mean (s.d.)	adapted to 'normal weight' mean (s.d.)	adapted to 'overweight' mean (s.d.)
age	20.4 (1.86)	20.6 (1.94)	20.1 (1.55)	20.3 (2.09)	21.4 (2.03)	21.5 (2.08)	21.2 (2.16)	21.5 (1.89)
body dissatisfaction score (0–54)	30.9 (7.2)	30.7 (5.44)	32.3 (7.71)	29.8 (8.20)	39.4 (3.89)	40.0 (4.94)	39.1 (3.06)	39.1 (3.43)
measured height (m)	1.66 (0.06)	1.66 (0.06)	1.66 (0.07)	1.65 (0.05)	1.66 (0.07)	1.65 (0.08)	1.68 (0.07)	1.65 (0.06)
measured weight (kg)	60.0 (6.52)	61.2 (4.99)	59.6 (7.72)	59.1 (6.60)	60.5 (7.04)	60.9 (8.26)	61.7 (7.77)	60.1 (6.47)
measured BMI (kg m^{-2})	21.8 (1.77)	22.3 (1.60)	21.5 (1.77)	21.7 (1.89)	22.1 (1.85)	22.3 (1.98)	21.9 (1.84)	22.1 (1.77)
PHQ-9 score	3.8 (3.05)	3.4 (2.39)	5.0 (3.60)	2.9 (2.73)	6.5 (4.17)	7.65 (4.83)	6.47 (4.31)	5.3 (2.90)
baseline rating of size of computer images (1–7)	4.78 (0.84)	4.80 (0.72)	4.86 (0.89)	4.68 (0.90)	5.09 (0.99)	4.94 (0.91)	5.19 (1.04)	5.14 (1.05)
post-training rating of size of computer images (1–7)	4.54 (0.93)	5.03 (0.84)	4.48 (0.82)	4.13 (0.92)	4.57 (0.91)	4.85 (0.77)	4.69 (0.64)	4.16 (1.12)
day 2 rating of size of computer images (1–7)	—	—	—	—	4.83 (0.77) ($N = 84$)	5.02 (0.75) ($N = 28$)	5.0 (0.63) ($N = 26$)	4.52 (0.82) ($N = 30$)
baseline rating of own size (0–10)	5.99 (1.01)	5.87 (0.96)	6.12 (0.97)	5.99 (1.13)	6.73 (1.01)	7.01 (0.96)	6.85 (1.02)	6.33 (0.96)
post-training rating of own size (0–10)	5.74 (1.07)	5.88 (1.09)	5.79 (0.97)	5.53 (1.15)	6.26 (1.20)	6.82 (1.03)	6.27 (1.22)	5.68 (1.09)
day 2 rating of own size (0–10)	—	—	—	—	6.58 (1.04) ($N = 84$)	6.89 (0.97) ($N = 28$)	6.73 (1.10) ($N = 26$)	6.15 (0.91) ($N = 30$)
baseline satisfaction with own size (0–10)	5.28 (2.06)	5.60 (2.06)	4.77 (1.64)	5.48 (2.37)	3.92 (1.68)	3.72 (1.87)	3.85 (1.62)	4.20 (1.56)
post-training satisfaction with own size (0–10)	5.79 (1.96)	5.77 (2.12)	5.65 (1.61)	5.96 (2.14)	4.45 (1.71)	3.81 (1.66)	4.32 (1.55)	5.25 (1.63)
day 2 satisfaction with own size	—	—	—	—	3.99 (1.58) ($N = 84$)	3.54 (1.53) ($N = 28$)	3.73 (1.50) ($N = 26$)	4.63 (1.54) ($N = 30$)
chocolate consumed post-training(g)	7.71 (6.16)	6.10 (5.6)	8.30 (6.10)	8.73 (6.64)	6.00 (7.57)	3.9 (4.85)	8.48 (9.15)	5.86 (7.82)

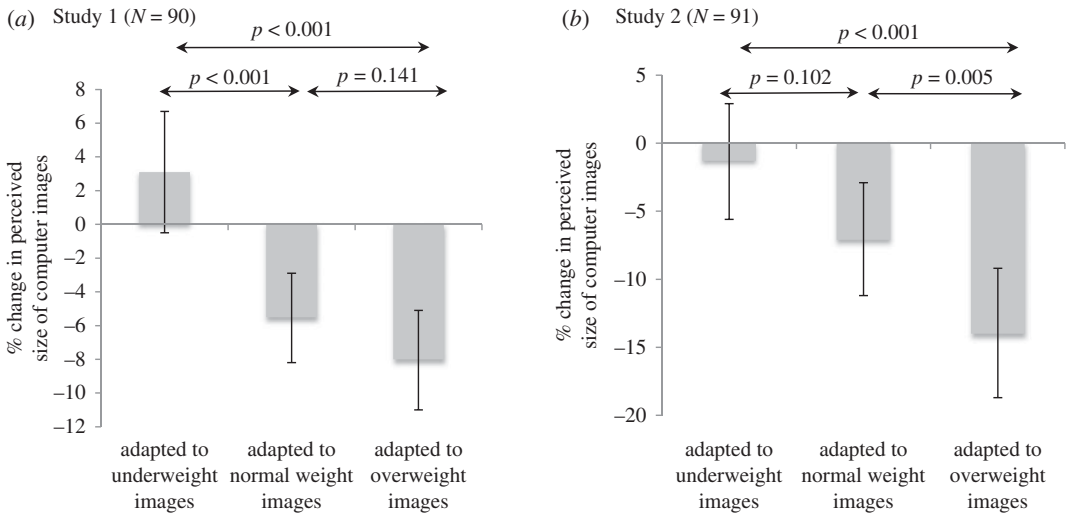


Figure 3. Mean percentage perceived change in size of computer images, pre- to post-task, by group (error bars represent 95% confidence intervals).

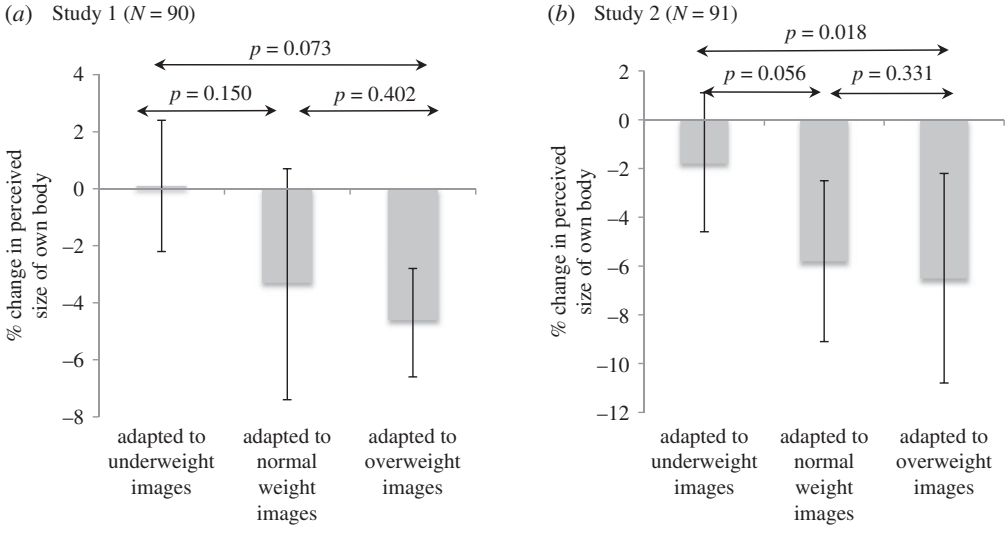


Figure 4. Mean percentage perceived change in own body size, pre- to post-task, by group (error bars represent 95% confidence intervals).

3.4. Satisfaction with own body size (Hypothesis 3)

There was evidence for group differences in mean satisfaction with own size following exposure to different sized images ($p = 0.046$), with those exposed to 'underweight' images being less satisfied than those exposed to 'normal' or 'overweight' images (figure 5*a*). Adjusting for potential confounders made no substantial difference to the results.

3.5. Chocolate consumption (Hypothesis 4)

The weight of chocolate consumed was not normally distributed, so a binary variable was created of eating 'no' versus 'any' chocolate. Treating the three experimental groups as a categorical variable, logistic regression found no differences in chocolate consumption between groups ($p = 0.203$) (figure 6*a*).

3.6. Sensitivity analysis excluding outliers

Removing outliers resulted in excluding four participants with scores for individual measurements that were more than three standard deviations from the mean: one participant who reported her own size at

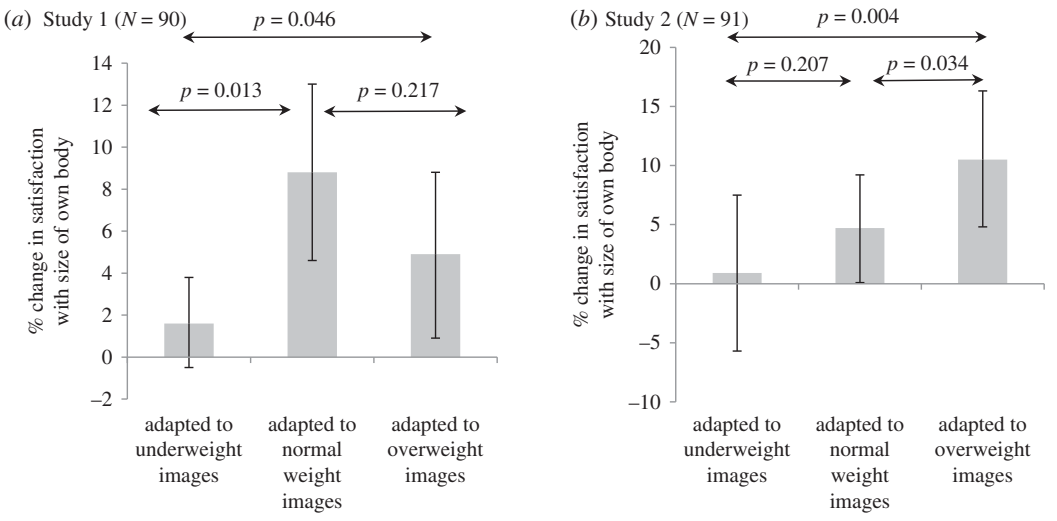


Figure 5. Mean percentage change in satisfaction with own body size, pre- to post-task by group (positive numbers indicate higher levels of satisfaction following task) (error bars represent 95% confidence intervals).

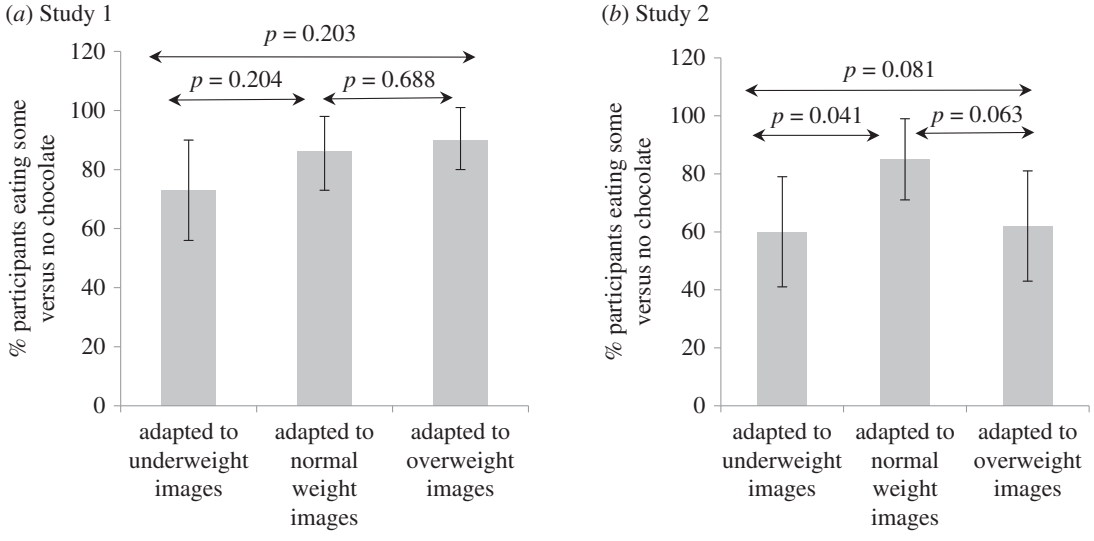


Figure 6. Percentage of participants eating some versus no chocolate following the task (error bars represent 95% confidence intervals).

baseline as 10 on the 10-point VAS scale, one who reported a change in her perception of her own size of 4 points on the 10-point VAS scale, and two who reported an increase in satisfaction of 3.5 and 3.85 respectively on the 10-point VAS scale. Analyses were repeated as described above, but excluding these outliers (electronic supplementary material, table S3).

There were no differences in results for judgement of the size of the computer images before and after the task. In judgement of own size, the statistical evidence for a difference between groups strengthened considerably ($p = 0.004$). There were no major differences in satisfaction with own size following removal of outliers. Adjusting for confounding variables did not substantially alter any results.

4. Methods: Study 2

Study 2 replicated Study 1 in a group with high baseline body dissatisfaction, in order to see whether results could be replicated in a group with more relevance to clinical populations. Differences in methodology were slight and are outlined below. The study protocol was preregistered on the Open Science Framework (https://osf.io/sngb6/).

4.1. Participants

Inclusion criteria required a baseline body dissatisfaction score of greater than 35 on the Body Dissatisfaction subscale of the Eating Disorder Inventory [29]; this value was chosen as it corresponded to the top quartile of scores in Study 1. Criteria were also widened to include those with a reported body mass index (BMI) of 18 to 25 $kg\,m^{-2}$ as we anticipated it being harder to recruit sufficient numbers with a high BD score, and in Study 1 we had to exclude many participants with a BMI between 18 and 19. Reimbursement was increased to £10, due to the requirement to attend follow-up the next day.

4.2. Measures and stimuli

Measures were identical to Study 1; participants completed the same measures at 24 h follow-up. Stimuli were identical to Study 1.

4.3. Procedure

The experiment was conducted by four authors (E.L., E.B., M.S. and H.A.). Day 1 study procedure was identical, except that participants were not debriefed. Participants attended 24 h follow-up. They looked at themselves in a full-length mirror, then completed the same measures as previously. Finally, they repeated the computer task using the 'normal' weight women, in order to reverse any increase in body dissatisfaction caused by exposure to 'underweight' women. They were then debriefed.

4.4. Statistical analysis

Outcomes and analysis were identical to Study 1, with additional analysis of results from 24 h follow-up, to investigate whether between-group differences in perception of computer images, and perception of and satisfaction with own size were present after 24 h. The data forming the basis of the results presented here are available open access from the data.bris Research Data Repository (https://data.bris.ac.uk/) doi:10.5523/bris.o1pqic3gep9n12c5mh0q0ojwj. A power calculation showed that three groups of 31 participants gives 90% power to detect between randomization group differences of 0.5 on a 10-point scale at a significance level of 0.05.

We also subsequently conducted a series of mega-analyses, in which we repeated all the analyses we have described, incorporating the results from both Study 1 and Study 2 into each analysis. The analyses were identical to those described for the individual studies, except that we included a variable to indicate whether each data point was from Study 1 or Study 2.

5. Results: Study 2

5.1. Descriptive data

We recruited 96 women between 29 July 2014 and 9 August 2015 (see figure 2 for CONSORT diagram). Four participants were excluded as they had been erroneously recruited with Body Dissatisfaction scores $\leq$35, despite inclusion criteria requiring a score greater than 35. One participant withdrew her consent following participation. Table 1 shows descriptive data.

5.2. Perception of size of computer images (Hypothesis 1)

There was evidence of a difference between the means for the three groups in their judgement of the size of the images shown following exposure to different sized images, consistent with the previous study, with those exposed to 'overweight' images during the task viewing subsequent images as smaller than those shown normal or 'underweight' images ($p < 0.001$) (figure 3*b*; electronic supplementary material, table S2). Adjusting for potential confounders (baseline PHQ-9, age, BMI) did not alter these results substantially (electronic supplementary material, table S2).

5.3. Perception of own body size (Hypothesis 2)

There was evidence of a difference between the means for the three groups in their judgement of their own size following exposure to different sized images, with those shown 'overweight' images subsequently judging their own body as smaller than those shown 'normal' or 'underweight'

images ($p = 0.018$) (figure 4*b* and electronic supplementary material, table S2). Adjusting for potential confounders did not alter these results substantially.

5.4. Satisfaction with own body size (Hypothesis 3)

There was evidence for group differences in mean satisfaction with own size following exposure to different sized images, with those shown overweight images subsequently being more satisfied with their own bodies than those shown 'underweight' or 'normal weight' images ($p = 0.004$) (figure 5*b* and electronic supplementary material, table S2). Adjusting for potential confounders made no substantial difference to the results.

5.5. Chocolate consumption (Hypothesis 4)

The weight of chocolate consumed was not normally distributed, so a binary variable was created of eating 'no' chocolate versus eating 'any' chocolate. Treating the three groups as a categorical variable, logistic regression analysis found weak evidence for between-group differences in chocolate consumption, with those in the 'underweight' and 'overweight' groups both eating less chocolate than those in the 'normal weight' group ($p = 0.081$) (figure 6*b*).

5.6. Twenty-four hour follow-up (Hypothesis 5)

There was some loss to follow-up: 84 of the 91 participants attended follow-up (92%), of whom 26 were in the group shown 'normal weight' images, 28 in the group shown 'underweight' images and 30 in the group shown 'overweight' images. Analysing using only data from those participants who attended on both days, we found that between-group differences in perceived size of computer images persisted at 24 h follow-up ($p < 0.001$), with differences in the same direction as in the post-task follow-up (electronic supplementary material, table S2). Results persisted following adjustment for confounding variables. There was no evidence that changes in perception of own size persisted to follow-up ($p = 0.311$). Differences in satisfaction with own size did persist ($p = 0.032$), although the strength of the evidence for this difference reduced following adjustment for confounders ($p = 0.080$). Repeating analyses including all participants, carrying forward the last available observation (i.e. those collected immediately after the task) in those who did not attend follow-up, did not substantially alter these results.

5.7. Sensitivity analysis excluding outliers

Removing outliers resulted in excluding five participants with scores for individual measurements that were more than three standard deviations from the mean (one scoring 10 and one scoring 3 for pre-test own size, two with change in size scores of -3.9 and $+2$ and one with a mean post-task score of the computer images of 1). Analyses were repeated excluding these outliers. There was no difference in results for judgement of the size of the computer images before and after the task. In judgement of own size, statistical evidence for a difference between groups strengthened considerably ($p = 0.005$) (electronic supplementary material, table S3); the difference in perception of own size following exposure to 'overweight' versus 'normal weight' images remained stable. Adjusting for confounding variables did not substantially alter any of these results.

5.8. Mega-analysis of Studies 1 and 2

Combining the results from Studies 1 and 2 strengthened the findings throughout, and we found strong statistical evidence in support of hypotheses 1 to 3 (electronic supplementary material, table S2). Adjusting for confounders, and repeating the analyses excluding outliers (as defined separately for Study 1 and Study 2) did not change these findings.

6. Discussion

The aims of the study were to test in a sample six times larger than that previously used whether exposure to bodies of different sizes would extend beyond changing perception of others' bodies to also change perception of own body size. We found that this was the case in Study 2 (participants with high baseline levels of body dissatisfaction), and in mega-analyses of Studies 1 and 2. Furthermore, we tested the

hypothesis that these changes in perception of own size would also change satisfaction with own size, and this was demonstrated and replicated across the two studies. Finally, we found that changes in satisfaction with own body size lasted until 24 h follow-up.

We confirmed findings from previous studies [13,14,17,18,28] that exposing women to images of women's bodies of different sizes alters their perception of the size of subsequent bodies, with exposure to larger bodies leading to subsequent perception of other bodies as smaller. We confirmed that this is found in those with, and without, high baseline levels of body dissatisfaction. In Study 2, and in the mega-analysis, we found good evidence that exposing participants to images of 'overweight' as opposed to 'underweight' or 'normal weight' bodies leads to them perceiving their own size as smaller. We also showed in both studies that satisfaction with own size changes following exposure to images of women's bodies of different sizes, with exposure to larger bodies resulting in increased satisfaction. This finding was stronger in the group of participants who were dissatisfied with their bodies at baseline. Differences in satisfaction with own body size, and perception of body size of others persisted to 24 h follow-up, although differences in perception of own size did not persist. We did not find any overall effect of exposure to images of different sizes on chocolate consumption, which may be due to individual variation in whether a change in satisfaction with own size leads to the consumption of more or less chocolate.

Studying the differences between groups more closely, changes in perception of own and others' size appears to be predominantly driven by decreases in perception of own size following exposure to normal or overweight images. Exposure to 'underweight' images had less effect on subsequent perception of own and others' size. This may be because our 'underweight' images are in fact very similar in size to the media images surrounding us, and may be larger than average media images. For example, a study of Internet models found that over 50% had a BMI of $\leq$18 [35]. It is likely that body size norms [36] are strongly affected by previous exposure to media images, so further exposure to 'underweight' images may make little difference.

The current studies build on previous research in demonstrating in a larger sample than that previously used, that exposure to bodies of different sizes can alter perception of and satisfaction with own body, and this occurs in all women and to a greater extent in those with high baseline body dissatisfaction. Experimenters and participants were blind to which training study was which, and the study was tightly controlled, in that all groups saw images of the same women, manipulated to appear in different sizes. The second study replicated the methods of the first, and was preregistered with defined hypotheses and primary outcome measures.

However, there were some limitations to the study. First, participants were almost all students at the University of Bristol, and we do not have data on their ethnicity or socio-economic status; they therefore may not be representative of the national population of 18–25 year olds. Second, the images created were not very sophisticated, merely varying the overall width of the bodies rather than altering their proportions in any other way; we did not pilot the images to see whether they were viewed as realistic, and they were all of women of White ethnicity. However, the images are more sophisticated than those used in some previous studies in that they are photographs of real women, rather than 2D or 3D sketches. Third, our 'normal' images comprised images of women with a known BMI of 22 to 23 $kg\,m^{-2}$, but we do not know the BMI of the groups we created by altering the widths of the images. The mean BMI of the participants (21.8 $kg\,m^{-2}$ in Study 1; 22.1 $kg\,m^{-2}$ in Study 2) was at the lower end of the range of the BMI of the 'normal' weight group (22 to 23 $kg\,m^{-2}$) whose images we used. Fourth, at baseline, participants rated the images with a BMI of 22 to 23 $kg\,m^{-2}$ as being towards the overweight end of the scale (mean of 4.78 on a 7-point scale (Study 1); 5.09 (Study 2)), despite this BMI being in the middle of the healthy range of BMI. Therefore, those completing the 'normal' weight version of the task may have been exposed to images they considered to be slightly overweight. This does appear to be the case: exposure to 'normal' weight images led to viewing subsequent computer images as smaller, perceiving own size as smaller and being more satisfied with own size. Fifth, the wording of the task in which participants judged their own size included the phrases 'too fat' and 'too thin', which may have drawn on their feelings of satisfaction with their body as well as their perception of their actual size. Sixth, participants may have guessed the purpose of the study, and results may have been affected by demand characteristics; we did not check participant awareness of the study hypothesis. Seventh, the follow-up time was only 24 h, and we therefore do not know whether effects may last longer than this. Eighth, the choice of chocolate as a snack may not have been ideal, both because it is a food that some people crave, and because the craving often involves both approach and avoidance, as well as feelings of guilt [37].

Our findings are potentially important in view of the current media environment, where women are constantly exposed to images of underweight women. We have demonstrated here that relatively short

exposure to images of normal women's bodies of different sizes can change perception of the size of both own and others' bodies, and satisfaction with own size, with particularly marked effect in those who feel more negative about their bodies at baseline. It is possible that the change in satisfaction with own size may be due to exposure to different-sized bodies changing the internalized 'ideal' body size to which women compare their own bodies, as well as to a change in perceived own size. In either case, the results of our study suggest that encouraging media outlets to choose models of a normal body size would be likely to lead to women perceiving their own size as smaller and feeling more positive about their own body size: this may in turn reduce levels of body dissatisfaction, a known risk factor for eating disorders and obesity. Some evidence suggests that choosing larger models would not diminish the effectiveness of advertising [38].

The effect of presenting images of larger women in changing perception of own body size, and increasing body size satisfaction, particularly in highly body-dissatisfied women, also suggests that such an intervention may have a role to play in the prevention of eating disorders and obesity. Intervening to reduce high levels of body dissatisfaction may be a useful adjunct in, for example, eating disorder prevention programmes.

Future research could usefully investigate whether such an intervention is effective in patients with a diagnosed eating disorder. In such patients, perceiving own body size as larger than it is in reality is a core part of the disorder [39], which is resistant to current treatments [40] and is predictive of relapse [1].

7. Conclusion

In conclusion, our results put forward a strong public health message that increasing the number of normal and larger women in the media may reduce levels of body dissatisfaction in the population, and thus would have the potential to reduce rates of weight gain, obesity and eating disorders.

Ethics. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. Ethics approval was obtained from the Faculty of Science Research Ethics Committee at the University of Bristol (reference nos. 270912588, 2706148085). Participants gave written consent to participate. The study was conducted according to Good Clinical Practice guidelines.

Data accessibility. The data forming the basis of the Study 1 results presented here are available from the data.bris Research Data Repository (https://data.bris.ac.uk/data) doi:10.5523/bris.rydgtj6qorrf1ndeh2e3c74g8, with author permission; results from Study 2 are available open access from the data.bris Research Data Repository (https://data.bris.ac.uk/) doi:10.5523/bris.o1pqic3gep9n12c5mh0q0ojwjp.

Authors' contributions. H.B., K.S.B., G.L., I.S.P.-V. and M.R.M. conceived of and designed Study 1; H.B., G.L., I.S.P.-V., M.R.M., A.S., M.R.B., C.J.H. and R.P. conceived of and designed Study 2; H.B. conducted the statistical analysis; H.B. wrote the first draft of the manuscript; R.C., K.S.B., G.L., I.S.P.-V., M.R.M., A.S., M.R.B., C.J.H. and R.P. helped with subsequent drafts. All authors gave final approval for publication.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the Elizabeth Blackwell Institute for Health Research, University of Bristol and the Wellcome Trust Institutional Strategic Support Fund (Clinical Primer Award to H.B.), April–Sept 2014; and the Wellcome Doctoral Training Fellowship (H.B.) at the University of Oxford, Oct 2014–present. M.R.M. is a member of the UK Centre for Tobacco Control Studies, a UKCRC Public Health Research: Centre of Excellence. Funding from British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council and the National Institute for Health Research, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. This work was supported by the Medical Research Council Integrative Epidemiology Unit at the University of Bristol (MC_UU_12013/6).

Acknowledgements. We would like to thank the study participants for their time.

References

1. Jacobi C, Hayward C, de Zwaan M, Kraemer HC, Agras WS. 2004 Coming to terms with risk factors for eating disorders: application of risk terminology and suggestions for a general taxonomy. *Psychol. Bull.* **130**, 19–65. (doi:10.1037/0033-2909.130.1.19)
2. Haines J, Neumark-Sztainer D, Wall M, Story M. 2007 Personal, behavioral, and environmental risk and protective factors for adolescent overweight. *Obesity* **15**, 2748–2760. (doi:10.1038/oby.2007.327)
3. Sonneville KR, Calzo JP, Horton NJ, Haines J, Austin SB, Field AE. 2012 Body satisfaction, weight gain and binge eating among overweight adolescent girls. *Int. J. Obesity* **36**, 944–949. (doi:10.1038/ijo.2012.68)
4. Quick V, Wall M, Larson N, Haines J, Neumark-Sztainer D. 2013 Personal, behavioral and socio-environmental predictors of overweight incidence in young adults: 10-yr longitudinal findings. *Int. J. Behav. Nutr. Phy.* **10**, 37. (doi:10.1186/1479-5868-10-37)
5. van den Berg P, Neumark-Sztainer D. 2007 Fat 'n happy 5 years later: is it bad for overweight girls to like their bodies? *J. Adolesc. Health* **41**, 415–417. (doi:10.1016/j.jadohealth.2007.06.001)
6. Bearman SK, Martinez E, Stice E, Presnell K. 2006 The skinny on body dissatisfaction: a longitudinal study of adolescent girls and boys. *J. Youth Adolesc.*

35, 217–229. (doi:10.1007/s10964-005-9010-9)

7. Frederick DA, Peplau LA, Lever J. 2006 The swimsuit issue: correlates of body image in a sample of 52,677 heterosexual adults. *Body Image* **3**, 413–419. (doi:10.1016/j.bodyim.2006.08.002)
8. Ricciardelli LA, McCabe MP. 2001 Dietary restraint and negative affect as mediators of body dissatisfaction and bulimic behavior in adolescent girls and boys. *Behav. Res. Ther.* **39**, 1317–1328. (doi:10.1016/S0005-7967(00)00097-8)
9. Swami V, Stieger S, Harris AS, Nader IW, Pietschnig J, Voracek M, Tovee MJ. 2012 Further investigation of the validity and reliability of the photographic figure rating scale for body image assessment. *J. Pers. Assess.* **94**, 404–409. (doi:10.1080/00223891.2012.660293)
10. Muennig P, Jia H, Lee R, Lubetkin E. 2008 I think therefore I am: perceived ideal weight as a determinant of health. *Am. J. Public Health* **98**, 501–506. (doi:10.2105/AJPH.2007.114769)
11. Stice E, Schupak-Neuberg E, Shaw HE, Stein RI. 1994 Relation of media exposure to eating disorder symptomatology: an examination of mediating mechanisms. *J. Abnorm. Psychol.* **103**, 836–840. (doi:10.1037/0021-843X.103.4.836)
12. Thompson JKH, Leslie J, Madeline A, Stacey T-D. 1999 *Exacting beauty: theory, assessment, and treatment of body image disturbance*. Washington, DC: American Psychological Association.
13. Glauert R, Rhodes G, Byrne S, Fink B, Grammer K. 2009 Body dissatisfaction and the effects of perceptual exposure on body norms and ideals. *Int. J. Eat. Disord.* **42**, 443–452. (doi:10.1002/eat.20640)
14. Winkler C, Rhodes G. 2005 Perceptual adaptation affects attractiveness of female bodies. *Br. J. Psychol.* **96**, 141–154. (doi:10.1348/000712605X36343)
15. Robinson E, Christiansen P. 2014 The changing face of obesity: exposure to and acceptance of obesity. *Obesity* **22**, 1380–1386. (doi:10.1002/oby.20699)
16. Robinson E, Kirkham TC. 2014 Is he a healthy weight? Exposure to obesity changes perception of the weight status of others. *Int. J. Obesity* **38**, 663–667. (doi:10.1038/ijo.2013.154)
17. Hummel D *et al.* 2013 Neural adaptation to thin and fat bodies in the fusiform body area and middle occipital gyrus: an fMRI adaptation study. *Hum. Brain Mapp.* **34**, 3233–3246. (doi:10.1002/hbm.22135)
18. Hummel D, Rudolf AK, Untch KH, Grabhorn R, Mohr HM. 2012 Visual adaptation to thin and fat bodies transfers across identity. *PLoS ONE* **7**, e43195. (doi:10.1371/journal.pone.0043195)
19. Grabe S, Ward LM, Hyde JS. 2008 Role of the media in body image concerns among women: a meta-analysis of experimental and correlational studies. *Psychol. Bull.* **134**, 460–476. (doi:10.1037/0033-2909.134.3.460)
20. Wegner BS, Hartmann AM, Geist CR. 2000 Effect of exposure to photographs of thin models on self-consciousness in female college students. *Psychol. Rep.* **86**, 1149–1154. (doi:10.2466/pr0.2000.86.3c.1149)
21. Kalodner CR. 1997 Media influences on male and female non-eating-disordered college students: a significant issue. *Eat. Disorder. J. Treat. Prevent.* **5**, 47–57. (doi:10.1080/10640269708249203)
22. Cash TF, Cash DW, Butters JW. 1983 Mirror, mirror, on the wall . . . contrast effects and self-evaluations of physical attractiveness. *Pers. Soc. Psychol. B.* **9**, 351–358. (doi:10.1177/0146167283093004)
23. Cattarin JA, Thompson JK, Thomas C, Williams R. 2000 Body image, mood, and televised images of attractiveness: the role of social comparison. *J. Soc. Clin. Psychol.* **19**, 220–239. (doi:10.1521/jscp.2000.19.2.220)
24. Wilcox K, Laird JD. 2000 The impact of media images of super-slender women on women's self-esteem: Identification, social comparison, and self-perception. *J. Res. Pers.* **34**, 278–286. (doi:10.1006/jrpe.1999.2281)
25. Ogden J, Mundray K. 1996 The effect of the media on body satisfaction: the role of gender and size. *Eur. Eat Disord. Rev.* **4**, 171–182. (doi:10.1002/(SICI)1099-0968(199609)4:3<171::AID-ERV132>3.0.CO;2-U)
26. Lin LF, Kulik JA. 2002 Social comparison and women's body satisfaction. *Basic. Appl. Soc. Psych.* **24**, 115–123. (doi:10.1207/S15324834BASP2402_4)
27. Crouch A, Degelman D. 1998 Influence of female body images in printed advertising on self-ratings of physical attractiveness by adolescent girls. *Percept. Motor Skill.* **87**, 585–586. (doi:10.2466/pms.1998.87.2.585)
28. Gledhill LJ, Cornelissen KK, Cornelissen PL, Penton-Voak IS, Munafo MR, Tovee MJ. 2017 An interactive training programme to treat body image disturbance. *Br. J. Health Psychol.* **22**, 60–76. (doi:10.1111/bjhp.12217)
29. Garner DM. 1991 *Eating disorder inventory-2: professional manual*. Odessa, FL: Psychological Assessment Resources.
30. Schoemaker C, van Strien T, van der Staak C. 1994 Validation of the eating disorders inventory in a nonclinical population using transformed and untransformed responses. *Int. J. Eat. Disord.* **15**, 387–393. (doi:10.1002/eat.2260150409)
31. Kroenke K, Spitzer RL, Williams JB. 2001 The PHQ-9: validity of a brief depression severity measure. *J. Gen. Intern. Med.* **16**, 606–613. (doi:10.1046/j.1525-1497.2001.016009606.x)
32. Baker JD, Williamson DA, Sylve C. 1995 Body-image disturbance, memory bias, and body dysphoria: effects of negative mood induction. *Behav. Ther.* **26**, 747–759. (doi:10.1016/S0005-7894(05)80043-7)
33. Taylor MJ, Cooper PJ. 1992 An experimental-study of the effect of mood on body size perception. *Behav. Res. Ther.* **30**, 53–58. (doi:10.1016/0005-7967(92)90096-Y)
34. Rhodes G, Jeffery L, Evangelista E, Ewing L, Peters M, Taylor L. 2011 Enhanced attention amplifies face adaptation. *Vision Res.* **51**, 1811–1819. (doi:10.1016/j.visres.2011.06.008)
35. Owen PR, Laurel-Seller E. 2000 Weight and shape ideals: thin is dangerously in. *J. Appl. Soc. Psychol.* **30**, 979–990. (doi:10.1111/j.1559-1816.2000.tb02506.x)
36. Rhodes G, Jeffery L, Boeing A, Calder AJ. 2013 Visual coding of human bodies: perceptual aftereffects reveal norm-based, opponent coding of body identity. *J. Exp. Psychol. Human Percept. Perform.* **39**, 313–317. (doi:10.1037/a0031568)
37. Cartwright F, Stritzke WG. 2008 A multidimensional ambivalence model of chocolate craving: construct validity and associations with chocolate consumption and disordered eating. *Eat. Behav.* **9**, 1–12. (doi:10.1016/j.eatbeh.2007.01.006)
38. Halliwell E, Dittmar H. 2004 Does size matter? The impact of model's body size on women's body-focused anxiety and advertising effectiveness. *J. Soc. Clin. Psychol.* **23**, 104–122. (doi:10.1521/jscp.23.1.104.26989)
39. Cash TF, Deagle 3rd EA. 1997 The nature and extent of body-image disturbances in anorexia nervosa and bulimia nervosa: a meta-analysis. *Int. J. Eat. Disord.* **22**, 107–125. (doi:10.1002/(SICI)1098-108X(199709)22:2<107::AID-EAT1>3.0.CO;2-J)
40. Alleva JM, Sheeran P, Webb TL, Martijn C, Miles E. 2015 A meta-analytic review of stand-alone interventions to improve body image. *PLoS ONE* **10**, e0139177. (doi:10.1371/journal.pone.0139177)

The Morphed Photographic Figure Scale:

Creation and validation of a novel set of realistic female body stimuli

Andy Skinner[1,2]*, Helen Bould[3]*, Andy Woods[4], Peter Etchells[5], Ian Penton-Voak[2], Andrew P Clark[6], Catherine J Harmer[3], Rebecca Park[3], Matthew Broome[3,7,8], Marcus Munafò[1,2]

*Joint first authors

[1]MRC Integrative Epidemiology Unit at the University of Bristol,

[2]School of Experimental Psychology, University of Bristol

[3]Department of Psychiatry, University of Oxford

[4]Crossmodal Research Laboratory, University of Oxford

[5]College of Liberal Arts, Bath Spa University

[6]Department of Life Sciences, Brunel University

[7]Warneford Hospital, Oxford Health NHS Foundation Trust, Oxford, UK

[8]Division of Mental Health and Wellbeing, Warwick Medical School, University of Warwick,

Corresponding Author: Andrew Skinner, School of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, United Kingdom. T: +44.117.9288581; F: +44.117.9288588; E: andy.skinner@bristol.ac.uk

Abstract

Research exploring how people with eating disorders or obesity perceive their bodies, and interventions targeting this rely on high quality figural rating scales. We present data on a new figural scale, the Morphed Photographic Figure Scale (MPFS). It comprises eight sequences of photographic quality images of female bodies varying from underweight to overweight. Each sequence is centred on a veridical photographic image of a female body, and we used image morphing techniques to produce a number of increasingly overweight and increasingly underweight versions of the same body. Morphing provides body images with highly realistic representations of changes in adiposity and muscle tone, and enables manipulation of body weight without introducing idiosyncratic variations in anthropomorphic features, which could impact judgments of physical appearance in unpredictable ways. All images were rated in an online study (n = 377) as underweight, overweight or neither. Within each sequence, rating data showed a clear rank ordering from underweight to overweight. Overweight participants were more likely to rate images as underweight, and less likely to rate them as overweight. Participants with a high level of concern about their own weight (measured using the Eating Disorders Examination Questionnaire) were more likely to rate images as overweight. We anticipate the MPFS will have application in a range of research domains, including assessment of body image disturbances, and the design of new interventions targeting body image.

Introduction

Figural rating scales have been used for some time to understand how people perceive their body size. This is a key aspect of research exploring eating disorders, obesity, body dissatisfaction and body image disturbance. Figural rating scales consist of sequences of images of bodies that range from underweight to overweight in a number of steps. A variety of such scales have been designed, using either line drawings or photographs. Each has particular benefits and limitations. Here, we present a new figural image set, the Morphed Photographic Figure Scale (MPFS), which combines the benefits of photographic quality images with advances in computer image manipulation to address a number of the limitations in previous scales.

Early image scales used contour line drawings (or 'silhouettes'), produced by artists. Despite widespread use of scales like the Body Rating Scale [1], they have a number of limitations. They do not provide realistic representations of the human form, often having poorly defined features and disproportionate limbs [2]. Perhaps more importantly, they fail to accurately represent true, localised variations in fat deposits, muscle tone and skin texture that occur when individuals lose or gain weight [3].

Photographic image scales, such as the Photographic Figure Rating Scale (PFRS; [4]), address these criticisms by using photographic images to provide realistic representations of bodies. In the PFRS, each step in the sequence is an image of a different person, with individuals selected by BMI to give a sequence progressing from underweight to overweight. These images provide more realistic representations of body shapes, but the use of a range of different individuals to form the sequence

introduces possible confounds. Idiosyncrasies in a range of anthropomorphic features means there are variations in factors such as leg-to-body ratio, which could, for example, affect judgments of physical attractiveness [5].

Ideally, a figural rating scale would maintain the realistic representation of body shape provided by the PFRS, but would also control for individual factors like variation body proportions that are relatively weight independent (e.g., leg-to-body ratio), and would expose key areas of the body where changes in weight are seen, such as the stomach. The Morphed Photographic Figure Scale (MPFS) does exactly this. A set of high quality photographic images of real women's bodies, dressed in a standardised way in sports bra and shorts, were morphed using techniques originally designed to produce high quality manipulations of images of faces [6-9]. The nature of the transformation used in the morphing procedure, which is made by averaging the body images of many individuals, enables realistic changes in both shape and the texture, without introducing unwanted changes to anthropomorphic features, which may influence judgment of appearance or weight. The combination of the outfits worn, and the high quality image manipulations, provides realistic representations of changes in fat deposits and muscle tone in important parts of the body not seen in other scales (e.g., the stomach). In line with previous recommendations [2, 10], we heavily blurred the facial features of the images, helping to ensure attention is not drawn away from the body.

Recently, figural rating scales have been used in studies that move beyond simply measuring perception of body size, to designing interventions targeted at treating individuals with body dissatisfaction [11]. We anticipate the MPFS will also be of use in these areas.

An online rating study was used to assess the validity of the MPFS, in which participants rated each image as 'over-weight', 'under-weight', or 'neither over or under weight'. To demonstrate the utility of the MPFS, we also used it to test for between group differences in judgments about weight. We hypothesised that male subjects, and those who were older, had higher body mass index (BMI), were more satisfied with their own weight, and did not have a history of an eating disorder would be more likely to perceive photographs of women as being under or normal weight, and that conversely women, and those with lower BMI, who were less satisfied with their own weight, and who had a history of an eating disorder would be more likely to perceive the photographs as being overweight.

Methods

Photographic Body Images

Models (mean age = 22 years old, SD = 3.0) were recruited from the undergraduate population at University of Bristol and provided with plain black underclothing. The photographic stimuli were single frame images captured using a Panasonic HDC-HS700 high-definition camera mounted on a tripod at a height of approximately 150 cm from the floor, and 250 cm from the model. The camera shutter was set to 1/50, image resolution set to 1920 × 1080, and luminance 55%. Lighting was natural daylight.

Body Mass Index (BMI) was calculated using height measurements from a stadiometer, and weight measurements from a Tanita BC-418 body composition analyser.

Body Image Morphing

The morphing process used to increase and decrease body size was performed using PsychoMorph [12]. PsychoMorph enables independent manipulation of the shape, texture and colour of images. Images to be manipulated are marked up with a template comprising 312 fiducial points delineating important elements of the image. For bodies, these are the perimeter outlines of the torso, arms and legs, the outlines of the clothing, and positions of certain body features, such as the navel. A body marked up with a template is shown in Figure 1 (left panel).

To increase and decrease body size, PsychoMorph uses a transform derived from two reference body images – one underweight, the other overweight. The differences in shape, texture and colour between these two reference bodies forms the transform, and this transform can be applied in varying amounts to a real body image to either increase or decrease its size.

To keep levels of idiosyncratic variation within our sequences to a minimum, we used averaged bodies rather than individual body images as the references that formed our transform. The averaging process preserves the shared characteristics of the body images (e.g. muscle tone) from which they were produced, but removes individual anthropometric variations. To make the average for the underweight body we first rank-ordered the complete set of body images available to us by BMI. We took the 10 bodies with the lowest BMIs (m=17.3), marked up each one with a template in PsychoMorph, and applied the averaging function in PsychoMorph (see Supplementary Details for PsychoMorph settings). We repeated this with the bodies with the 10 highest BMIs (m=25.2) to make an overweight body average. The average underweight and overweight bodies are shown in Figure 1 (right panel).

Post Processing

To remove unwanted elements from the images the 'mask' function in PsychoMorph was used to set the background in the lower part of the images to black. Image artifacts introduced by the morphing procedure (e.g. noise around edges) were removed using GIMP. Finally, images were cropped to 500 x 1400 pixels.

Stimuli Set

The complete Morphed Photographic Figure Scale set comprises eight sequences, each based on one veridical image of an individual. For each sequences we applied +20%, +40%, +60% and +80% of the PsychoMorph underweight <-> overweight transform to make four larger versions of the individual, and -20%, -40%, -60% and -80% of the transform to make four slimmer versions, giving a total of nine images per sequence.

Four of the sequences are based on models with a Body Mass Index (BMI) in the range 16-19 kg/m^2, the other four are based on models with BMIs 20-23 kg/m^2. Individual models in these ranges were selected on the basis of having natural poses, free from arrangements of limbs that are known to be problematic for the morphing procedure (e.g. hands occluding thighs). An example of one MPFS sequence is shown in Figure 2 (top panel). The complete set of eight sequences is shown in Supplementary Materials.

Online Rating of Stimuli

Participants were recruited via email and poster advertisements to University of Oxford students and staff. They then completed an online task (implemented using the Xperiment online experimental platform, http://www.xperiment.mobi), in which they were shown the 72 body images created using the morphing technique described above. Images were presented in a randomised order, and the accompanying text asked: "Is this woman over-weight, under-weight or neither over- nor under-weight?". Participants selected a weight category by clicking on one of the category words. Participants completing the study were offered the opportunity to be entered into a prize draw to win an iPad Air. The study received ethics approval from the Central University Research Ethics Committee at the University of Oxford (MSD-IDREC-C1-2015-020).

We calculated the percentage of images that each participant coded as "underweight", "neither under nor over weight" and "overweight". We then compared whether the percentage of images viewed as underweight, and percentage of images viewed as overweight differed by gender, age, previous or current eating disorder, BMI, time of

last meal, ethnicity, and three body satisfaction measures (EDEQ Weight Concern, EDEQ Shape Concern, Body Satisfaction Score), initially in a descriptive table and then by regression analysis. The analyses were repeated, adjusting for all variables except ethnicity.

Results

The online task was completed by 337 participants (73% female). Data on descriptive variables were complete for 336 people; however, two people had improbably high scores for BMI (>100 kg/m^2) and they were therefore also excluded (see Figure 3, CONSORT diagram). The mean age was 23 years old (SD 5.9) and mean BMI was 22.4 kg/m^2 (SD 3.4). Further details of demographics are given in Table 1.

Not all participants gave an answer regarding the weight status of all images, but they were not excluded from the study as we were calculating percentage of images that were coded as under-, over- or neither under nor overweight. Mean and standard deviation scores for the percentages of images coded as being under-, over- or neither under- nor overweight are given in Table 1, both for the whole sample and by gender, age, ethnicity, BMI band, history of an eating disorder, and three measures of body satisfaction (EDEQ Weight Concern, EDEQ Shape Concern, Body Satisfaction Score).

Examples of the ratings obtained for an individual image are shown in Figure 2 (bottom panel). Ratings for all sequences are included in Supplementary Materials. In addition to showing how each individual image in the eight sequences was rated, this rating data also illustrates how ratings changed across sequences. In all eight sequences, there was a clear rank order progression, with monotonic decreases in 'underweight' ratings and monotonic increases in 'overweight' ratings when moving from underweight to overweight images.

Full results of regression analyses comparing the different categories are given in Table 2. Those with a higher BMI were more likely to perceive women as being underweight,

and less likely to perceive them as being overweight. This difference was robust to adjustment for other variables, and indicated that 0.3% more images are viewed as underweight with each 1 point increase in observer BMI (95% CI 0.02, 0.57, p=0.03), and 0.5% fewer images are perceived as overweight with each 1 point increase in observer BMI (95% CI -0.93, -0.11, p=0.01).

Higher EDEQ Weight Concern scores were strongly associated with rating a higher percentage of women as overweight, which again was robust to adjustment: 7.33% more images are viewed as overweight by those with an EDEQ Weight Concern score of >4 than those with a score of 0-4 (95% CI 3.59, 11.06), p<0.0001 (4 is the cut-off for clinical significance on each EDEQ subscale).

There was no evidence for any differences in the percentage of images coded as under or overweight by gender, age, presence or absence of current/previous eating disorder, time since last meal, EDEQ Shape Concern or Body Satisfaction score.

Discussion

The Morphed Photographic Figure Scale (MPFS) is a novel set of figural images produced by the application of computer image manipulation techniques to photographic images of real bodies. The scale is unique, in that its photorealistic images provides accurate representations of the changes in fat deposits, muscle tone and overall appearance that result from weight gain or loss (in body areas often covered in other scales), but are free from idiosyncratic variations in anthropomorphic features that may affect judgments of physical attractiveness.

We provide validation data, indicating how subjects rated each image as either underweight, overweight or neither. Other approaches to rating and validating figural stimuli have been proposed for use in specific applications (e.g. [13]). Here, we opted for a less specific approach in order to enable a broad range of researchers to select stimuli for use in a variety of applications. The rating data show a clear rank ordering from underweight to overweight within each of the eight sequences in the MPFS, indicating good construct validity.

To demonstrate the utility of the MPFS, we have used it to explore between group differences in perception of weight status. Participants with a higher BMI were found to be more likely to judge images as underweight, and less likely to judge them as overweight; those with weight concern scores in the clinical range were more likely to judge images as being overweight. These results are in line with previous studies, in that it has been found previously that observers with lower own BMI are more likely to overestimate the BMI of images of others [14]. This study also found that own BMI was a more important factor than having a diagnosed eating disorder, and, in contrast to our

results, found that satisfaction with own body did not have an effect on estimation of others' body size. Some previous work has found no evidence of gender differences in ideal weight for women [15], which is consistent with our findings that weight categories are not perceived differently by men and women; this is in contrast to earlier studies using numerical descriptions of weight, which did find gender differences in what would be considered underweight [16].

The main limitation of the MPFS is that it currently comprises only female body image scales. One reason for this is that the processes used to produce these very high quality stimuli are extremely resource intensive, and this limited the number of stimuli that could be produced. Specifically, the quality of the averaging and morphing image manipulations are highly dependent on precise placement of the hundreds of fiducial markers on each of the many body images used in the process. Placement can be automated to a degree, but still requires considerable further input and refinement from skilled operators. This meant it was necessary to focus available resources on production of a limited set of stimuli. We recognise that issues of body dissatisfaction and dysmorphia relate to both males and females, but the rate is acknowledged to be higher in females [17,18], so we therefore focused initially on completion of a female version of the scale.

In addition, there is ongoing debate about the extent to which male views of body image are influenced by the overall changes in appearance that accompany weight gain and loss. The prevailing view is that upper body muscularity is a more important predictor of male physical attractiveness [19, 20], and recent male body scales based on artist drawings of bodies have manipulated male bodies independently along dimensions of

both weight and muscularity [21]. Methods will need to be developed to enable independent manipulation of weight *and* muscularity in morphing male body images, and further experimental work is needed to refine and optimise these methods before embarking on production of a male version of the MPFS.

The MPFS scale is also limited in that it only depicts the bodies of young female adults, and has no images of older females. Our model pool was limited, and we lacked the images necessary to produce figural scales for older women. There do not appear to be any age-related differences in how the images are rated. Studies are divided as to whether the incidence of body image disturbance is particularly high among young adults of college age [22, 23], relatively stable across the lifespan [24], or increase with age [25]. Whilst subsequent subscales using models of different ages would be valuable, the MPFS presented here is ideally suited to the broad range of research targeting body image disturbance in young adult women.

In conclusion, here we present the Morphed Photographic Figure Scale, a novel figural scale that offers considerable improvements over existing scales for researchers exploring perception of body size, and for interventions targeting body dysmorphia and dissatisfaction. Our aim is to make the MPFS available to researchers in these and any related domains, and to support further development and validation of the MPFS in a variety of fields of interest.

Acknowledgements

Image collection was funded by Leverhulme Trust Research Grant F/00 182/CJ. Stimulus development and rating was supported by the Medical Research Council (MC_UU_12013/6 and MC_UU_12013/7). Dr Helen Bould is supported by a Wellcome Trust Doctoral Training Fellowship.

References

1. Stunkard, A.J., T. Sorensen, and F. Schulsinger, *Use of the Danish Adoption Register for the study of obesity and thinness.* Res Publ Assoc Res Nerv Ment Dis, 1983. **60**: p. 115-20.
2. Thompson, M.A. and J.J. Gray, *Development and validation of a new body-image assessment scale.* J Pers Assess, 1995. **64**(2): p. 258-69.
3. Gardner, R.M., L.M. Jappe, and L. Gardner, *Development and Validation of a New Figural Drawing Scale for Body-Image Assessment: The BIAS-BD.* Journal of Clinical Psychology, 2009. **65**(1): p. 113-122.
4. Swami, V., et al., *Initial examination of the validity and reliability of the female photographic figure rating scale for body image assessment.* Personality and Individual Differences, 2008. **44**(8): p. 1752-1761.
5. Swami, V., D. Einon, and A. Furnham, *The leg-to-body ratio as a human aesthetic criterion.* Body Image, 2006. **3**(4): p. 317-23.
6. Skinner, A.L. and C.P. Benton, *Anti-expression aftereffects reveal prototype-referenced coding of facial expressions.* Psychol Sci, 2010. **21**(9): p. 1248-53.
7. Benton, C.P. and A.L. Skinner, *Deciding on race: a diffusion model analysis of race-categorisation.* Cognition, 2015. **139**: p. 18-27.
8. Rhodes, G. and L. Jeffery, *Adaptive norm-based coding of facial identity.* Vision Res, 2006. **46**(18): p. 2977-87.
9. B. Tiddeman, D.M.B.a.D.P., *Computer Graphics in Facial Perception Research.* IEEE Computer Graphics and Applications, 2001. **21**(5): p. 42-50.
10. Altabe, M., *Issues in the assessment and treatment of body image disturbance in culturally diverse populations*, in *Body image, eating disorders, and obesity* J.K. Thompson, Editor. 1996, American Psychological Association: Washington, DC. p. 129-148.
11. Bould, H., et al., *Effects of visual adaptation on perception of and satisfaction with own body size: two randomised studies.* The Lancet, 2015. **386, Supplement 2**: p. S24.
12. Tiddeman, B., M. Burt, and D. Perrett, *Prototyping and transforming facial textures for perception research.* Ieee Computer Graphics and Applications, 2001. **21**(5): p. 42-50.
13. Gardner, R.M. and D.L. Brown, *Body image assessment: A review of figural drawing scales.* Personality and Individual Differences, 2010. **48**(2): p. 107-111.
14. Tovee, M.J., J.L. Emery, and E.M. Cohen-Tovee, *The estimation of body mass index and physical attractiveness is dependent on the observer's own body mass index.* Proc Biol Sci, 2000. **267**(1456): p. 1987-97.
15. Crossley, K.L., P.L. Cornelissen, and M.J. Tovee, *What is an attractive body? Using an interactive 3D program to create the ideal body for you and your partner.* PLoS One, 2012. **7**(11): p. e50601.
16. Oakes, M.E., C.S. Slotterback, and E.K. Mecca, *Gender differences in perceptions of normal body weights.* Current Psychology, 2003. **22**(2): p. 93-99.
17. Phillips KA, Menard W, Fay C. *Gender similarities and differences in 200 individuals with body dysmorphic disorder.* Compr Psychiatry. 2006;**47**:77–87
18. Rief W, Buhlmann U, Wilhelm S, Borkenhagen A, Brahler E. *The prevalence of body dysmorphic disorder: a population-based survey.* Psychol Med. 2006;**36**:877–885.

19. Coy, A.E., J.D. Green, and M.E. Price, *Why is low waist-to-chest ratio attractive in males? The mediating roles of perceived dominance, fitness, and protection ability.* Body Image, 2014. **11**(3): p. 282-9.
20. Swami, V. and M.J. Tovee, *Male physical attractiveness in Britain and Malaysia: A cross-cultural study.* Body Image, 2005. **2**(4): p. 383-93.
21. Novella, J., J.T. Gosselin, and D. Danowski, *One Size Doesn't Fit All: New Continua of Figure Drawings and Their Relation to Ideal Body Image.* J Am Coll Health, 2015. **63**(6): p. 353-60.
22. Bucchianeri, M.M., et al., *Body dissatisfaction from adolescence to young adulthood: findings from a 10-year longitudinal study.* Body Image, 2013. **10**(1): p. 1-7.
23. Neighbors, L.A. and J. Sobal, *Prevalence and magnitude of body weight and shape dissatisfaction among university students.* Eat Behav, 2007. **8**(4): p. 429-39.
24. Swami, V., et al., *Associations Between Women's Body Image and Happiness: Results of the YouBeauty.com Body Image Survey (YBIS).* Journal of Happiness Studies, 2015. **16**(3): p. 705-718.
25. Frederick, D.A., L.A. Peplau, and J. Lever, *The swimsuit issue: Correlates of body image in a sample of 52,677 heterosexual adults.* Body Image, 2006. **3**(4): p. 413-9.

Table 1: Percentage Perceived as Over, Under or Normal Weight by category

		N	Percentage of images perceived as underweight		Percentage of images perceived as neither over nor underweight		Percentage of images perceived as overweight	
			Mean %	95% Confidence Intervals	Mean %	95% Confidence Intervals	Mean %	95% Confidence Intervals
Total		**334**	**18.5**	**17.6, 19.3**	**47.0**	**45.6, 48.4**	**30.6**	**29.3, 31.9**
By gender	Male	90	18.5	16.7, 20.3	48.0	45.2, 50.8	29.0	26.6, 31.5
	Female	244	18.4	17.4, 19.4	46.7	45.1, 48.3	31.1	29.6, 32.7
By ethnicity	Asian	23	20.6	17.1, 24.1	45.0	38.8, 51.2	27.6	21.1, 34.2
	Black	2	11.7	-9.5, 32.8	60.6	53.5, 67.6	27.2	20.2, 34.3
	Mixed	24	15.9	12.6, 19.2	49.1	43.8, 54.3	28.8	24.2, 33.4
	Other	5	13.3	6.6, 20.0	44.7	26.9, 62.5	37.6	16.7, 58.4
	White	280	18.6	17.7, 19.6	47.0	45.5, 48.5	30.9	29.5, 32.3
By current/previous eating disorder	No	305	18.5	17.6, 19.4	47.3	45.9, 48.8	30.3	28.9, 31.7
	Yes	29	18.1	15.0, 21.2	44.1	39.5, 48.8	33.3	28.7, 37.9
By age bands	<21	185	18.6	17.4, 19.8	46.7	44.9, 48.6	30.6	28.8, 32.4
	≥21	149	18.2	17.0, 19.5	47.4	45.3, 49.6	30.5	28.6, 32.4
By hours since last ate	0-1	118	17.7	16.2, 19.2	47.9	45.4, 50.5	30.7	28.5, 33.0
	1.5-3	114	19.2	17.8, 20.5	45.1	43.0, 47.2	31.1	29.1, 33.7
	3.5-20	102	18.5	16.8, 20.2	48.2	45.5, 50.8	29.5	27.1, 31.8
By BMI bands	Underweight (<18.5)	21	16.8	13.1, 20.6	49.6	42.9, 56.3	30.4	25.6, 35.2
	Normal (18.5-25)	262	18.2	17.3, 19.1	46.5	45.0, 48.0	31.3	29.8, 32.8
	Overweight (>25)	51	20.4	17.7, 23.1	48.7	44.4, 53.0	27.0	23.7, 30.3
By EDEQ weight subscale	0-4	231	18.4	17.3, 19.4	48.9	47.1, 50.6	28.8	27.3, 30.3
	>4	103	18.7	17.2, 20.2	42.9	40.9, 45.0	34.6	32.2, 36.9
By EDEQ shape subscale	0-4	259	18.2	17.2, 19.2	48.2	46.6, 49.8	30.0	28.5, 31.4
	>4	75	19.3	17.4, 21.2	43.1	40.8, 45.5	32.7	29.8, 35.5
By Body Satisfaction score	1-2.8	102	18.7	17.1, 20.3	45.1	42.6, 47.5	31.9	29.4, 34.5
	2.9-3.6	118	18.4	16.9, 19.8	46.9	44.6, 49.3	30.1	28.0, 32.2
	3.7-5	114	18.4	16.9, 19.8	48.9	46.5, 51.3	29.8	27.6, 32.0

Table 2: Between Group Differences in Perceived Weight Scores (positive scores indicate more bodies perceived as overweight)

	Differences in percentage of images perceived as underweight				Differences in percentage of images perceived as overweight			
Difference in perceived weight scores	Unadjusted difference (95% confidence intervals) (N=334)	P value	Adjusted difference (95% confidence intervals) (N=334)	P value	Unadjusted difference (95% confidence intervals) (N=219)	P value	Adjusted difference (95% confidence intervals) (N=218)	P value
Male	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-
Female	-0.07 (-2.01, 1.86)	0.94	0.22 (-1.88, 2.33)	0.84	2.10 (-0.84, 5.03)	0.16	-0.45 (-3.56, 2.66)	0.78
No current/ previous ED	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-
Current/ previous ED	-0.36 (-3.41, 2.69)	0.82	-0.35 (-3.60, 2.88)	0.82	2.99 (-1.64, 7.62)	0.21	0.90 (-3.89, 5.69)	0.71
Age <21	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-
Age ≥21	-0.42 (-2.15, 1.30)	0.63	-0.89 (-2.68, 0.91)	0.33	-0.10 (-2.73, 2.53)	0.94	0.76 (-1.89, 3.41)	0.57
Hours since last ate: 0-1	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-
Hours since last ate: 1.5-3	1.44 (-0.62, 3.5)	0.17	1.30 (-0.80, 3.40)	0.22	0.65 (-2.48, 3.79)	0.68	1.00 (-2.10, 4.11)	0.53
Hours since last ate: 3.5-20	0.79 (-1.33, 2.91)	0.47	0.64 (-1.52, 2.79)	0.47	-1.28 (-4.51, 1.95)	0.44	-1.52 (-4.71, 1.67)	0.35
BMI	**0.29 (0.04, 0.55)**	**0.02**	**0.30 (0.02, 0.57)**	**0.03**	**-0.43 (-0.82, -0.05)**	**0.028**	**-0.52 (-0.93, -0.11)**	**0.01**
EDEQ weight 0-4	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-
EDEQ weight >4	0.32 (-1.54, 2.18)	0.74	-0.34 (-2.87, 2.19)	0.79	**5.80 (3.03, 8.55)**	**<0.0001**	**7.33 (3.59, 11.06)**	**<0.0001**
EDEQ shape 0-4	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-
EDEQ shape >4	1.14 (-0.92, 3.19)	0.28	1.19 (-1.57, 3.96)	0.40	2.74 (-0.37, 5.86)	0.08	-1.03 (-5.11, 3.06)	0.62
BD score 1-2.8	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-	1.00 (1.00, 1.00)	-
BD score 2.9-3.6	-0.29 (-2.42, 1.84)	0.79	0.26 (-2.18, 2.71)	0.83	-1.84 (-5.06, 1.39)	0.26	-0.26 (-3.87, 3.35)	0.89
BD score 3.7-5	-0.31 (-2.45, 1.84)	0.78	0.63 (-2.08, 3.35)	0.65	-2.11 (-5.37, 1.13)	0.20	0.75 (-3.26, 4.76)	0.71

Figure 1 Left panel: Body image marked up with PsychoMorph template. Right panel: Average underweight and overweight bodies used to make PsychoMorph transform.

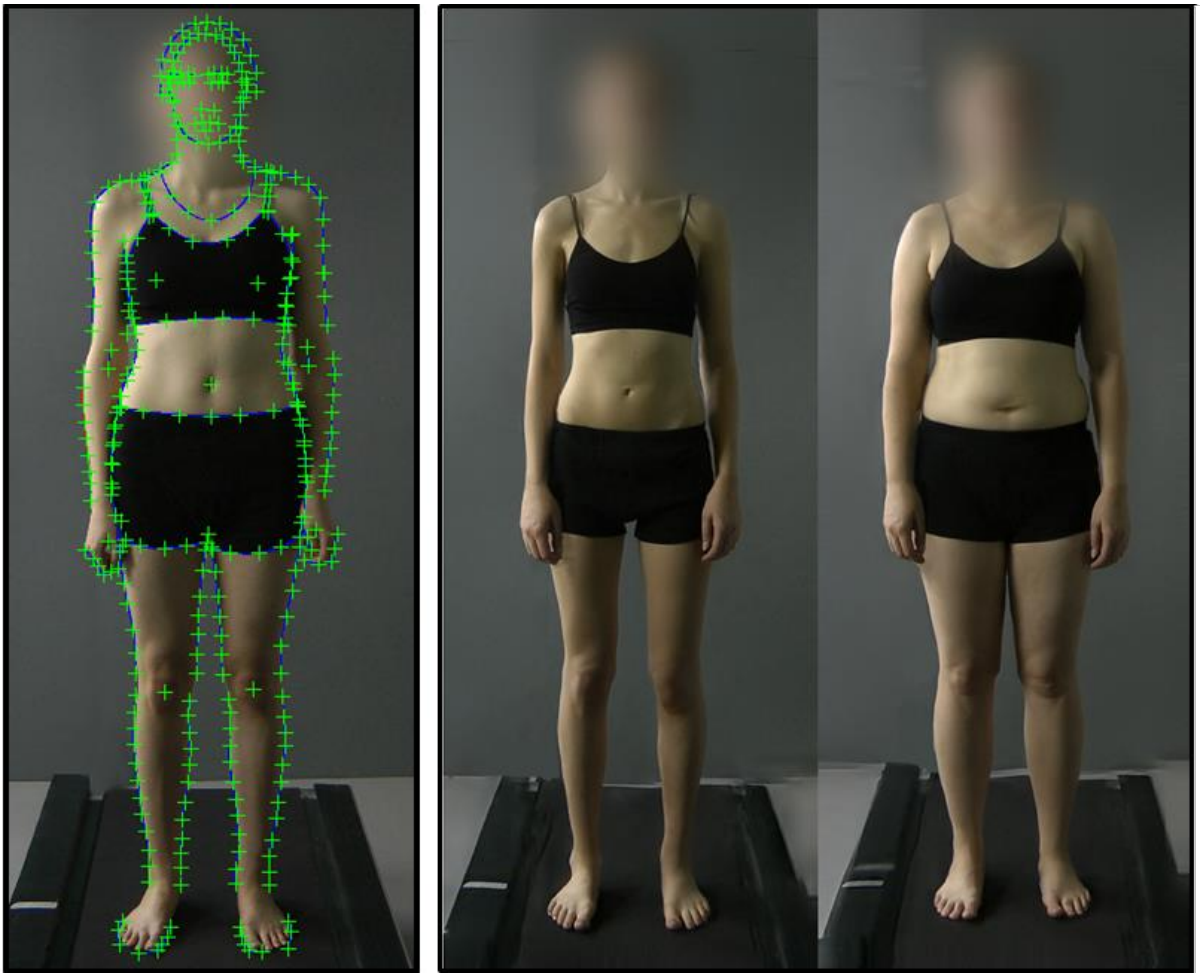


Figure 2 Top panel: Example MPFS sequence (sequence P2). The central image (Ver) is the veridical image of the individual, BMI 16.9. The four to the left are morphed slimmer, the four to the right morphed larger. Bottom panel: Ratings of images in the sequence as underweight, overweight or neither.

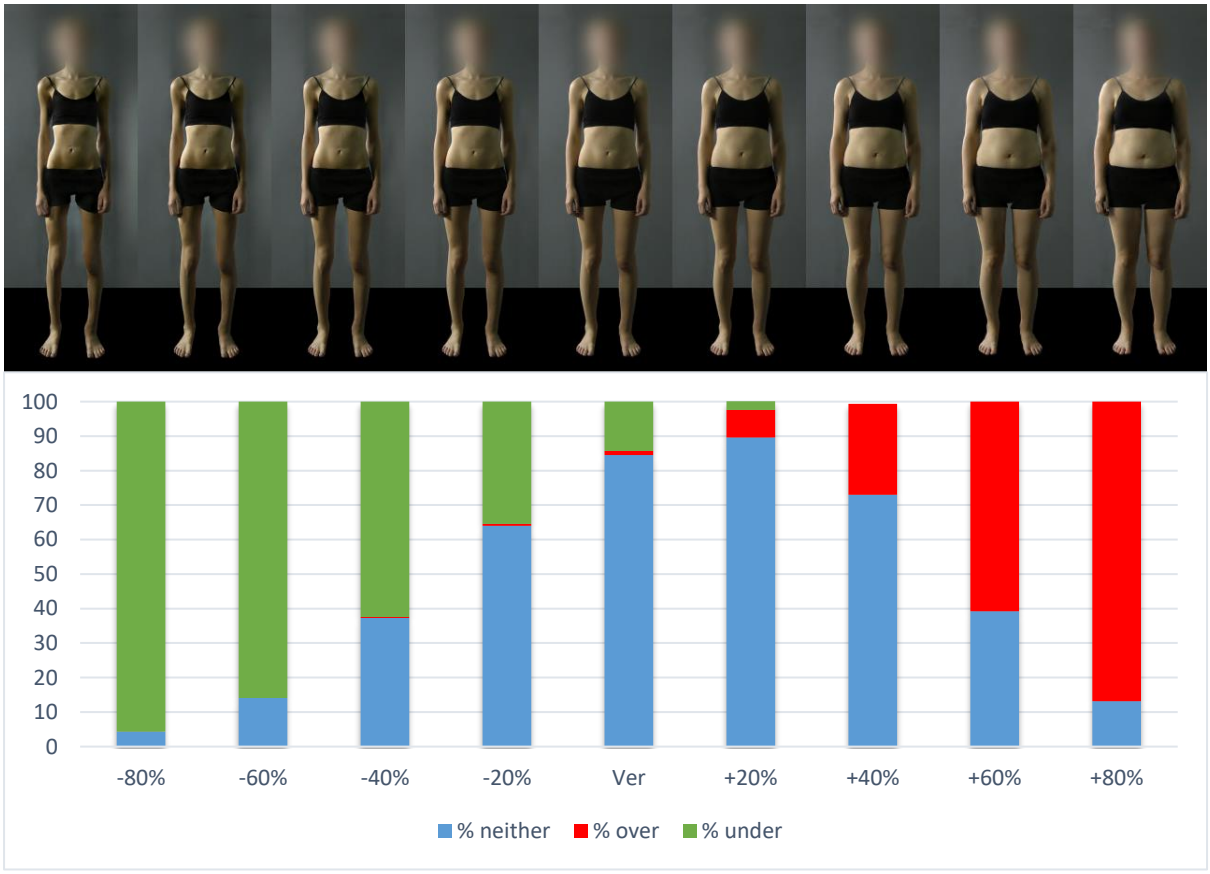
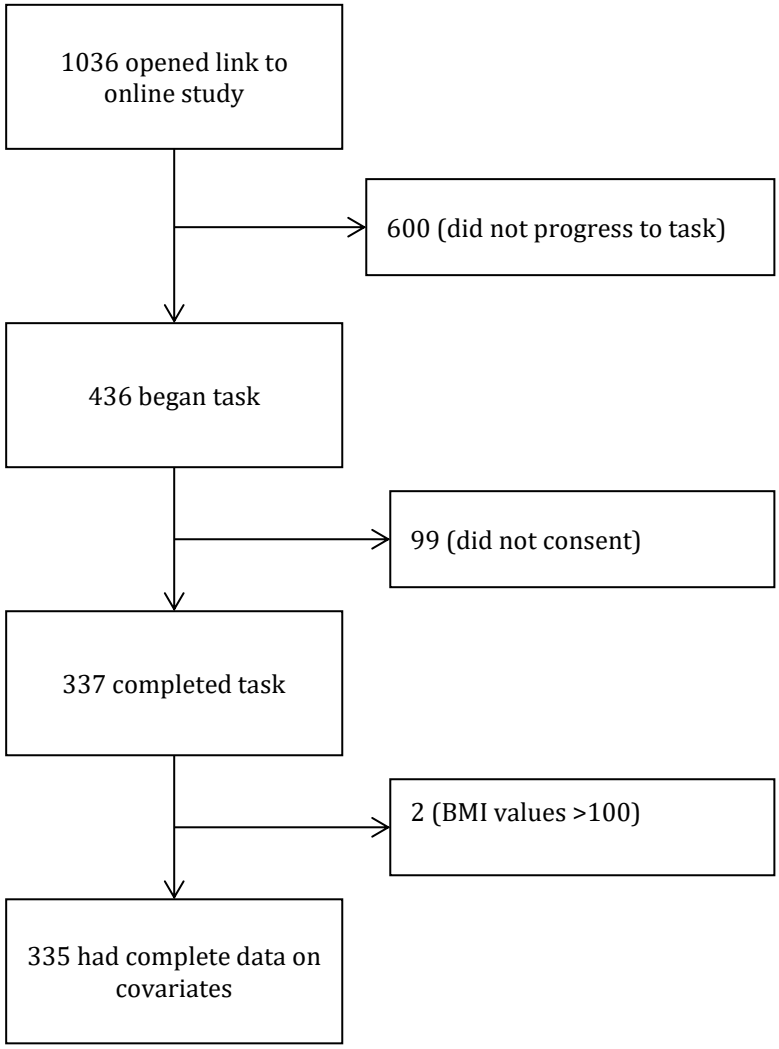


Figure 3: CONSORT diagram



Supplementary Materials

1. PsychoMorph Settings

We used PsychoMorph version 4, with the following options set for the transform and averaging operations:

1.1 Transform Options:

Select Shape, Colour and Texture

1.2 Averaging Options:

Texture processing: Wavelet MRF

Shape normalisation: Full

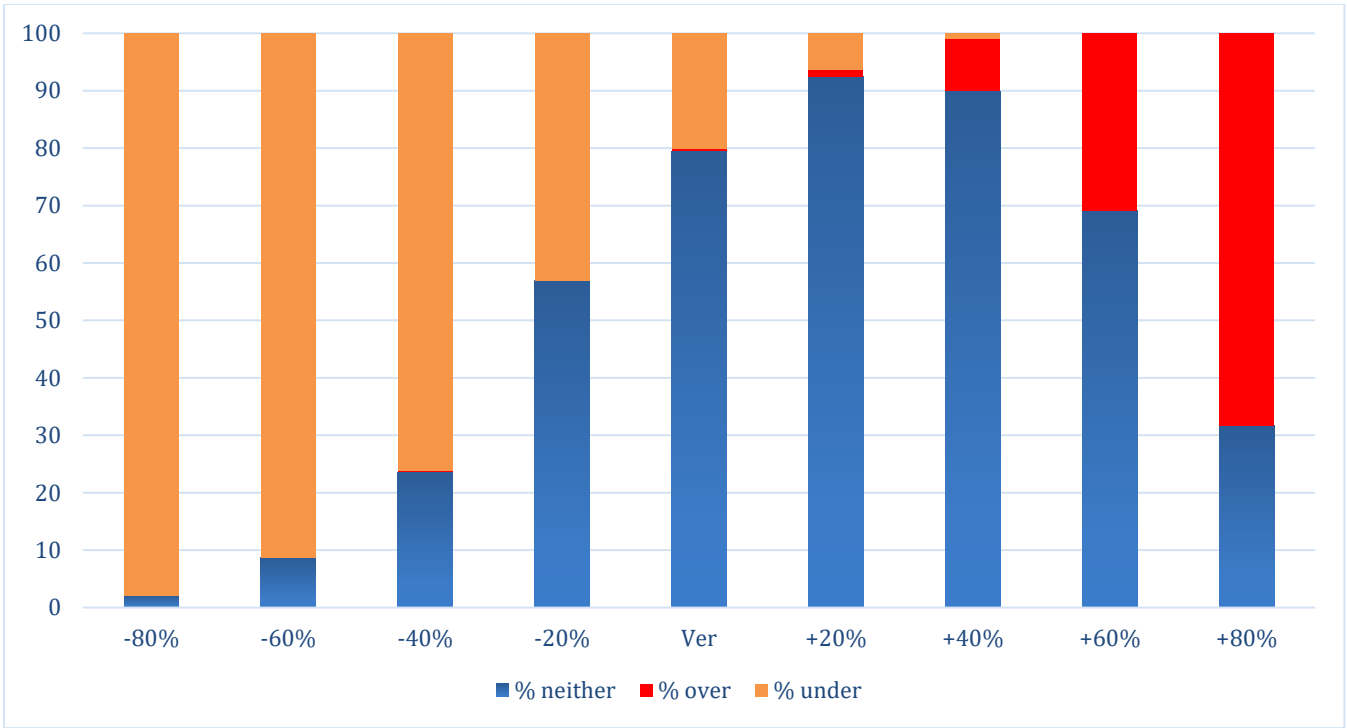
Warping options: Multiscale

Select Recalculate Template

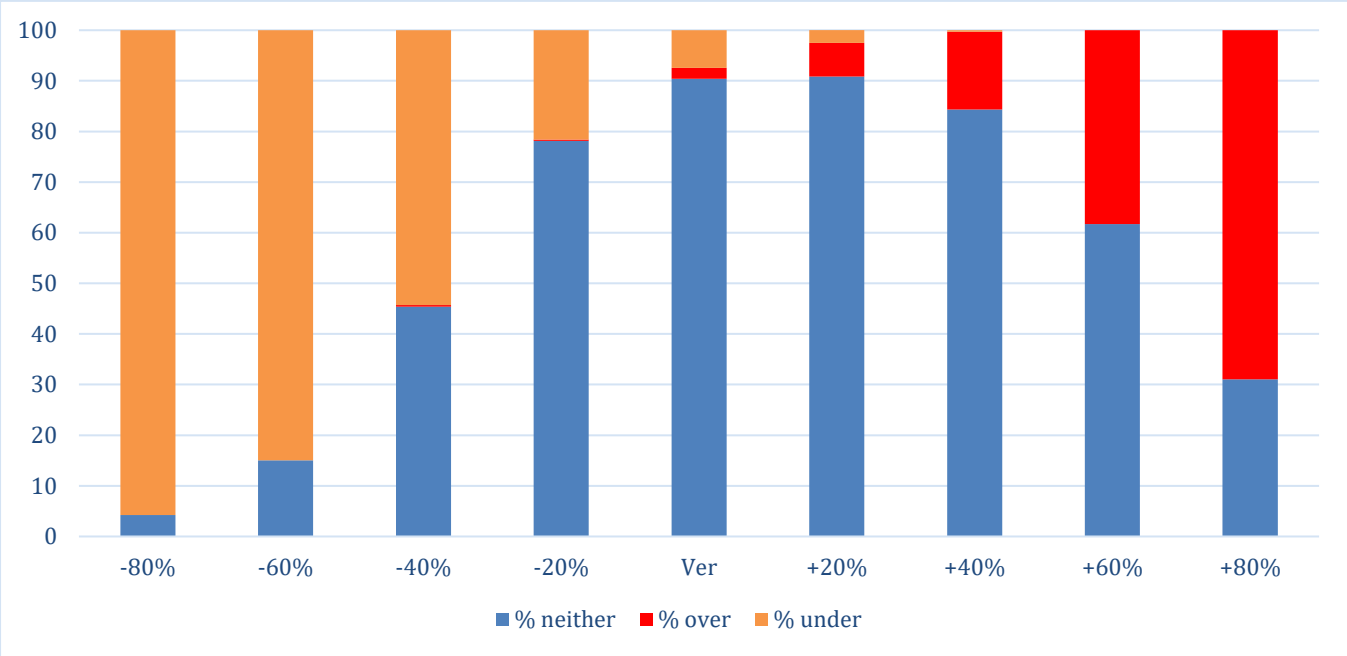
Select Auto Resize Template

2. Complete Stimuli Set

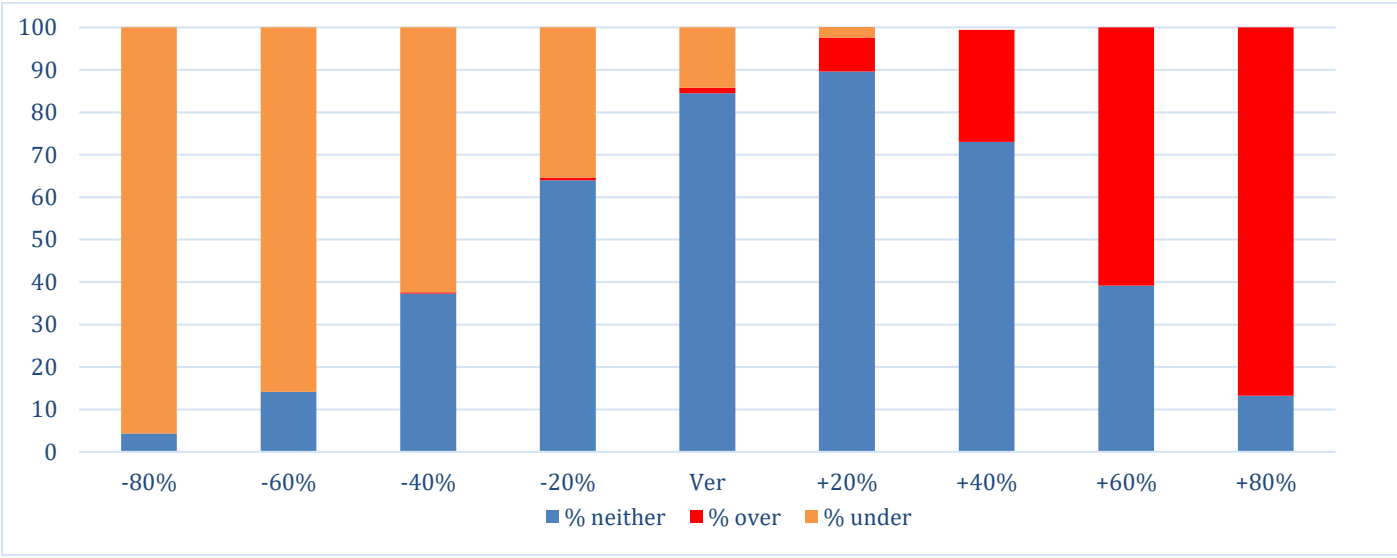
P1 (veridical individual BMI 16.0)



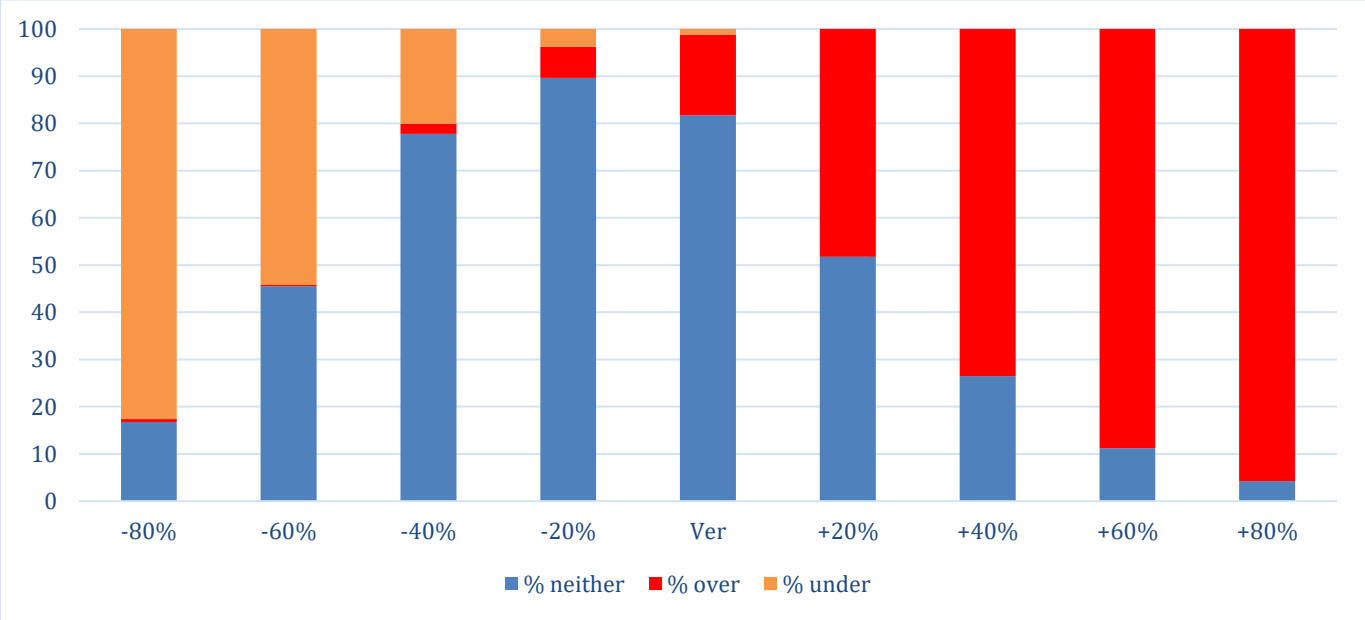
P2 (veridical individual BMI 16.9)



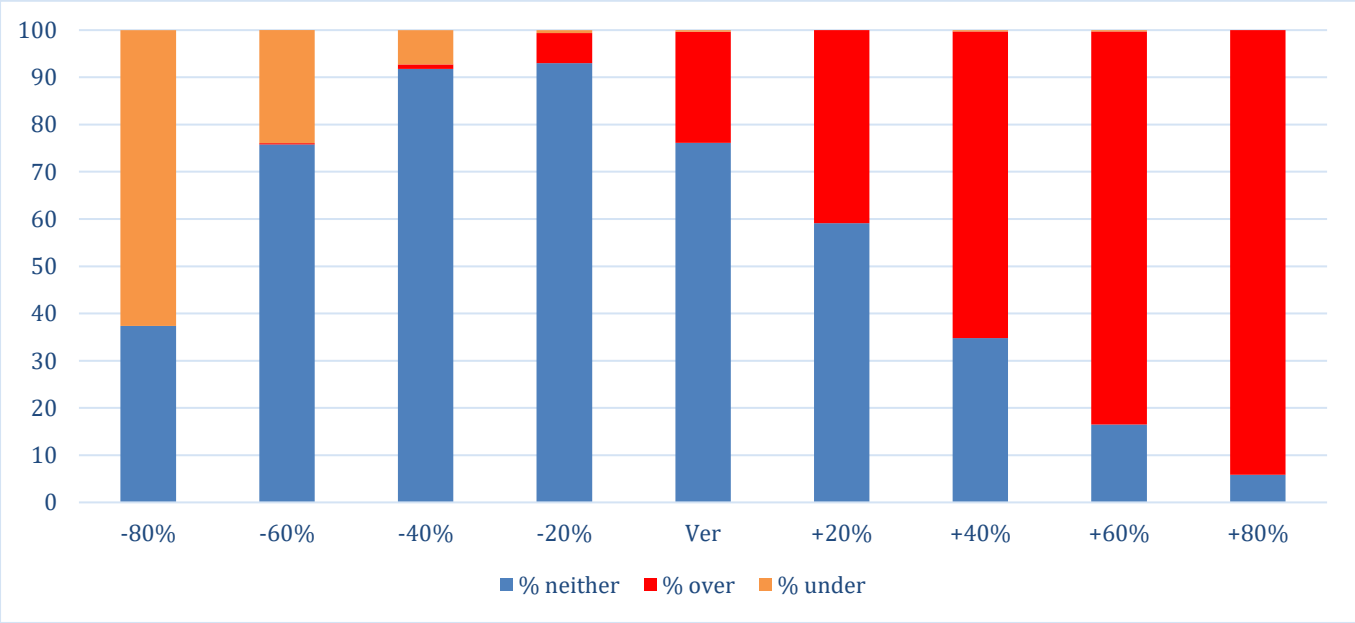
P3 (veridical individual BMI 17.4)



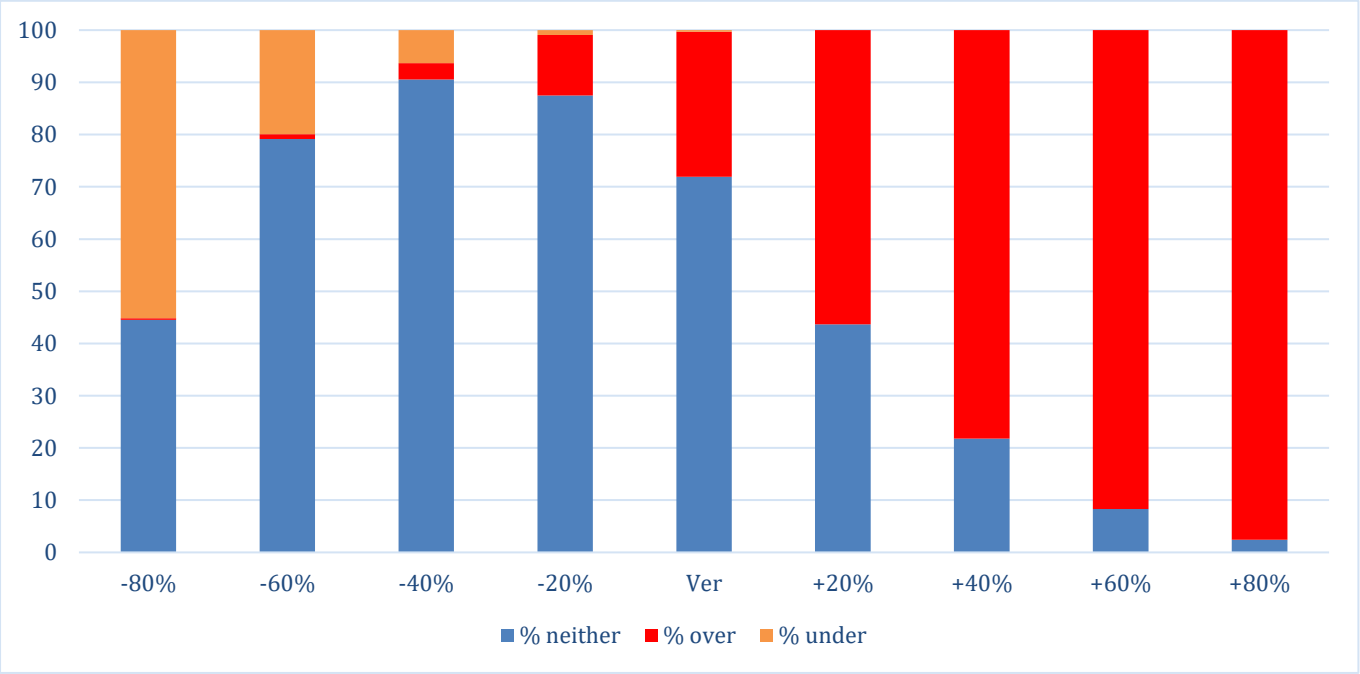
P4 (veridical individual BMI 18.1)



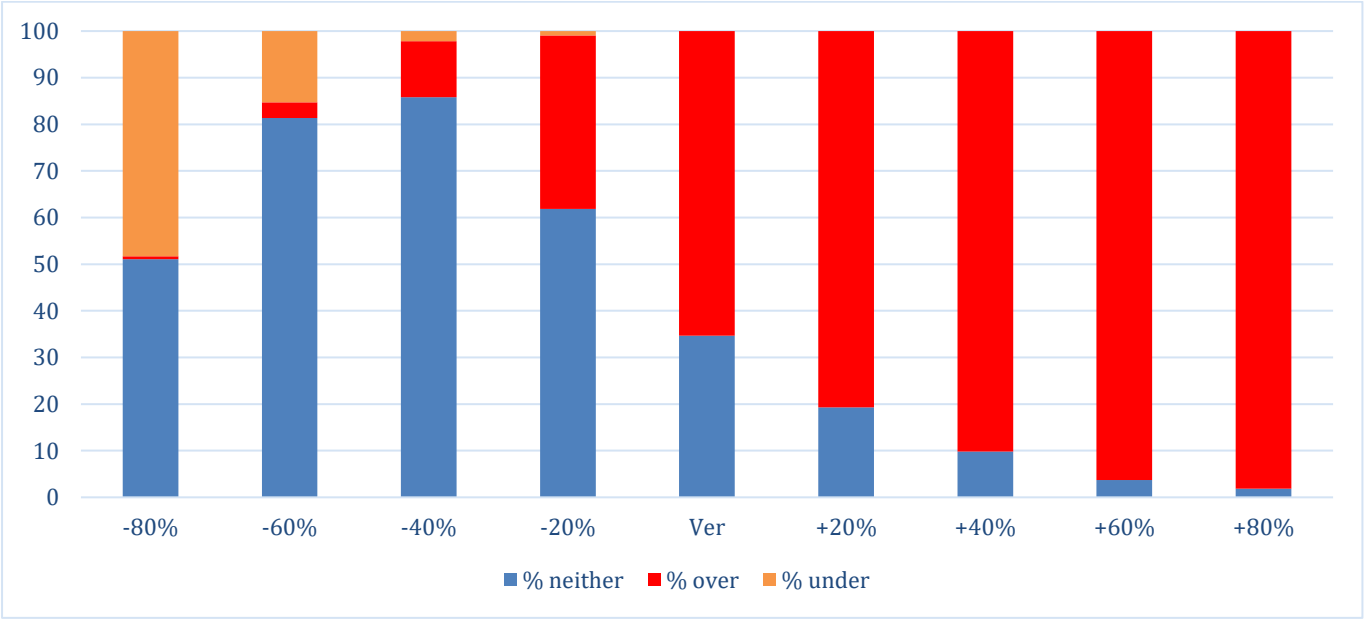
P5 (veridical individual BMI 22.0)



P6 (veridical individual BMI 22.1)



P7 (veridical individual BMI 22.6)



P8 (veridical individual BMI 22.6)