

Running head: *Training Program for Body Image Disturbance*

An Interactive Training Program to treat Body Image Disturbance.

Lucinda J. Gledhill¹, Katri K. Cornelissen², Piers L. Cornelissen^{2*},
Ian S. Penton-Voak³, Marcus R. Munafò^{3,4} and Martin J. Tovée⁵

¹ Institute of Neuroscience, Newcastle University, UK.

² Department of Psychology, Northumbria University, UK

³ School of Experimental Psychology, University of Bristol, UK.

⁴ MRC Integrative Epidemiology Unit at the University of Bristol, UK.

⁵ School of Psychology, University of Lincoln, UK

Word Count: 5737 (excluding abstract, references and figure & table directions and legends).

**Requests for reprints should be addressed to; Prof. Piers Cornelissen, Psychology Department, Northumbria University, Newcastle Upon Tyne, NE1 8ST, UK.. (e-mail: piers.cornelissen@northumbria.ac.uk).

Declaration of Conflicting Interests: I. S. P.-V. and M. R. M. are directors of Jericoe Ltd., which develops software for assessing and modifying emotion perception. The other authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Abstract

Objectives: Anorexia nervosa (AN) is a life-threatening mental health condition. A core feature is a disturbance of body image, such that sufferers see themselves as fatter than they actually are. **Design:** We tested the effectiveness of a novel training program to recalibrate our participants' perception of body size. **Methods:** In a novel adaptation of a cognitive bias training program, participants judged the body size of a series of female bodies and were given feedback to improve their accuracy over 4 daily training sessions. In Study 1, we recruited young women with high concerns about their body size for a randomised controlled study. In Study 2, we then applied the training program to a case series of women with atypical AN. **Results:** In Study 1, the training program significantly improved the body size judgements of women with high body concerns compared to controls. We also found evidence of improved body image and reduced eating concerns in this group. In Study 2, the program again recalibrated the body size judgements of women with atypical AN. We also saw evidence of a clinically meaningful reduction in their body size and eating disordered concerns. **Conclusions:** This training has the potential to be a valuable treatment used together with more traditional talking therapies.

Key Words: Body Size Training, Body Size judgements, Anorexia Nervosa Syndrome Disorder, Anorexia Nervosa.

Introduction

Anorexia nervosa (AN) is a serious psychiatric illness with significant psychological and physiological impact, affecting up to 1% of the female population (Treasure et al., 2010). Current therapeutic regimes have only limited success in treating this condition where long-term mortality can reach 10% and the relapse rate may be as high as 40% over the first 12-months post-discharge (Berkman et al., 2007; Carter et al., 2012). Diagnostic criteria for AN include a distorted evaluation of personal body size (American Psychiatric Association, 2013), which is a core component of psychological models of the disorder (Cash & Deagle, 1997; Fairburn et al., 2003). Body image distortion is one of the most persistent of all eating disorder symptoms, its severity predicts long-term treatment outcome, and its persistence increases the risk of relapse (Channon & DeSilva, 1985; Fairburn et al., 2003;). Most studies have found that people with AN overestimate body size, have negative feelings towards their body and classify lower weight bodies as fat relative to controls (Cornelissen et al., 2013, 2015; Williamson et al., 1993; Gardner & Bokenkamp, 1996; Tovée et al. 2000, 2003; George et al. 2011). This overestimation extends to judgements of other women's bodies (Cornelissen et al., 2016). It is this overestimation and disparagement of body size, coupled with a morbid dread of becoming overweight, which may fuel a drive for thinness through abnormal eating patterns and associated behaviours (such as excessive exercise, purging etc.).

The principal treatment for body image disturbance is cognitive behavioural therapy (CBT), to modify dysfunctional thoughts, feelings and behaviours that contribute to a negative body image. Previous studies in eating disordered and non-eating disordered women have suggested that being exposed to their own body in a mirror (mirror exposure therapy) can help desensitise a patient to their body appearance (Key et al., 2002; Delinsky & Wilson, 2006; Vocks et al., 2008) . However, this does not directly tackle the issue of body size over-estimation. Other interventions include: fitness training to improve physical capacity and shift attention from appearance to functionality (Farrell et al., 2006; Jarry & Cash, 2011), media literacy to challenge the impact of images of thin bodies (Martin & Lichtenberger, 2002; Ginis & Bassett, 2011), self-esteem enhancement to improve self-worth (Grabe et al., 2008; Irving & Berel, 2001), and psychoeducation (O'Dea, 2004. O'Dea & Yager, 2011). However, a recent meta-analysis suggested that once corrections for bias (both within and across studies) in the data were applied, the effect sizes of these treatments were relatively small, strongly suggesting the need for new additional therapies to address negative body image (Alleva et al., 2015).

Therefore, in the current study, we tested the effectiveness of a novel body training program to increase individual participants' categorical boundaries for thin/fat bodies towards fatter bodies, and we tested whether this reduced participants' body size and eating concerns. This approach has

been developed from a face training program, used to modify biases in emotion recognition in order to encourage the perception of happiness over anger in ambiguous expressions in adolescents who are at high risk of criminal offending and delinquency (Penton-Voak et al., 2013). In the original study, participants were presented images of faces and had to make a 2-alternative forced choice (2-AFC) decision whether a particular face was happy or angry. The face stimuli were selected at random from a sequence of images which morphed smoothly from clearly happy at one end of the continuum to clearly angry at the other end. Images in the middle of the sequence (intermediate in their expression) could be judged either way. In such a task, any given individual will tend to perceive a category boundary somewhere along this continuum which, for them, constitutes a transition from happy to angry. However, there is no absolute location in the continuum which corresponds to a 'correct' category boundary; instead the location of this perceptual boundary is subjective and varies across different individuals. Indeed, adolescents at high risk of offending, unlike controls, tend to categorise even the intermediate faces as angry. However, by giving appropriately structured feedback, their angry-happy categorical boundary can be shifted towards the 'happy' end of the spectrum. Penton-Voak et al. (2013) found that this resulted in a decrease in self-reported anger and aggression and in independently rated aggressive behaviour. The same training algorithm has also been used to recalibrate the perception of happiness over sadness in ambiguous facial expressions in people reporting high levels of depressive symptoms to improve mood (Penton-Voak et al., 2011; Adams et al., 2013).

We have previously shown that body size judgements are also made in a categorical manner; observers show a clear categorical boundary in making 2-AFC judgements about whether a body is fat or thin (Tovee et al., 2012). Moreover, observers show a clear enhancement in discriminating between bodies in matching-to-sample tasks if the discrimination is between bodies at the categorical boundary compared to discriminating between bodies from within the same category (Tovee et al., 2012). Therefore, we tested the effectiveness of a modified version of the training paradigm from Penton-Voak et al. (2013) to shift participants' thin-fat categorical boundary and asked whether this shift would lead to a general improvement in body image concerns. It is important to emphasize that this is not the same as trying to train participants to achieve some fixed, normative goal, constituting a 'normal' body size. Instead, the intention was to move an individual's thin-fat categorical boundary towards fatter bodies, irrespective of their starting point on the thin to fat continuum.

We hypothesise that our training will shift a persons' thin-fat categorical boundary, so that they will categorise as thin, bodies that they had previously categorised as fat and that this categorical shift will be accompanied by a significant reduction in their body image and eating disordered concerns. To test these hypotheses we carried out two studies. In the first, we recruited

young women with high concerns about their body size for a randomised controlled study to determine whether the training alters the perceptual position of the thin-fat categorical boundary with associated improvements in body image and eating concerns. In the second study, we then applied the training to a case series of women with atypical anorexia nervosa (aAN) to determine whether the training would also lead to a change in categorical boundary in this population and their more deep-seated body image concerns.

Study 1

Participants

Forty participants for this study were recruited from undergraduate students at ***** University. Potential participants, who self-reported no history of eating disorders, were asked to provide their height and weight and to fill in the 16-item Body Shape Questionnaire BSQ (Evans & Dolan, 1993). This psychometric tool indexes the degree of preoccupation and negative attitude toward body weight and body shape. Only those who achieved a BSQ score of 60 or greater (i.e. substantial body shape concerns) were eligible to participate. Participants were randomly assigned to two conditions. Twenty female participants (mean age: 18.15 years; SD: 0.37), were assigned to the intervention condition, while 20 females (mean age: 19.00 years; SD: 1.26), were assigned to the control condition. Data from previous studies of similar training programs indicate an effect size of $d \sim 1$ for the effect of training on perceptual position of the categorical boundary, suggesting a total sample size of $n = 34$ would be sufficient to achieve 80% power at an alpha level of 5%. Table 1 describes the participants' characteristics. Multiple t-tests were computed, using both Satterthwaite's correction for unequal variance and Bonferroni's correction for multiple comparisons, to compare characteristics between the intervention and control groups. None exceeded the critical value of $p < .0083$. Ethical approval for this study was granted by the Faculty of Medical Sciences ethics committee at ***** University (00620/2013). All procedures contributing to study 1 and 2 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.

Table_1

Measures

Psychometric and anthropometric measurements

To assess participants' attitudes to body shape, weight and eating we used the Eating Disorders Examination Questionnaire (EDE-Q), which is a self-report version of the Eating Disorder Examination (EDE) structured interview (Fairburn & Beglin, 1994). This is commonly used as a screening questionnaire for eating disordered behaviour and has been normed for young women and undergraduates (Mond et al., 2006; Luce, Crowther & Pole, 2008). The questionnaire contains four subscales reflecting the severity of aspects of the psychopathology of eating disorders: (i) the Restraint (EDE-restraint) subscale investigates the restrictive nature of eating behaviour; (ii) the Eating Concern (EDE-eating concerns) subscale measures preoccupation with food and social eating; (iii) the Shape Concern (EDE-shape concerns) subscale investigates dissatisfaction with body shape and (iv) the Weight Concern (EDE-weight concerns) subscale assesses dissatisfaction with body weight. The EDE-Q (range 0-6) also measures overall disordered eating behaviour. Furthermore, it provides frequency data on key behavioural features of eating disorders. We also used the Beck Depression Inventory (BDI) (range 0-63; Beck et al., 1961) that measures participants' level of depression and the Rosenberg Self-Esteem Scale (RSE) (range 0-30; Rosenberg, 1965) that measures self-esteem. In addition, we calculated the participants' body mass index (BMI) from their weight obtained with a set of calibrated scales and their height obtained with a stadiometer.

Stimulus image preparation

We used computer-generated imagery (CGI) methods to create a sequence of 3D images of a model whose body shape changes systematically with increasing BMI [for details see Cornelissen et al. (2015)]. The advantages of the CGI stimuli are: i) the identity of the person in the image is maintained over a wide BMI range; ii) the body shape changes at different BMI levels are realistic and iii) the 3D rendered stimulus images are high definition and photorealistic. We used a set of 15 images ranging in BMI from 15.4 to 33.7 (drawn from the image database in Cornelissen et al., 2015).

Perceptual training paradigm

An E-Prime (<http://www.pstnet.com/eprime.cfm>) script was modified from Penton-Voak et al. (2012, 2013) to run the training paradigm on a Windows PC with a 19" LCD monitor panel (1600x1200 native pixel resolution, 32-bit colour depth). Each trial of the baseline and training conditions began with a central fixation cross which was shown for 1500-2500 ms (randomly jittered). This was replaced by an image of a body for 150 ms followed immediately by a mask of visual noise for 150 ms. Finally, the mask was replaced by a prompt screen, containing a "?", to indicate that participants should make their judgement of "fat" or "thin" (a 2-alternative forced

choice) and respond by key-press. No time limits for the response were imposed on the task. The baseline condition comprised 45 trials in all (3 presentations of each of the 15 stimulus images) in randomized order. Participant responses were used to calculate the categorical boundary/midpoint at which they shifted from perceiving thinness to perceiving fatness in the body sequence 50% of the time (see upper row of Fig. 1). Each training session typically took 30-45 minutes.

Figure_1

Trials from the training phase differed from the baseline procedure in that feedback (i.e. “Incorrect! That body was fat” or “Correct. That body was thin”) was provided to the participant following their keyboard response. The training phase was made up of 6 blocks, with 31 trials in each block. From pilot testing, bodies 1-2 and 14-15 were almost always classified as ‘thin’ and ‘fat’ respectively, so these were only presented once in training. Bodies 3-5 were responded to less frequently as being ‘thin’, and 11-13 less frequently as being ‘fat’ and were therefore presented twice. The remaining bodies, 6-10, were presented three times each as responses to these bodies were the least clear-cut.

In the intervention condition, the nature of the feedback given was ‘inflationary’ and was designed to shift a participant’s categorical boundary by two bodies higher up along the image sequence (from low to high BMI) than their baseline measure. In this way participants were re-trained to judge bodies near their categorical boundary, which they had previously judged as fat during baseline measurement, to be thin. By contrast, in the control condition, the feedback to participants was consistent with their categorical boundary as measured at baseline, and was intended merely to reinforce their existing categorical boundary.

Procedure

The training took place over 4 days. On Day-1, participants first completed the EDE-Q, RSE and BDI questionnaires and had their height and weight measured. Then they carried out the first baseline and training sequences for the categorical perception task. On Days 2 and 3, participants carried out the baseline and the training sequences only. On Day-4, participants completed the baseline and training sequences, followed by the EDE-Q questionnaire. To test retention of the training on Day-14, participants carried out the baseline sequence and then completed the EDE-Q questionnaire.

Results

We used PROC MIXED (SAS v9.3) to fit a multi-level model to the perceptual training data which included three main effects (i.e. group: intervention versus control; training: baseline versus post-training threshold; test day: 1, 2, 3, 4, 14) and all possible two and three-way interactions. In addition, based on significant reductions in $-2\log$ likelihood, we permitted both individual slope and intercept variation for each subject and specified an ‘unstructured’ variance-covariance structure for the G-matrix. This model allowed us to compute post-hoc pairwise tests, which were controlled for multiple comparisons, as illustrated in Figure 2A and 2B. We found statistically significant main effects of test day ($F_{4, 304} = 6.69, p < .0001$) and training ($F_{1, 204} = 15.11, p = .0001$) on perceptual thresholds, but not for group ($F_{1,38} = 1.46, p = .23$). In addition, we found statistically significant interactions: *group x test day* ($F_{4, 304} = 16.62, p < .0001$) and *group x training* ($F_{1, 304} = 15.93, p < .0001$), but not *training x test day* ($F_{3,304} = 0.14, p = .93$) nor *training x group x test day* ($F_{3,304} = 0.08, p = .97$).

Figure_2

Fig. 2A shows a plot of the LSmeans for the categorical boundary, derived from the multi-level model, as a function of training day. The data are plotted separately for the control (red and pink) and intervention groups (blue and cyan), each split according to whether the measurement was the pre-training baseline (cyan and pink) for that day or the post-training value (red and blue). We found negligible difference between the baseline and post-training thresholds for the controls, as illustrated in Figure 2B. This plot shows the LSmean difference between baseline and post-training measurements as a function of training day, separately for the intervention and control groups. The error bars represent the 95% CIs for the pairwise comparisons. By contrast, Figure 2B shows a significant effect of training on each training day for the intervention group. This result, together with the significant *group x test day* and *group x training* interactions, shows that training causes an accumulating shift in participants’ categorical boundary towards heavier bodies for the intervention group, but not the control group. Finally, we found that the comparison between Day-1 baseline and Day-14 baseline showed a statistically significantly increase in BMI at the category boundary for the intervention group ($t_{53.7} = 4.42, p < .0001$). In comparison, the controls showed a small reduction that was marginally statistically significant ($t_{53.7} = -1.86, p = .07$). This suggests that, in this sample of non-eating disordered participants, all of whom have high body shape concerns, the perceptual training effect on their categorical boundaries persisted in the intervention group for at least two weeks.

For the psychological measures, we used PROC MIXED (SAS v9.3) to fit separate multi-level models to participants’ global EDE-Q scores, as well as the EDE-restraint, EDE-eating

concerns, EDE-shape concerns and EDE-weight concerns sub-scale scores, measured on: Day-1 before the perceptual training, Day-4 after perceptual training and on Day-14 after the baseline perceptual measurement. Each model contained group and test day as main effects, together with the interaction *group x test day*. Continuous outcome and explanatory variables were centred for these analyses by converting them to z-scores. As before, based on significant reductions in -2log likelihood, we permitted both individual slope and intercept variation for each subject and specified an ‘unstructured’ variance-covariance structure for the G-matrix. Each model allowed us to compute post-hoc pairwise comparisons, all of which are controlled for multiple comparisons, and they are all reported in Table 2. The LSmeans for the global EDE-Q scores for the intervention (blue) and control (red) groups, derived from the multi-level model, are plotted as a function of training day in Figure 1C. The LSmean differences in global EDE-Q scores between the control and intervention groups, together with their 95% CIs, are plotted as a function of training day in Figure 1D.

Table_2

Table 2 shows that, with the exception of the EDE-eating concerns scores, we found statistically significant reductions in EDE-Q, EDE-restraint, EDE-weight concerns and EDE-shape concerns scores for the intervention group compared to controls for test days 4 and 14 but not day-1. Broadly, these results suggest the training has a positive effect on participants’ attitudes to body shape, weight and some aspects of their eating behaviour. Finally, we compared Day-1 with Day-14 scores on the EDE-Q and its subscales. Overall, the EDE-Q global score showed a modest, albeit marginally significant, reduction for the intervention group ($t_{75.5} = -1.81, p = .07$), and no difference for controls ($t_{75.5} = 0.24, p = .81$). The EDE-weight concerns subscale of the EDE-Q also showed a marginally significant reduction for the intervention group ($t_{73.3} = -1.94, p = .06$), and a significant increase for controls ($t_{73.3} = 2.16, p = .03$). This suggests that, in a non-eating disordered sample of women, all of whom have high concerns about body shape, the perceptual intervention shifts attitudes to body shape, weight and eating in a way that would be beneficial for people with eating disorders.

Discussion

Study 1 demonstrates that the training program significantly shifts the thin-fat categorical boundary in individuals with high body concerns. This change is specific to the intervention group who received inflationary feedback. There was a significant reduction in the EDE-Q scores,

particularly on the restraint, weight and shape sub-scales, suggesting that the training not only shifts categorical boundary but also generalises to impact on body size and shape concerns and eating restraint. These changes were retained at 2 weeks post training, suggesting that these represent statistically significant and long lasting changes to attitudes to body size and eating. This suggests that the training has the potential to be used to improve body image. However, it can be argued that women with an eating disorder have more deep-seated concerns that may be harder to modify. To test the training's effectiveness and feasibility we therefore recruited an eating disordered population for study 2.

Study 2

Rationale

In Study 2, we recruited an eating disordered cohort of women to determine whether it was also possible to alter their thin-fat categorical boundary and whether there would be an associated positive change in their mental state. In the absence of a control group, we wanted to control for the possibility that changes in the psychological scores over time might occur merely as a result of habituation to the task, given that the EDE-Q was repeated over multiple test days. For this reason, we included another cognitive task, the Wechsler Adult Intelligence Scale (WAIS-II) Digit Span, the responses to which should not be influenced by concerns about body shape and weight. Nevertheless, because we administered this task as many times as the psychological questionnaires about body shape, weight and eating, we should also expect there to be changes in participants' responses which are related to practice effects. Therefore, we included Digit Span as a covariate in our analyses of the psychological data, on the grounds that it should control for such influences, and any real changes in attitude towards body image should survive this statistical control.

Participants

We recruited 21 female out-patients into the study (mean age $M = 27.71$, $SD = 7.48$), all of whom had a current diagnosis of AN, but whose BMI tends to be higher than is typically the case for in-patients (mean BMI $M = 19.51$, $SD = 3.30$) and so no longer fit the strict DSM-5 diagnostic criteria with respect to BMI, although they still have very high body size concerns and eating disordered behaviours. Means and standard deviations for BSQ and EDE-Q were: 62.48 (18.77) and 3.63 (1.60) respectively. We therefore refer to this group as suffering from atypical anorexia (aAN), following Cornelissen et al. (2015). Data from previous studies of similar training programs indicate an effect size of $d_z > 1$ for the pre- to post-training change in perceptual position of the

categorical boundary, suggesting a total sample size of $n = 10$ would be sufficient to achieve 80% power at an alpha level of 5%.

Three aAN participants failed to adhere to the training regime, therefore we ran two different analyses. In the first, based on an intention to treat, we included all 21 participants. In the second analysis, we excluded these three participants.

The experimental procedures and methods were approved by the local ethics committee at ***** University. Participants were recruited through two third sector organizations: the Beating Eating Disorders Organisation and the Northern Initiative on Women and Eating Organisation.

Measures

We used the same psychometric, anthropometric and psychophysical measures as in study 1. However, we did not include the BDI and RSE because Experiment 2 included one more measurement point than study 1, and we did not want to over-burden participants. We also modified the EDE-Q. Participants fill in the EDE-Q at day 1 (before the training), at day 4 (after the training, at day 7 and day 30. The EDE-Q asks participants to report concerns/behaviours which occurred over the previous 28 days. This time scale was retained on day 1 and 30, but on day 4 we altered the time to the last 24 hours, and on day 7 we altered the scale to the last 7 days. This allows us to detect the changes the training may have had on body image and eating concerns, and which might otherwise be masked by the longer time frame which also covers the pre-training period. In addition, as mention above, we included the Digit Span task from the WAIS-II IQ test battery (Wechsler, 1981), which assesses short-term memory (mean Digit Span on initial testing was $M = 13.85$, $SD = 3.82$).

Procedure

On Day-1, participants first completed the Digit Span task, the BSQ and EDEQ questionnaires and had their BMI measured. They then carried out the first baseline and training sequences for the categorical perception task. On Days 2 and 3, participants carried out the baseline and the training sequences only. On Day-4, participants completed the Digit Span task and the EDEQ questionnaire, followed by the baseline and training sequences. On Days 7 and 30, participants completed the Digit Span task and the EDEQ questionnaire and carried out the baseline sequence only.

Results

Based on intention to treat, PROC MIXED (SAS v9.3) was used to fit a multi-level model to the perceptual training data for all 21 participants. This included two main effects (i.e. training: baseline versus post-training threshold; test day: 1, 2, 3, 4, 7, 30) and the two-way interaction between them. Based on significant reductions in -2log likelihood, we permitted both individual slope and intercept variation for each subject and specified an ‘unstructured’ variance-covariance structure for the G-matrix. As before, this allowed us to compute post-hoc pairwise tests, which were controlled for multiple comparisons, and these are illustrated in Figure 3A and 3B. We found statistically significant main effects of test day ($F_{5,16.6} = 4.07$, $p = .01$) and pre-/post-training ($F_{1,64} = 21.53$, $p < .0001$) on perceptual thresholds, but no significant interaction between them ($F_{3,64} = 1.13$, $p = .35$).

Figure_3

Figure 3A shows a plot of the LSmean categorical boundary from the perceptual task, derived from the multi-level model, as a function of training day. The data are plotted separately for the pre-training baseline for that day (cyan) or the post-training value (blue). We found statistically significant differences between the pre-training baseline and post-training thresholds for days 3 and 4, but not 1 and 2, as illustrated in Figure 3B. Figure 3A shows a somewhat surprising difference between the Day-1 post-training threshold (19.53) and a *higher* Day-2 pre-training threshold (19.85). However, this difference was not statistically significant ($t_{38.8} = 0.95$, $p = .348$). Overall, these results, together with the significant main effects of test day and group, show that training is associated with an accumulating shift in the categorical boundary towards heavier bodies in the aAN participants. Finally, we found that the comparison between Day-1 baseline and Day-30 baseline showed a statistically significantly increase in categorical boundary ($t_{21.9} = 3.06$, $p = .006$). This suggests that, in this sample of aAN participants, the change in perceptual training persisted for at least a month.

For the psychological measures, we used PROC MIXED (SAS v9.3) to fit separate multi-level models to participants’ EDE-Q, EDE-restraint, EDE-eating concerns, EDE-shape concerns and EDE-weight concerns scores, measured on Days 1, 4, 7 and 30. Continuous outcome and explanatory variables were centred for these analyses by converting them to z-scores. Each model comprised the main effect of test day, together with chronological age and Digit Span as covariates. As before, based on significant reductions in -2log likelihood, we permitted both individual slope and intercept variation for each subject and specified an ‘unstructured’ variance-covariance structure for the G-matrix.

The first analysis, based on an intention to treat, and which included the 3 participants who failed to carry out the training program as designed, did not produce statistically significant effects of any explanatory variable for EDE-Q or any of its four sub-scores. The analysis for the 18 aAN participants who did complete the program as required, was more successful. The post-hoc pairwise comparisons from these models are all reported in Table 3, and the data for EDE-Q are plotted in Figure 2C and 2D.

Table_3

Table 3 shows that, in the smaller sample of 18 aAN participants who completed the training regime as prescribed, perceptual training was associated with significant reductions in body shape, weight and eating concerns (with the exception of EDE-restraint), even when Digit Span was controlled for. Moreover, these reductions persisted up to a month from initial testing.

A key question in this test of the training regime is whether the perceptual training and the alteration in the categorical boundary may be linked to the observed changes in psychological profile. If so, then the degree to which the categorical boundary is shifted should be proportional to the change in psychological profile. In the case of study 1, the intervention produces a largely uniform shift in the boundary with comparatively little variation between individuals in the shift (mean change in BMI units 2.33 SD 2.32) and the change in the global EDE-Q score (mean change -0.28 SD 0.33) and so there is no significant correlation between boundary shift and psychological change. In the case of study 2, where participants' body image concerns are more deeply established, there is a greater variation in the effect of the training on the categorical boundary position (mean change 2.75 SD 3.56) and their psychological scores (mean change -0.45 SD 1.03) consistent with a greater resistance to change in some aAN participants. In this case, there is a substantial correlation between how far the boundary shifts and the change in the psychological scores ($r = -0.63$, $p = .004$), suggesting the two may be linked. This correlation with the categorical shift holds for all the EDE-Q subscales, not just those linked to body size and shape (EDE-eating concerns, $r = -0.65$; EDE-restraint, $r = -0.56$; EDE-shape concerns, $r = -0.53$; EDE-weight concerns, $r = -0.52$).

Discussion

In this set of studies, we tested the use of a novel cognitive bias modification technique to shift the body size perception, reduce body image concerns (a core feature of AN), and explored improvement in eating disordered concerns. We have already shown that bodies are judged in a

categorical manner (Tovee et al., 2012) and the first step in testing this training was to determine whether it could shift the categorical boundary in thin-fat judgements of bodies varying in BMI. In study 1, we recruited young women with high body size concerns (but without a formal diagnosis of AN) to test whether their thin-fat categorical boundary could be altered. The results showed a significant shift in the categorical boundary in the intervention group relative to the controls, and this shift was retained 2 weeks after the training. The EDE-Q scores (a general measure of eating disordered concerns which is often used as a screening questionnaire) showed a significant reduction in the intervention condition relative to controls, and this reduction persisted 2 weeks after the end of training.

In study 2 we recruited a cohort of women with aAN. These women have long-standing problems with body image and eating disordered behaviour and thus potentially might be more resistant to the training altering their thin-fat categorical boundary and therefore any associated improvement in body and eating disordered concerns. The follow up period was also extended from 2 weeks to a month to test retention of the training. This second study did not have a control group, but we compensated for this by using a digit span task as a control task. As this task is unrelated to body image and eating disorders and should be unaffected by the body training, it can act as a proxy for any habituation / practice effects caused by repeating the questionnaire measures. As in study 1, there was a significant shift in the categorical boundary during training that was retained a month later. There was also a significant change in the EDE-Q scores over the course of training and which was also retained a month later. These changes were still statistically significant even when variance in the digit span task was taken into account. What is important here is not that women with aAN can be trained to alter the categorical boundary at which they classify a body to be “fat”, but that this change is retained and seems to generalise to other aspects of their body image and eating disordered concerns. The degree to which the categorical boundary shifts is significantly correlated with the change in the EDE-Q score, which is consistent with a causal relationship between the boundary change and the psychological scores. This relationship is true for all the subscales, not just those related to body size or shape. This suggests an effect of the training beyond simply changing the participants’ judgements of body size, but has a more global effect on body and eating concerns. It also provides support for the hypothesis that biases in the perception of body size play a causal role in the maintenance of body image concerns and eating disordered behaviour. One interpretation of these results is that this modification establishes a virtuous cycle, whereby the change in perception of body size may lead to changes in behaviour that are then reciprocated and reinforced.

Three of the women with AN recruited into study 2 did not complete the training. This may represent resistance to treatment that might change their body perception (an interpretation

consistent with the lack of statistical significance of the intention to treat analysis if these women are included). Alternatively, it is possible the requirement to attend the training on four consecutive days may have proved too difficult.

A key question is whether the reductions that we observed in psychological concerns over a month in the aAN participants are clinically meaningful. With respect to EDE-Q, Bardone-Cone et al. (2010) operationalize recovery in eating disordered patients as a reduction in all four sub-scale scores to within 1 *SD* of age-matched community norms. Mond et al. (2006) report such norms for the age group 23 to 27 years based on a sample of 908 women: EDE-Q $M = 1.56$, $SD = 1.26$; EDE-restraint $M = 1.34$, $SD = 1.39$; EDE-eating concerns $M = 0.81$, $SD = 1.10$; EDE-shape concerns $M = 1.84$, $SD = 1.50$; EDE-weight concerns $M = 2.24$, $SD = 1.61$. The mean EDE-Q and sub-scale scores from our aAN sample (mean age 26.8 years) at Day-30, were respectively: EDE-Q $M = 2.63$; EDE-restraint $M = 2.45$; EDE-eating concerns $M = 2.06$; EDE-shape concerns $M = 2.83$; EDE-weight concerns $M = 3.18$. Therefore, with the exception of EDE-eating concerns, which missed the criterion by only 0.15 units, the perceptual training regime reported here produced reductions in EDE-Q scores that were clinically meaningful, when defined in this way.

Although the statistical results of these two exploratory studies should be treated with caution with respect to clinical outcomes due to low power, they indicate robust perceptual effects that suggest an exciting new way of treating biases in the judgement of female body size processing characteristic of AN, and provide evidence that perceptual biases play a causal role in AN as has been suggested for some other disorders (e.g. depression, Harmer et al. 2009). As such, perceptual training has the potential to be a valuable cost-effective adjunctive treatment for AN used together with more traditional talking therapies (cognitive behavioural therapy, mindfulness etc.). The training program could be given at 6 monthly intervals to reinforce a normalised body perception, as our research has suggested that body image concerns increase as BMI increases during treatment and potentially could undermine recovery (Cornelissen et al., 2015). This “top-up” training could be accomplished by adapting the body training for use through a mobile friendly website to run on a patient’s PC, laptop, tablet or smartphone or through a downloadable app. Additionally, this program could be used in non-clinical groups. Body dissatisfaction is widespread amongst girls and young women in western countries, where around 50% report being dissatisfied with their bodies and this dissatisfaction is a key predictor of the development of low self-esteem, depression, and eating disorders (e.g. Bearman, Presnell & Martinez, 2006; Monteath & McCabe, 1997). A downloadable form of the training could be used to improve body image judgements in this general population. However, a limitation of the current studies for clinical use is that the training has only been tested on a single cohort of women with AN. A larger scale randomised control trial is needed to properly test its feasibility and effectiveness in a clinical environment.

References

- Adams, S., Penton-Voak, I.S., Harmer, C.J., Holmes, E.A. & Munafò, M.R. (2013). Effects of emotion recognition training on mood among individuals with high levels of depressive symptoms: study protocol for a randomized controlled trial. *Trials*, 14, 161. <http://dx.doi.org/10.1186/1745-6215-14-161>.
- 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, e0139177. <http://dx.doi.org/10.1371/journal.pone.0139177>
- Bardone-Cone, A.M., Harney, M.B., Maldonado, C.R., Lawson, M.A., Robinson, D.P, Smith, R & Tosh, A. (2010). Defining Recovery from an Eating Disorder: Conceptualization, Validation, and Examination of Psychosocial Functioning and Psychiatric Comorbidity. *Behaviour Research and Therapy*, 48, 194–202. <http://dx.doi.org/10.1016/j.brat.2009.11.001>.
- Bearman, S. K., Presnell, K., & Martinez, E. (2006). The skinny on body dissatisfaction: A longitudinal study of adolescent girls and boys. *Journal of Adolescence*, 35, 217–229. <http://dx.doi.org/10.1007/s10964-005-9010-9>
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J. & Erbaugh J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561–571.
- Berkman, N.D., Lohr, K.N. & Bulik, C.M. (2007). Outcomes of eating disorders: a systematic review of the literature. *International Journal of Eating Disorders*, 40, 293-309. <http://dx.doi.org/10.1002/eat.20369>
- Carter, J.C., Mercer-Lynn, K.B., Norwood, S.J., Bewell-Weiss, C.V., Crosby, R.D., Woodside, D.B. & Olmsted, M.P. (2012). A prospective study of predictors of relapse in anorexia nervosa: Implications for relapse prevention. *Psychiatry Research*, 200, 518–523. <http://dx.doi.org/10.1016/j.psychres.2012.04.037>

Cash, T.F. & Deagle, E.A. (1997). The nature and extent of body-image disturbances in anorexia nervosa and bulimia nervosa: A meta-analysis. *International Journal of Eating Disorders*, 22, 107-126. [http://dx.doi.org/10.1002/\(SICI\)1098-108X\(199709\)22:2<107](http://dx.doi.org/10.1002/(SICI)1098-108X(199709)22:2<107)

Channon, S. & DeSilva, W. (1985). Psychological correlates of weight gain in patients with anorexia nervosa. *Journal of Psychiatric Research*, 19, 267-271.

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

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

Cornelissen, K.K., Gledhill, L., Cornelissen, P.L. & Tovée MJ. (2016). Visual biases in judging body weight. *British Journal of Health Psychology*, 21, 555-569, <http://dx.doi.org/10.1111/bjhp.12185>

Delinsky, S.S. & Wilson, G.T. (2006), Mirror exposure for the treatment of body image disturbance. *International Journal of Eating Disorders*, 39,108–116. <http://dx.doi.org/10.1002/eat.20207>

Evans, C. & Dolan, B. (1993). Body Shape Questionnaire: derivation of shortened “alternate forms”. *International Journal of Eating Disorders*, 13, 315-321. [http://dx.doi.org/10.1002/1098-108X\(199304\)13:3<315::AID-EAT2260130310>3.0.CO;2-3](http://dx.doi.org/10.1002/1098-108X(199304)13:3<315::AID-EAT2260130310>3.0.CO;2-3)

Fairburn, C.G. & Beglin, S.J. (1994). Assessment of eating disorders: Inter-view or self-report questionnaire? *International Journal of Eating Disorders*, 16, 363–370. [http://dx.doi.org/10.1002/1098-108X\(199412\)16:4<363::AID-EAT2260160405>3.0.CO;2](http://dx.doi.org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2).

Fairburn, C.G., Cooper, Z. & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A “transdiagnostic” theory and treatment. *Behaviour Research and Therapy*, 41, 509-528. [http://dx.doi.org/10.1016/S0005-7967\(02\)00088-8](http://dx.doi.org/10.1016/S0005-7967(02)00088-8)

- Farrell, C., Shafran, R. & Lee, M. (2006). Empirically evaluated treatments for body image disturbance: a review. *European Eating Disorders Review*, 14, 5, 289-300. <http://dx.doi.org/10.1002/erv.693>
- Gardner, R.M., Jones, L.C. & Bokenkamp, E.D. (1996). Comparison of three psychophysical techniques for estimating body-size perception. *Perceptual and Motor Skills*, 80, 1379-1390. <http://dx.doi.org/10.2466/pms.1995.80.3c.1379>
- George, H.R., Cornelissen, P.L., Hancock, P.J., Kiviniemi, V.V. & Tovée, M.J. (2011). Differences in eye-movement patterns between anorexic and control observers when judging body size and attractiveness. *British Journal of Psychology*, 102, 340-354. <http://dx.doi.org/10.1348/000712610X52429>
- Ginis, K.A.M. & Bassett, R.L. (2011). Exercise and changes in body image. In T.F. Cash & L. Smolak (Ed.s), *Body Image: A Handbook of Science, Practice, and Prevention* (pp. 415-423). Guilford Press: New York.
- Grabe, S., Ward, L.M. & Hyde, J.S. (2008). The role of the media in body image concerns among women: a meta-analysis of experimental and correlational studies. *Psychology Bulletin*, 134, 460-76. <http://dx.doi.org/10.1037/0033-2909.134.3.460>
- Harmer, C.J., Goodwin, G.M. & Cowen, P.J. (2009). Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *The British Journal of Psychiatry*, 195, 102-108. <http://dx.doi.org/10.1192/bjp.bp.108.051193>
- Irving, L. M. & Berel, S. R. (2001). Comparison of media-literacy programs to strengthen collage women's resistance to media images. *Psychology of Women Quarterly*, 25, 103-111. <http://dx.doi.org/10.1111/1471-6402.00012>
- Jarry, J.L., & Cash, T.F. (2011). Cognitive-behavioral approaches to body image change. In T.F. Cash & L. Smolak (Eds.), *Body Image: A Handbook of Science, Practice, and Prevention* (pp. 415-423). New York: Guilford Press.
- Key, A., George, C.L., Beattie, D., Stammers, K., Lacey, H. & Glenn Waller, G. (2002). Body image treatment within an inpatient program for anorexia nervosa: The role of mirror exposure in

the desensitization process. *International Journal of Eating Disorders*, 31, 185-190.

<http://dx.doi.org/10.1002/eat.10027>

Martin, K.A., & Lichtenberger, C.M. (2002). Fitness enhancement and changes in body image. In T. F. Cash & T. Pruzinsky (Eds.), *A handbook of theory, research, and clinical practice* (pp. 414-421). New York: The Guilford Press.

Mond, J.M., Hay, P.J., Rodgers, B. & Owen, C. (2006). Eating Disorder Examination Questionnaire (EDE-Q): Norms for young adult women. *Behavior Research Therapy*, 44, 53–62. <http://dx.doi.org/10.1016/j.brat.2004.12.003>.

Monteath, S. A., & McCabe, M. P. (1997). The influence of societal factors on the female body image. *Journal of Social Psychology*, 137, 708 –727. . <http://dx.doi.org/10.1080/00224549709595493>.

O'Dea, J. (2004). Evidence for a self-esteem approach in the prevention of body image and eating problems among children and adolescents. *Eating Disorders*, 12, 225-239. . <http://dx.doi.org/10.1093/heapro/dan004>

O’Dea, J.A. & Yager, Z. . (2011). School-based Psychoeducational Approaches to Prevention . In Thomas F. Cash and Linda Smolak (Eds.), *Body Image: A Handbook of Science, Practice, and Prevention*, (2nd ed., pp. 434–441), New York, USA: Guilford.

Penton-Voak, I.S., Bate, H., Lewis, G. & Munafò, M.R. (2012). Effects of emotion perception training on mood in undergraduate students: randomised controlled trial. *British Journal of Psychiatry*, 201, 71-2. <http://dx.doi.org/10.1192/bjp.bp.111.107086>

Penton-Voak, I.S., Thomas, J., Gage, S.H., McMurrin, M., McDonald, S. & Munafò, M.R. (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychological Science*, 24, 688-697. <http://dx.doi.org/10.1177/0956797612459657>

Probst, M., Vandereycken W, Vanderlinden J, Van Coppenolle H. (1998). The significance of body size estimation in eating disorders: Its relationship with clinical and psychological variables. *International Journal of Eating Disorders*, 24, 167-174. [http://dx.doi.org/10.1002/\(SICI\)1098-108X\(199809\)24:2<167::AID-EAT6>3.0.CO;2-C](http://dx.doi.org/10.1002/(SICI)1098-108X(199809)24:2<167::AID-EAT6>3.0.CO;2-C)

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

Tovée, M.J., Emery, J.L. & Cohen-Tovée, E.M. (2000). The estimation of body mass index and physical attractiveness is dependent on the observer's own body mass index. *Proceedings of the Royal Society B: Biological Sciences*, 267, 1987–1997. <http://dx.doi.org/10.1098/rspb.2000.1240>

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

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

Treasure, J., Claudino, A.M. & Zucker, N. (2010). Eating disorders. *Lancet*, 375, 583–593.
[http://dx.doi.org/10.1016/S0140-6736\(09\)61748-7](http://dx.doi.org/10.1016/S0140-6736(09)61748-7).

Vocks, S., Wachter, A., Wucherer, M. & Kosfelder, J. (2008). Look at yourself: Can body image therapy affect the cognitive and emotional response to seeing oneself in the mirror in eating disorders? *European Eating Disorders Review*, 16, 147–154. <http://dx.doi.org/10.1002/erv.825>

Williamson, D.A., Cubic, B.A. & Gleaves, D.H. (1993). Equivalence of body image disturbances in anorexia and bulimia nervosa. *Journal of Abnormal Psychology*, 102, 177-180.
<http://dx.doi.org/10.1037/0021-843X.102.1.177>

Table 1: Demographic and questionnaire data from the participants in Study 1.

| | Intervention (n=20) | | Control (n=20) | |
|---------------|---------------------|-----------|----------------|-----------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| Age (years) | 18.20 | 0.37 | 19.00 | 1.27 |
| BMI | 25.40 | 5.12 | 23.90 | 3.01 |
| Screening BSQ | 68.20 | 7.25 | 73.00 | 11.52 |
| EDE-Q | 3.09 | 1.01 | 3.67 | 0.83 |
| BDI | 16.70 | 9.04 | 17.80 | 12.01 |
| RSE | 16.00 | 4.83 | 13.90 | 4.45 |

Note. BMI = Body Mass Index. BSQ = Body Shape Questionnaire. BDI = Beck Depression Inventory. EDE-Q = Eating Disorder Examination Questionnaire global score. BDI = Beck Depression Inventory. RSE = Rosenberg Self Esteem scale.

Table 2: Summary table of the psychological scores from study 1.

| Measure | Test day | Intervention | Control | Difference Z-score | 95% CI Z-score | p-value |
|---------------------|----------|------------------------|------------------------|-----------------------|-------------------|---------|
| | | Raw score Mean (SD) | Raw score Mean (SD) | | | |
| EDE-Q | 1 | 3.09 (1.01) | 3.67 (0.83) | 0.56 | -0.012 – 1.13 | .06 |
| | 4 | 2.81 (0.97) | 3.77 (0.82) | 0.93 | 0.37 – 1.49 | .002 |
| | 14 | 2.64 (1.08) | 3.73 (0.91) | 1.04 | 0.44 – 1.68 | .001 |
| EDE-restraint | 1 | 2.55 (1.31) | 3.36 (1.41) | 0.58 | -0.04 – 1.21 | .07 |
| | 4 | 2.31 (1.23) | 3.41 (1.31) | 0.79 | 0.20 – 1.38 | .01 |
| | 14 | 2.13 (1.18) | 3.41 (1.38) | 0.92 | 0.33 – 1.51 | .003 |
| EDE-eating concerns | 1 | 2.30 (1.43) | 3.00 (1.54) | 0.50 | -0.18 – 1.18 | .14 |
| | 4 | 1.94 (1.19) | 2.81 (1.15) | 0.62 | 0.09 – 1.16 | .12 |
| | 14 | 1.82 (1.37) | 2.59 (1.44) | 0.55 | -0.09 – 1.19 | .09 |
| EDE-shape concerns | 1 | 3.42 (1.13) | 4.13 (1.00) | 0.63 | 0.03 – 1.23 | .04 |
| | 4 | 3.29 (1.19) | 4.21 (0.78) | 0.81 | 0.24 – 1.38 | .006 |
| | 14 | 3.00 (1.24) | 4.18 (0.83) | 1.04 | 0.45 – 1.64 | .001 |
| EDE-weight concerns | 1 | 4.08 (1.00) | 4.18 (0.81) | 0.09 | -0.50 – 0.68 | .75 |
| | 4 | 3.70 (1.02) | 4.66 (0.82) | 0.97 | 0.37 – 1.57 | .002 |
| | 14 | 3.59 (0.94) | 4.73 (0.81) | 1.15 | 0.58 – 1.72 | .0002 |

Note. EDE-Q = Eating Disorder Examination Questionnaire global score. EDE-Q res = Eating Disorder Examination Questionnaire eating restraint subscale. EDE-Q eat = Eating Disorder Examination Questionnaire eating concern subscale. EDE-Q sc = Eating Disorder Examination Questionnaire body shape concern subscale. EDE-Q wc = Eating Disorder Examination Questionnaire weight concern subscale.

Table 3: Summary table of the psychological scores on different test days, and the statistical comparisons between them, from study 2.

| Measure | Test day | Mean (SD) (raw score) | Comparison Point | Difference (z-score) | 95% CI (z-score) | p-value |
|---------------------|----------|--------------------------|---------------------|-------------------------|---------------------|---------|
| EDE-Q | 1 | 3.43 (1.64) | Day 1 vs 4 | 0.65 | 0.13 – 1.18 | .02 |
| | 4 | 2.69 (1.74) | Day 1 vs 7 | 0.57 | 0.02 – 1.12 | .04 |
| | 7 | 2.91 (1.84) | Day 1 vs 30 | 0.74 | 0.20 – 1.28 | .008 |
| | 30 | 2.63 (1.68) | | | | |
| EDE-restraint | 1 | 3.16 (1.69) | Day 1 vs 4 | 0.59 | -0.07 – 1.25 | .08 |
| | 4 | 2.50 (1.77) | Day 1 vs 7 | 0.65 | -0.02 – 1.32 | .06 |
| | 7 | 2.44 (1.87) | Day 1 vs 30 | 0.66 | -0.004 – 1.33 | .05 |
| | 30 | 2.45 (1.76) | | | | |
| EDE-eating concerns | 1 | 2.89 (2.09) | Day 1 vs 4 | 0.79 | 0.14 – 1.43 | .02 |
| | 4 | 2.03 (1.86) | Day 1 vs 7 | 0.66 | -0.03 – 1.35 | .06 |
| | 7 | 2.37 (2.00) | Day 1 vs 30 | 0.84 | 0.22 – 1.47 | .01 |
| | 30 | 2.06 (1.69) | | | | |
| EDE-shape concerns | 1 | 4.08 (1.47) | Day 1 vs 4 | 0.73 | 0.14 – 1.32 | .02 |
| | 4 | 3.15 (1.77) | Day 1 vs 7 | 0.59 | -0.06 – 1.25 | .07 |
| | 7 | 3.45 (1.95) | Day 1 vs 30 | 0.76 | 0.15 – 1.36 | .01 |
| | 30 | 3.18 (1.77) | | | | |
| EDE-weight concerns | 1 | 3.58 (1.97) | Day 1 vs 4 | 0.38 | -0.15 – 0.92 | .16 |
| | 4 | 3.10 (2.11) | Day 1 vs 7 | 0.33 | -0.23 – 0.89 | .25 |
| | 7 | 3.26 (2.16) | Day 1 vs 30 | 0.54 | 0.001 – 1.09 | .05 |
| | 30 | 2.83 (1.87) | | | | |

Note. EDE-Q = Eating Disorder Examination Questionnaire global score. EDE-Q res = Eating Disorder Examination Questionnaire eating restraint subscale. EDE-Q eat = Eating Disorder Examination Questionnaire eating concern subscale. EDE-Q sc = Eating Disorder Examination Questionnaire body shape concern subscale. EDE-Q wc = Eating Disorder Examination Questionnaire weight concern subscale.



Figure 1: The middle row shows part of the body sequence varying in BMI. The top row illustrates the results from a baseline assessment and the position of the categorical boundary prior to training. The bottom row illustrates the results from the post-training test session, showing that the categorical boundary has shifted relative to the pre-training result.

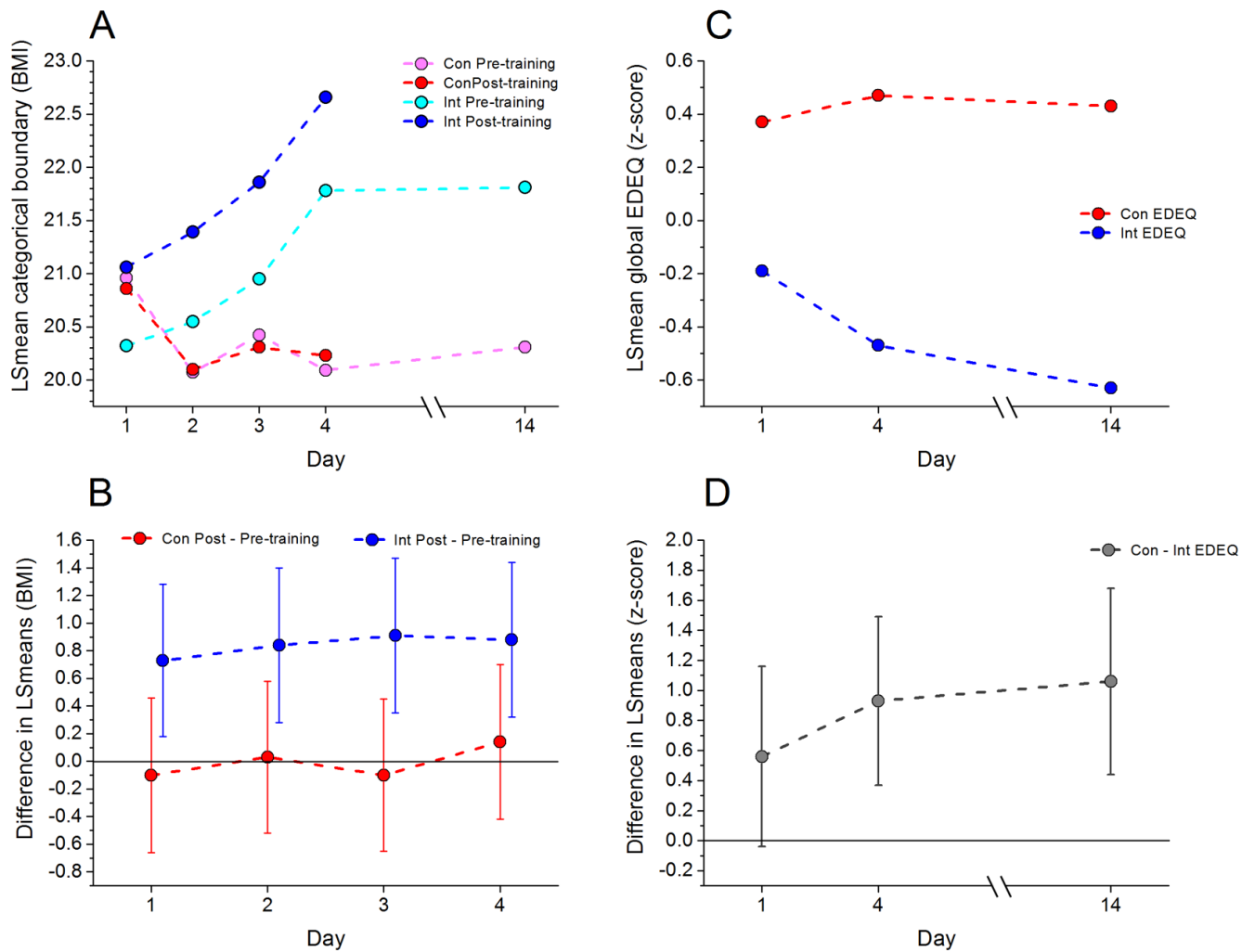


Figure 2: A) A plot of the mean value of BMI at the categorical boundary, predicted from the multi-level model as a function of measurement day. Magenta and red circles represent control group pre- and post-training thresholds respectively. Cyan and blue circles represent intervention group pre- and post-training thresholds respectively. B) A plot of the predicted differences between pre- and post-training categorical threshold, with 95% CIs, as a function of training day. Confidence intervals that straddle zero are not significant at $p < .05$. Blue circles represent the intervention group and red circles the control group. C) A plot of predicted global EDE-Q z-scores as a function of measurement day. Blue circles represent the intervention group and red circles the control group. D) A plot of the predicted differences in global EDE-Q z-scores between the control and training groups as a function of measurement day, with 95% CIs, as a function of training day. Confidence intervals that straddle zero are not significant at $p < .05$.

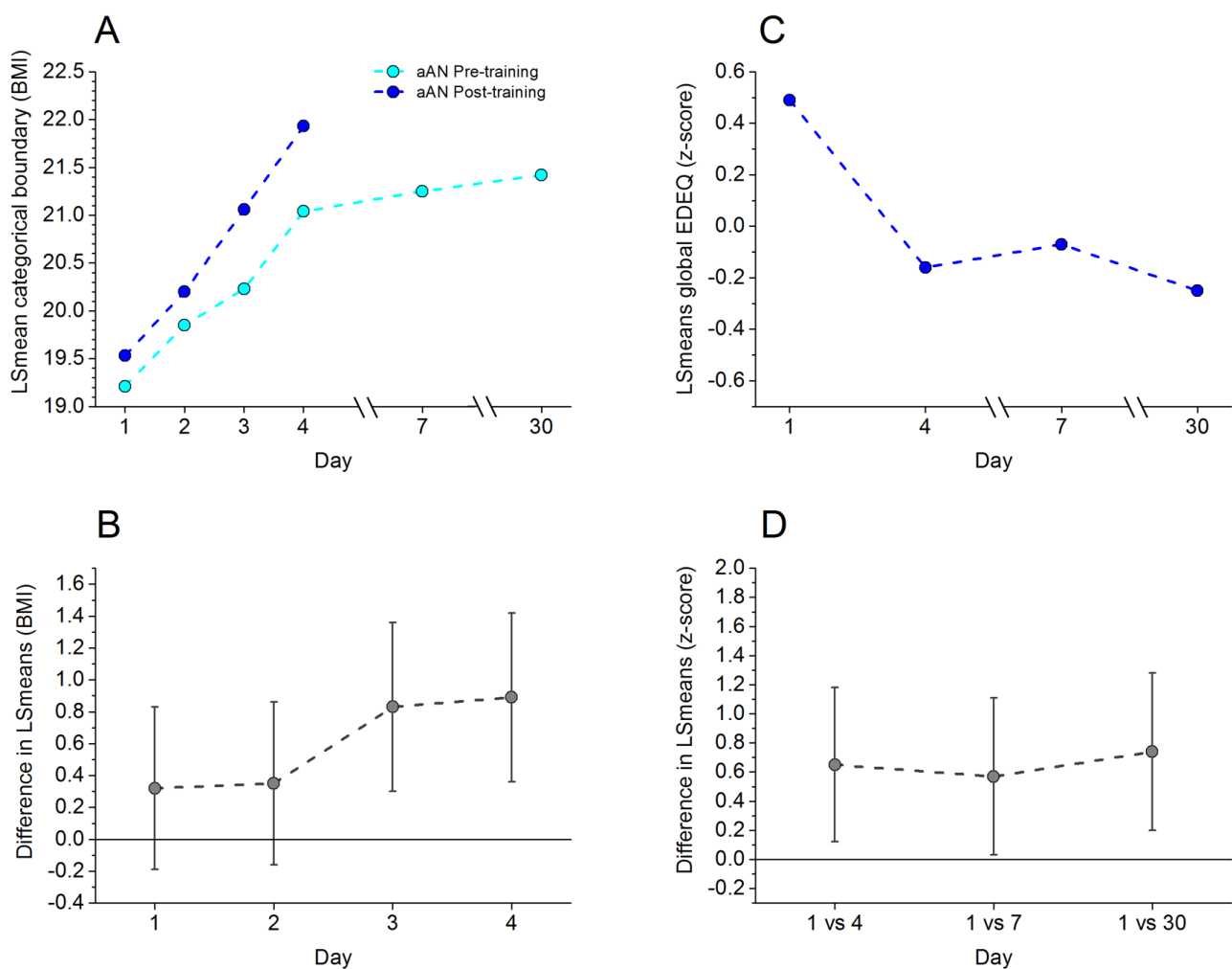


Figure 3: A) A plot of the mean value of BMI at the categorical boundary, predicted from the multi-level model as a function of measurement day. Cyan and blue circles represent pre- and post-training thresholds respectively for the aAN participants in experiment 2. B) A plot of the predicted differences between pre- and post-training categorical threshold, with 95% CIs, as a function of training day. Confidence intervals that straddle zero are not significant at $p < .05$. C) A plot of the mean predicted global EDE-Q z-scores as a function of measurement day for aAN participants in experiment 2. D) A plot of the predicted differences in global EDE-Q z-scores between the baseline measurement on day 1 and the post-training measurements on days 4, 7 and 30, with 95% CIs. Confidence intervals that straddle zero are not significant at $p < .05$.