**Experimental analysis of the impact of body comparison on non-clinical women**

**Abstract**

**Background and Objectives:** Two experimental studies examined the impact that body comparison has on women’s body satisfaction and self-esteem. The two studies differed in the use of a human comparator or an avatar (non-human) comparator. The independent variables were the type of body comparison (upward, downward and neutral) and the perceived personality of the comparator.

**Methods:** Each study used a within-participant design. Participants compared themselves to images of women or avatars (previously rated as ‘attractive’, ‘neutral’ or ‘unattractive’), with an accompanying ‘positive’ or negative’ personality descriptor. Participants rated their body satisfaction and self-esteem after each image.

**Results:** Upward comparison resulted in negative effects for the participant compared to downward and neutral conditions, but downward comparison’s impact was only found for human images. The described personality of the image had a more complex impact when using human images, but was still relevant for avatars.

**Limitations:** The sample for this study was lacking in diversity. Compliance within the experimental tasks was not strictly monitored. The use of a within-subject design might have allowed some participants to deduce the nature of the study.

**Conclusions:** These findings provide causal evidence for the theorised effects of body comparison on body satisfaction self-esteem. Upward comparison has especially problematic outcomes, even causing negative effects when comparing to a non-human avatar. Downward comparison had positive effects, but only for human images. Furthermore, personality played some moderating role in these outcomes. Social comparison theory needs to be reviewed in light of these findings, and practical implications are discussed.

**Keywords:** body image, safety behaviors, self-esteem, body comparison, eating disorders

**Experimental analysis of the impact of body comparison on non-clinical women**

* 1. **Introduction**

Body image and eating pathology are influenced by a number of safety behaviours. Such behaviours include body avoidance, checking and comparison. Each of these behaviours are hypothesised to reduce anxiety and related pathology in the short term (e.g., checking on the scales makes the person more confident that they have not gained weight at the time) but to worsen such pathology in the longer term (e.g., the person fears that the scales might be wrong and that they have actually gained weight, so have to keep checking). The causal impact of body checking and avoidance are well understood (e.g., Bailey & Waller, 2017; Shafran, Fairburn, Robinson & Lask. 2004), with each resulting in a worsening of body and eating concerns. However, the effects of body comparison are less well understood, and require experimental study to demonstrate the impact of this behaviour on body image and eating pathology.

Social comparison theory (Festinger, 1954) states that comparison is a normal human behaviour, used to assess our standing in life. However, when comparison becomes excessive, or out of context, more negative consequences can occur. Body comparison involves considering oneself in relation to others’ appearance and related characteristics (Laker & Waller, 2020), and takes two forms. Upward comparison occurs when the object of comparison is perceived as better than oneself (providing an aspiration figure, but impairing self-esteem in the short and long term), while downward comparison involves comparing oneself with a less positive person (enhancing self-esteem in the short term, at least). Two factors need to be borne in mind when understanding the likely impact of body comparison in everyday life. First, women tend to overestimate their own size and hold negative self-perceptions (e.g., Cattarin, Thompson, Thomas & Williams, 2000; Halliwell & Dittimar, 2005; Slade, 1995). Second, the western ‘thin-ideal’ (Delaney, O’Keefe & Skene, 1997) makes it more likely that women will focus on comparing themselves with more socially-valued, slimmer role models (e.g., on social media). Taking these biases into account, it is more likely that women will engage in upward comparison, making comparison more likely to have negative effects for them.

Correlational studies support the hypothesised relationship of body comparison with body dissatisfaction and eating pathology (Arigo, Schumacher & Martin, 2014; Arroyo, 2014; Cattarin et al., 2000; Leahey, Crowther & Ciesla, 2011). However, they do not show causal impact of these two forms of comparison, and commonly fail to differentiate upward from downward comparison. Furthermore, some ‘comparison’ studies only require the individual to expose themselves to images of other people, rather than requiring them to engage in active comparison (e.g., Bailey & Ricciardelli, 2010; Mills, Polivy, Herman & Tiggemann, 2002; Posavac, Posavac & Posavac, 1998). To summarise, social comparison theory makes different predictions about the outcomes of upward and downward body comparisons, but the literature to support these hypotheses is methodologically limited.

The few experimental body comparison studies that have been reported (Leahey, Crowther & Mickelson, 2007; Van den Berg & Thompson, 2007) support the proposal that upward comparison has negative outcomes, whereas downward comparison has positive results. However, two factors need further investigation for their causal role. First, the causal impact of appearance has been studied in isolation. Factor analysis (e.g., Laker & Waller, 2020) has shown that comparison involves two elements – body and personality. It is possible that body comparison’s impact will be influenced by other aspects of the individual with whom one is comparing oneself.

Second, body comparison has been studied in respect to images of real people. Therefore, it is possible that the impact of body comparison experiments is influenced by exposure to other bodies, rather than being the result of the comparison process per se. However, many representations are non-human (e.g., emojis on social media; avatars in games), and such representations could be hypothesised to have an impact on self-perception. For example, Daiini (2018) and Seung (2010) have each shown that avatars can become linked to self-identity. In comparison conditions, it could be argued that the less realistic the representation, the less likely that any comparison will result in positive or negative effects on the individual, as the comparison is less likely to allow the person to evaluate themselves against the comparator. Alternatively, it could be that the act of comparison with a more or less attractive representation is more important than the human nature of the image involved.

Given these two gaps in understanding of body comparison, two experimental studies are reported. Each compares the effect of upward and downward appearance and personality comparisons on self-esteem and body satisfaction. While the first uses human images, the second will determine whether the outcomes of comparison are similar or different when using non-human images (avatars, based on the same human images). It is hypothesised that upward body comparison against human images will have negative effects for the individual, while downward comparison will have positive effects. However, no hypothesised direction of effect is made regarding the impact of positive vs negative personality descriptors. Nor is any directional hypothesis made regarding the impact of avatars.

* 1. **Study 1 – Comparison with human images**

As outlined above, existing theory and research in the field of body comparison suggests that upward comparison should have negative effects on body image and self-esteem, while downward comparison should have positive effects (Leahey et al., 2007; Lin & Soby, 2016; Van den Berg & Thompson, 2007). This experimental study will determine whether that effect can be replicated, and whether the identified personality of the comparison figure impacts on that effect.

**2.2 Method**

**2.2.1 Ethics**

This study was reviewed and approved by the Research Ethics Committee of the Psychology Department of the University.

**2.2.2 Design**

This study used a within-subjects experimental design. The dependent variables were self-esteem and body satisfaction. The independent variables were three different types of physical comparison (upward, downward, neutral), and two types of personality descriptor (positive, negative) for the person being used as the comparator. Each participant completed six comparisons, presented in random order.

**2.2.3 Participants**

A sample of 40 adult females completed the study (mean age = 19.73 years; *SD* = 3.40; range = 18-37 years). The participants were students recruited from a Psychology course, and received credits for participation. All participants were asked if they currently had an eating disorder or had previously had an eating disorder. None disclosed that they had.

 A priori sample size analysis was carried out using G\*Power (v.3.1.9.2), focusing on the key repeated measures ANOVA factor of condition (type of comparison: three levels), as G\*Power does not generate sample size calculations based on multiple within-subject measures. The assumed effect size was f = 0.39, based on the Bailey & Waller (2017) effect size of partial eta2 = 0.129 (equivalent to *d* = 0.770 and *f* = 0.39). With that effect size, a power of 85%, and alpha = .05. a total sample of 39 was shown to be needed. Thus, with 40 participants, the study was adequately powered to detect medium to large effects on the key main within-subject condition of types of comparison. However, the sample size was also checked (for the same analysis, with identical assumptions of power and alpha) using the Cohen’s *d* criterion in G\*Power, which indicated that this sample would only be suitable to detect larger effect sizes (f = 0.54).

**2.2.4 Measures**

Single-item measures of body satisfaction and self-esteem were used to assess the immediate of the intervention on these state variables. Following each of the six experimental conditions, participants were asked five questions. Each was rated on a scale of 0-100 (0 = very negative; 100 = very positive), using a slider on the online presentation. The two dependent variables were questions related to self-esteem (‘What do you think about yourself right now?’), and to body satisfaction (‘What do you think about your own body right now?’). These were interspersed among three other questions (asking: how likely it was that they would be the person’s friend; how easy it would be to like the person; how good would it be to work with the person), in order to hide the true focus of the study.

**2.2.5 Procedure**

Participants viewed a series of six photographs of adult women (see below) - two attractive, two neutral, and two unattractive. These photographs are available in the supplementary material. To ensure that the participants made a comparison to each picture, they were asked to list the first three things about the person in each picture that they would compare themselves with.

**2.2.5.1 Development of the stimuli.** The authors selected a pool of photographs available from the internet, demonstrating a wide range of body types and appearances. These 105 pictures were rated by 22 female participants (mean age = 23.73, *SD* = 6.18), who did not take part in the main study. The participants were recruited via a university volunteer participation scheme and personal contacts of the author. Participants were provided with a £10 Amazon voucher upon completion of the ratings. They were asked to rate each picture on how attractive they thought the person was (1 = very unattractive, 5 = averagely attractive, 9 = very attractive). The two highest-rated, middle-rated, and lowest-rated pictures were selected. Given the thin-ideal internalization in western cultures (Low et al., 2003), the images of the ‘rated attractive’ women were slim, and the images of the ‘unattractive’ women were larger. Thus, the comparator images were selected primarily as per the participants’ ratings. However, as it was possible that the very lowest-rated pictures had potential for fat-shaming, the final selection of those photographs was taken from those with a slightly higher rating. Fat-shaming comes from the notion ‘fat is bad’ (Ravary, Baldwin & Bartz, 2019), and an individual is judged/mocked for being overweight. The methodological issue with fat-shaming is that there might already be a bias towards negative assumptions about the individual in the photograph (both physically and as a person), meaning that the result measured would not be due to comparison per se. The final selection of photographs resulted in a pool of six pictures, which had been rated as the two most attractive, the two nearest to average attractiveness, and two that were rated as near to least attractive.

Once the six photographs were selected (please see supplementary material for these images), each was given a brief descriptor of the person involved. The description (please see supplementary material for these descriptions), detailed the job the person in the picture did along with two adjectives, which were either positive or negative (depending on whether the picture was in the ‘Positive’ or ‘Negative’ condition). For example, a ‘positive’ description of the person in the picture would be ‘Hayley works as a cleaner. She is often described as kind and funny’, while a ‘negative’ description would be ‘Anna works as a shop assistant. Her colleagues describe her as bossy and impatient’.

**2.2.5.2 Main study.** Participants were female undergraduate students, recruited via an online research participation system. They received credits for participation. Sixty-two individuals activated the link, and 40 completed the measures (completion rate = 64.52%). They were informed that the study was on how people compare themselves with others, and that only females could register. Once the participant consented, they were provided with a link to the experimental study, where they were shown the first image (order randomised), were asked to select three things about the image that they would compare with themselves (to ensure that they engaged in active comparison rather than simply being exposed to the images), and were asked to answer the five single-item questions. They then moved on to the remaining images (time determined by the participant). At the end, they were not debriefed regarding the purpose of the study, in case they wished to participate in Study 2.

**2.2.6 Data analysis**

For each of the five questions, scores were calculated by using the mean response for appearance and personality comparison. To calculate mean for appearance comparison, the scores for the pairs of ‘attractive’ images, ‘neutral’ images and ‘unattractive’ images were summed and divided by two. To calculate the mean for personality comparison, the scores were summed for each ‘positive’ and ‘negative’ personality and divided by three. A two-way repeated measures 3 x 2 ANOVA was used to determine the impact of the three physical types (upward comparison relative to attractive images; downward comparison relative to unattractive images; and neutral comparison relative to the pictures that were rated as neither attractive nor unattractive) and of the two personality conditions (positive and negative) on the dependent variables of body image and self-esteem. Effect sizes were reported as partial *eta2*, where a score of ≥ 0.14 indicates a large effect, between 0.06 and 0.13 indicates a medium effect, and anything below that indicates a small effect (Cohen, 1992). Least significant difference (LSD) tests were used to compare the effects within factors. A series of paired *t*-tests were used to interpret the interaction of physical comparison and personality comparison for body image, reporting Cohen’s *d* as the effect size (calculated with correction for the correlation between measures).

**2.3 Results**

Table 1 shows levels of body satisfaction and self-esteem following upward comparison, downward comparison, and the neutral condition. The main and interaction terms are addressed below.

**2.3.1 Impact of appearance comparison on body satisfaction**

 The 3 x 2 ANOVA is detailed in Table 1. The main effects and interaction term are outlined below.

**2.3.1.1 Effect of comparison.** The repeated measures ANOVA showed that there was a significant main effect of direction of appearance comparison (*F* = 21.2; *p* < .001), with a very large effect size (partial eta2 = 0.53). All three conditions were significantly different from each other (LSD tests). Collapsing the two ‘personality’ levels, upward comparison was yielded a mean score of 38.36, the Neutral condition had a mean score of 52.18, and downward comparison yielded a mean score of 61.01. Thus, compared to the neutral condition, downward appearance comparison resulted in significantly more positive body satisfaction, while upward comparison resulted in significantly poorer body satisfaction.

**2.3.1.2 Effect of personality.** The ANOVA also showed that there was a significant effect of personality (*F* = 5.30; *p* < .05; partial eta2 = 0.12). Collapsing the appearance scores, self-comparison to people with ‘Positive’ or ‘Negative’ personality descriptors resulted in mean scores of 52.02 vs 49.00 respectively. Thus, the women had moderately better body satisfaction following comparison with a person with a more positive personality.

Table 1

Mean scores and ANOVA results for the impact on body satisfaction and self-esteem of upward and downward appearance comparison and positive and negative personality traits

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Condition** |  |  |  |  |  | **ANOVA** |
| **Appearance comparison** | Up | Neutral | Down |  | Direction of comparison for Appearance |  | Direction of comparison for Personality |  | Appearance x Personality |
| **Personality comparison** | Positive | Negative | Positive | Negative | Positive | Negative | *F* | *P* | df | LSD | Partial eta² | *F* | *P* | df | LSD | Partial eta² | *F* | *P* | df | Partial eta² |
| Body satisfaction |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| M | 37.68 | 39.03 | 54.75 | 49.60 | 63.63 | 58.38 | 21.2 | .001 | 2,38 | D>N>U | 0.53 | 5.30 | .05 | 1,39 | P>N | .12 | 5.20 | .01 | 2,38 | .21 |
| *(SD)* | (22.84) | (21.84) | (17.80) | (21.25) | (19.02) | (19.03) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Self -esteem |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| M | 47.98 | 47.58 | 57.68 | 58.08 | 62.48 | 63.73 | 15.2 | .001 | 2,38 | D>N>U | 0.45 | 0.08 | *NS* | 1,39 | - | - | 0.26 | *NS* | 2,38 | - |
| *(SD)* | (23.19) | (22.35) | (18.97) | (18.34) | (18.02) | (18.90) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Questions: Body satisfaction - What do you think about your own body right now?; Self-esteem - ‘What do you think about yourself right now?’

LSD data: body satisfaction (appearance – attractive vs neutral), mean difference = -13.83, *SE =* 2.31, *P <* .001; body satisfaction (appearance – attractive vs unattractive), mean difference = -22.65, *SE =* 3.45, *P <* .001; body satisfaction (appearance – neutral vs unattractive), mean difference = -8.83, *SE =* 1.87, *P <* .001. Body satisfaction (personality – positive vs negative), mean difference = 3.02, *SE =* 1.32, *P <* .05.

Self-esteem (appearance – attractive vs neutral), mean difference = -10.10, *SE =* 1.90, *P <* .001; body satisfaction (appearance – attractive vs unattractive), mean difference = -15.33, *SE =* 3.02, *P <* .001; body satisfaction (appearance – neutral vs unattractive), mean difference = -5.23, *SE =* 2.09, *P <* .05. Body satisfaction (personality – positive vs negative), mean difference = -4.17, *SE =* 1.50, *NS*.

**2.3.1.3 Interaction of comparison and personality.** Finally, the ANOVA showed a significant interaction between appearance comparison and personality (*F* = 5.20, *p* < .01; partial eta2 = 0.21). The ‘personality’ means for upward appearance comparison were similar (‘Positive’ *M* = 37.68, ‘Negative’ *M* = 39.03), showing that body satisfaction does not change according to whether an attractive person has positive or negative personality characteristics. However, the neutral condition means were more divergent (‘Positive’ *M* = 54.75, ‘Negative’ *M* = 49.60), as were those for downward comparison (‘Positive’ *M* = 63.63, ‘Negative’ *M* = 58.38). Paired *t*-tests were used to interpret this interaction. There was a significant effect of personality type (i.e., ‘Positive’ or ‘Negative’) on body satisfaction, but only when making a downward appearance comparison (*t* = 3.89, *p* < .001, *d* = 0.62). This effect was not found when making an upward (*t* = 0.82, *NS, d* = 0.13) or neutral (*t* = 1.67, *NS, d* = 0.26) appearance comparison.

**2.3.1.4 Summary.** When comparing themselves to a person who is perceived as attractive, women feel worse about the way they look regardless of the perceived personality of the person. However, individuals feel better about the way they look if they compare with someone seen as less attractive than themselves, and feel even better about their body if the less attractive person has a positive personality.

**2.3.2 Impact of appearance comparison on self-esteem**

 The ANOVA is detailed in Table 1. The main effects and interaction terms are outlined below.

**2.3.2.1 Effect of comparison.** There was a significant main effect of direction of appearance comparison (*F* = 15.2; *p* < .001) with a very large effect size (partial eta² = 0.45). All three conditions were significantly different from each other (LSD tests). Collapsing the two ‘personality’ levels, upward comparison was associated with a mean score of 47.78, the neutral condition yielded a mean score of 57.88, and the downward comparison condition yielded a mean score of 63.11, and all three conditions were significantly different. Thus, compared to the neutral condition, downward appearance comparison resulted in significantly better self-esteem, while upward comparison resulted in significantly poorer self-esteem.

**2.3.2.2 Effect of personality.** There was no significant effect of personality (*F* = .08; *NS*). Collapsing the appearance scores, self-comparison to people with ‘Positive’ personality descriptors resulted in a mean score of 56.05, and comparison to ‘Negative’ personality descriptors resulted in a mean score of 56.46. Therefore, the image being paired with a positive or negative personality did not impact the participants’ self-esteem.

**2.3.2.3 Interaction of comparison and personality.** Finally, there was no significant interaction between appearance comparison and personality (*F* = 0.26, *NS*). Thus, the personality of the comparator does not alter how one feels about one’s self when comparing with attractive or unattractive people.

**2.3.2.4 Summary.** There is a large effect of appearance comparison on self-esteem, which is not affected by the ‘personality’ of the comparison person.

**2.4 Discussion**

The aim of this first study was to establish whether upward and downward comparisons drive how an individual feels about their body image and self-esteem, and to determine whether any such impact was influenced by the personality of the comparator person. The outcome was that upward and downward body comparison had the anticipated outcomes, respectively worsening and enhancing body satisfaction and self-esteem in comparison to each other, and a neutral condition. However, while there was an impact of the personality of the comparator image, that only applied in the case of downward comparison, where a more positive personality enhanced body satisfaction among those who saw a less attractive image. The implications of these findings will be considered following Study 2.

**3.1 Study 2 – Comparison with non-human images**

 Given the outcomes above, it is evident that upward and downward body comparison with images of real people images results in the outcomes that are generally predicted by Social Comparison Theory (Festinger, 1954). However, as outlined earlier, this leaves the question of whether this is a result of the comparison process, or an impact of exposure to such images. Therefore, this study will consider whether comparison to non-human images (avatars, based on more or less attractive people) has a similar effect or not. Again, the ‘personality’ of the comparator will be considered, to determine whether it interacts with the attractiveness of the avatar.

**3.2 Method**

**3.2.1 Ethics**

This study was reviewed and approved by the Research Ethics Committee of the Psychology Department of the University.

**3.2.2 Design**

To ensure comparability, this online study used the same within-subjects experimental design as Study 1. Again, each person completed six body comparisons that were presented in random order, varying in the attractiveness and personality of the image used, and the same outcomes were measured.

**3.2.3 Participants**

Of the 54 adult females who began the survey, 41 completed the measures (completion rate = 75.9%). Their mean age was 19.73 years; *SD* = 3.40; range = 18-37 years). Seventeen of these participants were volunteers from the above study who had agreed to participate in a second study, but they had not been debriefed after Study 1, so they were unaware of the true nature of this study. The intention had been to run within-subject effects between the two studies to determine whether there was a difference, but not enough participants completed both studies. All participants were asked if they currently had an eating disorder or has previously had an eating disorder. None disclosed that they had. The a priori sample size analysis from Study 1 was applied in this study, indicating that 39 participants would be sufficient. While the actual effect size in Study 1 was larger than this, the same medium effect size was assumed in case body comparison with an avatar would yield a less strong effect than a human figure. Thus, with 41 participants, the study was adequately powered to detect medium effects. The participants were students recruited from a Psychology course, and received credits for participation.

**3.2.4 Measures**

The same five single-item measures as in Study 1 were used to assess the impact of each of the six conditions on body satisfaction and self-esteem

**3.2.5 Procedure**

Following reading the information sheet and giving consent, the study began. Each participant saw six images of adult female human-like avatars - two attractive, two neutral, and two unattractive.

**3.2.5.1 Development of the stimuli.** Six photographs were selected from the 105 used in Study 1, using similarly scored pictures from the larger set (rather than same six as before). These six pictures were then converted into human-like avatars, using the avatar creator on The SIMS 4 computer game (Maxis, 2014). Short descriptions (similar to the above study) accompanied each picture. This resulted in a pool of six avatars, based on two attractive photos, two photos that were near to average attractiveness, and two photos that were rated as near to least attractive. The avatar images are available from the authors.

**3.2.5.2 Main study.** The procedure in the main part of the study was identical to that used in Study 1.

**3.2.6 Data analysis**

The data analysis was identical to that in Study 1. It involved a 3 (attractiveness) x 2 (personality) repeated measures ANOVAs with post hoc LSD tests, and paired t-tests used to interpret any interactions. The dependent variables were the individual’s body dissatisfaction and self-esteem.

**3.3 Results**

Table 2 shows levels of body satisfaction and self-esteem following upward comparison, downward comparison, and the neutral condition.

**3.3.1 Impact of appearance comparison on body satisfaction**

 The 3 x 2 ANOVA is detailed in Table 2. The main effects and interaction terms are outlined below.

Table 2

Mean scores and ANOVA results for the impact on body satisfaction and self-esteem of upward and downward appearance comparison and positive and negative personality traits

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Condition** |  |  |  |  |  | **ANOVA** |
| Appearance | Up | Neutral | Down |  | Direction of comparison for Appearance |  | Direction of comparison for Personality |  | Appearance x Personality |
| Personality | Positive | Negative | Positive | Negative | Positive | Negative | *F* | *P* | df | LSD | Partial eta² | *F* | *P* | df | LSD | Partial eta² | *F* | *P* | df | Partial eta² |
| Body satisfaction |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| M | 55.46 | 47.88 | 67.76 | 62.27 | 66.78 | 63.41 | 19.8 | .001 | 2,39 | U<N=D | .50 | 8.67 | .01 | 1,40 | P>N | .18 | 1.59 | *NS* | 2,39 | - |
| *(SD)* | *23.45* | *25.18* | *19.71* | *20.47* | *21.24* | *20.07* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Self -esteem |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| M | 62.56 | 61.24 | 68.32 | 68.80 | 69.41 | 69.80 | 6.38 | .01 | 2,39 | U<N=D | .25 | .01 | *NS* | 1,40 | - | - | .46 | *NS* | 2,39 | - |
| *(SD)* | *23.84* | *25.73* | *21.69* | *22.20* | *22.45* | *22.40* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

LSD data: body satisfaction (appearance – attractive vs neutral), mean difference = -13.34, *SE =* 2.10, *P <* .001; body satisfaction (appearance – attractive vs unattractive), mean difference = -13.43, *SE =* 2.36, *P <* .001; body satisfaction (appearance – neutral vs unattractive), mean difference = -.09, *SE =* 1.18, *NS*. Body satisfaction (personality – positive vs negative), mean difference = 5.48, *SE =* 1.86, *P <* .01.

Self-esteem (appearance – attractive vs neutral), mean difference = -6.66, *SE =* 1.84, *P <* .01; body satisfaction (appearance – attractive vs unattractive), mean difference = -7.71, *SE =* 2.46, *P <* .01; body satisfaction (appearance – neutral vs unattractive), mean difference = -1.05, *SE =* 1.36, *NS*. Body satisfaction (personality – positive vs negative), mean difference = .15, *SE =* 1.88, *NS*.

**3.3.1.1 Effect of comparison.** There was a significant main effect of direction of appearance comparison (*F* = 19.8; *p* < .001) on body satisfaction, with a very large effect size (partial eta2 = 0.50). The upward comparison condition was significantly different from the other two (LSD tests). Collapsing the two ‘personality’ levels, upward comparison was associated with a mean score of 51.67, the neutral condition had a mean score of 65.02, and the downward condition yielded a mean score of 65.10. Thus, compared to the other conditions, upward appearance comparison resulted in significantly poorer body satisfaction.

**3.3.1.2 Effect of personality.** There was a significant and large effect of personality (*F* = 8.67; *p* < .01; partial eta2 = 0.18). Collapsing the appearance scores, comparison with people with ‘Positive’ or ‘Negative’ personality descriptors resulted in a mean score of 63.33 versus 57.85, such that body satisfaction was enhanced following comparison with an avatar with a more positive personality.

**3.3.1.3 Interaction of comparison and personality.** Finally, there was no significant interaction between appearance comparison and personality (*F* = 1.59, *NS*). Thus, the personality of the comparator does not differentially influence how one feels about one’s body when comparing with attractive or unattractive avatars.

**3.3.1.4 Summary.** When comparing themselves to an avatar who is perceived as attractive, women feel worse about their bodies. Individuals feel better about their body when comparing to a ‘nicer person’.

**3.3.2 Impact of appearance comparison on self-esteem**

 The ANOVA is detailed in Table 2. The main effects and interaction term are outlined below.

**3.3.2.1 Effect of comparison.** There was a significant main effect of direction of appearance comparison (*F* = 6.38; *p* < .01), with a large effect size (partial eta² = 0.25). Upward comparison was significantly different from the other two conditions (LSD tests). Collapsing the two ‘personality’ levels, upward comparison was associated with a mean score of 61.90, the neutral condition yielded a mean score of 68.56, and the downward comparison condition yielded a mean score of 69.61, and all three conditions were significantly different. Thus, upward appearance comparison resulted in significantly poorer self-esteem.

**3.3.2.2 Effect of personality.** There was no significant effect of personality (*F* = .01; *NS*). Collapsing the appearance scores, self-comparison to people with ‘Positive’ personality descriptors resulted in a mean score of 66.76, and comparison to ‘Negative’ personality descriptors resulted in a mean score of 66.61. Therefore, the image being paired with a positive or negative personality did not impact the participants’ self-esteem.

**3.3.2.3 Interaction of appearance comparison and personality.** Finally, there was no significant interaction between appearance comparison and personality (*F* = .46, *NS*). Thus, the personality of the comparator does not alter how one feels about one’s self when comparing with more or less attractive avatars.

**3.3.2.4 Summary.** When comparing themselves to an avatar who is perceived as attractive, women feel worse about themselves in terms of body satisfaction and general self-esteem. The described personality of the avatar does not affect their self-esteem.

**4.1 Discussion**

As with Study 1, upward appearance comparison resulted in lower body satisfaction and self-esteem compared to downward comparison and neutral conditions, even when the comparator was clearly non-human. In contrast, downward appearance comparison did not result in any positive effects. The perceived personality of the comparator avatar was relevant to body satisfaction, but regardless of the type of comparison.

**5.1 General Discussion**

These two experimental studies have tested the hypothesis that upward and downward appearance comparison will have negative and positive impacts on body image and self-esteem, consistent with Festinger’s (1954) social comparison theory. The nature of the comparator was varied across the two studies, to determine whether the same predictions hold for comparison with human and non-human images. Furthermore, the impact of the perceived personality of the comparator was investigated, to determine whether this apparently extraneous variable had a consistent moderating impact on appearance comparison.

The most consistent finding was that, as predicted by Festinger (1954), upward comparison resulted in poorer body satisfaction and self-esteem, presumably due to the hypothesised impact of providing an aspirational object. This effect applied regardless of the human nature of the comparator, and can only be assumed to be true when compared to downward comparison and neutral conditions. However, the outcome was more specific for downward comparison, which only resulted in better outcomes compared to upward comparison and neutral conditions (self-esteem and body satisfaction) where a human image was used. Taken alongside existing evidence (Leahey et al., 2007; Van den Berg & Thompson, 2007), it is clear that upward and downward body comparison have an impact on immediate state body image and self-esteem. However, these outcomes indicate that Festinger’s theory is not equally effective at explaining upward and downward appearance comparison. Upward comparison has negative effects across the board, while downward comparison has positive effects only when comparing with human images.

Furthermore, the impact of personality on the outcome of appearance comparison was unexpected. The enhanced level of body satisfaction following comparison with a less attractive person or avatar with a positive personality descriptor is not explicitly predicted by social comparison theory, which might have been expected to predict that a more negative personality would have resulted in more positive self-regard. The result might be explained by people gaining more satisfaction about feeling attractive compared with someone who has a nice personality, than compared with someone who is not a nice person.

To summarise, these experiments have shown that Festinger’s (1954) theory might need elaboration to account for the nature of the comparator. While the theory holds in terms of comparison with human images, it only partially accounts for the non-human comparison results. It appears that ‘attractive’ encourages upward comparison, but ‘unattractive’ only encourages downward comparison if the image is human. Even that latter link is moderated by the apparently irrelevant role of the personality of the comparator figure.

 While the experimental nature of this study is an important strength in reaching causal conclusions about body comparison’s effects, it is important to acknowledge the limitations of the work, and how they might be addressed in future research. First, the study is limited by its sample (young, adult females, with limited ethnic diversity), and a more diverse group should be considered in future research. The ‘attractive’ and ‘positive’ personality image for the first study was wearing a bikini, and this may have impacted the results due to the lack of clothing. However, all other ‘attractive’ images in studies 1 and 2 are fully clothed, therefore the effect of this is likely to be small. Compliance with the experimental task should also be monitored more closely than could be achieved using online methods, meaning that attentional and compliance factors should be examined in more closely controlled face-to-face settings. Such work might use methods such as eye-tracking to quantify the level and nature of comparison being undertaken, or might specify the nature of the comparison to be undertaken more clearly (e.g., face vs figure). It is also possible that the use of a within-subject design might have allowed some participants to deduce the nature of the study. Therefore, future studies might either use a between-subject design or examine for order effects. Finally, baseline measurements of self-esteem and body satisfaction were not taken, therefore it cannot be concluded that upward and downward comparison cause a difference in an individual’s normal self-esteem and body satisfaction – it can only be concluded that they differ from each other. Further studies would benefit from taking baseline measurements of self-esteem and body satisfaction.

 This initial causal evidence suggests possible targets for future research. For example, the study used extremely attractive or unattractive images, but it would be useful to determine whether these findings depended on those extremes (e.g., would moderately attractive/ unattractive images have yielded the same findings?). It would also be valuable to repeat the work with ‘attractive’ and ‘unattractive’ images that were matched for body size, to determine whether impactful comparison is based primarily on size or other features. Furthermore, this study has used avatars, but further research could determine whether other non-human comparators (e.g., emojis, cartoons, manga images) evoke similar response patterns.

Finally, it will be important to determine whether these findings are replicated among individuals with severe body image issues and among patients with eating disorders or with other body-related problems (e.g., body dysmorphic disorder). This is a particularly important consideration, as such clinical and pre-clinical groups tend to have inaccurate perceptions of their own size and appearance (e.g., Cattarin et al., 2000; Slade, 1995), which make it far more likely that any comparison will be upward and hence negative for the individual. If upward and downward comparison have similar effects in such groups, it would be useful to determine how such use of comparisons could be discouraged through the use of psychoeducation, prevention programmes, or behavioural experiments (e.g., Waller, Turner, Tatham, Mountford & Wade, 2019). Levels of change in comparison should be measured (e.g., Laker & Waller, 2020), to determine whether they relate improvements in body image and other eating characteristics.

**Funding sources**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Authors’ declaration** **of interest**

The authors have no interests to declare

**References**

Arigo, D., Schumacher, L., & Martin, L. M. (2014). Upward appearance comparison and the development of eating pathology in college women*. International Journal of Eating Disorders, 47,* 467-170. DOI: 10.1002/eat.22240.

Arroyo, A. (2014). Connecting theory to fat talk: body dissatisfaction mediates the relationship between weight discrepancy, upward comparison, body surveillance, and fat talk. *Body Image, 11,* 303-306. DOI: http://dx.doi.org/10.1016/j.bodyim.2014.04.006.

Bailey, N., & Waller, G. (2017). Body checking in non-clinical women: experimental evidence of a specific impact on fear of uncontrollable weight gain. *International Journal of Eating Disorders*, 50, 693-697. DOI:10.1002/eat.22676

Bailey, S. D., & Ricciardelli, L. A. (2010). Social comparisons, appearance related comments, contingent self-esteem and their relationships with body dissatisfaction and eating disturbance among women. *Eating Behaviors, 11,* 107-112. DOI: 10.1016/j.eatbeh.2009.12.001.

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,* 220-239. DOI: http://dx.doi.org/10.1521/jscp.2000.19.2.220.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159. DOI: 10.1037/0033-2909.112.1.155.

Daiiani, M. (2018, November 29). *Avatar versus self the correlation of Iranian digital role-players with self-estrangement and identity crisis.* 2nd National and 1st International Digital Games Research Conference: Trends, Technologies and Applications, Tehran, Iran. DOI: 10.1109/DGRC.2018.8712059.

Delaney, M. E., O’Keefe, L. D., & Skene, K. M. L. (1997). Development of a sociocultural measure of young women’s experiences with body weight and shape. *Journal of Personality Assessment, 69,* 63-80. DOI: 10.1207/s15327752jpa6901\_4.

Festinger, L. A. (1954). Theory of social comparison processes. *Human Relations, 7,* 117-140. DOI: http://dx/doi.org/10.1177/001872675400700202.

Halliwell, E., & Dittimar, H. (2005). The role of self-improvement and self-evaluation motive in social comparisons with idealized female bodies in the media. *Body Image, 2,* 249-261. DOI: http://dx.doi.org/10.1016/j.bodyim.2005.05.001.

Laker, V., & Waller, G. (2020). The development of a body comparison measure: the CoSS. *Eating and Weight Disorders*, *25*, 879-888. DOI: 10.1007/s40519-019-00698-5.

Leahey, T. M., Crowther, J. H., & Ciesla, J. A. (2011). An ecological momentary assessment of the effects of weight and shape social comparisons on women with eating pathology, high body dissatisfaction, and low body dissatisfaction. *Behavior Therapy, 42*, 197-210. DOI: http://dx.doi.org/10.1016/j.eatbeh.2016.06.006.

Leahey, T. M., Crowther, J. H., & Mickelson, K. D. (2007). The frequency nature and effects of naturally occurring appearance-focused social comparisons. *Behavior Therapy, 38,* 132-143. DOI: 0005-7894/07/0132-0143/$1.00/0.

Lin, L., & Soby, M. (2016). Appearance comparisons styles and eating disordered symptoms in women. *Eating Behaviors, 23,* 7-12. DOI: http://dx.doi.org/10.1016/j.eatbeah.2016.06.006.

Low, K. G., Charanasomboon, S., Brown, C., Hiltunen, G., Long, K., Reinhalter, K., & Jones, H. (2003). Internalization of the thin ideal, weight and body image concerns. *Social Behaviour and Personality, 31,* 81-89. DOI: 10.2224/sbp.2003.31.1.81.

Maxis. (2014). The Sims 4 [Computer video game]. Redwood City, CA: Electronic Arts.

Mills, J. S., Polivy, J., Herman, P. C., & Tiggemann, M. (2002). Effects of exposure to thin media images: Evidence of self-enhancement among restrained eaters. *Personality and Social Psychology Bulletin, 28,* 1687-1672. DOI: 10.1177/014616702237650.

Posavac, H. D., Posavac, S., & Posavac, E. J. (1998). Exposure to media images of female attractiveness and concern with body weight among young women. *Sex Roles, 38,* 187-201. DOI: 10.1023/A:1018729015490.

Ravary, A., Baldwin, M. W., & Bartz, J. A. (2019). Shaping the body politic: Mass media fat-shaming affects implicit anti-fat attitudes. *Personality and Social Psychology Bulletin 45,* 1580-1589. DOI: 10.1177/0146167219838550.

Seung, A. J. (2010). “I feel more connected to the physically ideal mini me than the mirror-image mini me”: Theoretical implications of the “malleable self” for speculations on the effects of avatar creation on avatar-self connection in Wii. *Cyberpsychology, Behaviour and Social Networking, 13,* 567-570. DOI: 10.1089/cyber.2009.0243.

Shafran, R., Fairburn, C. G., Robinson, P., & Lask, B. (2004). Body checking and its avoidance in eating disorders. *International Journal of Eating Disorders, 35,* 93-101. DOI: 10.1002/eat.10228.

Slade, P. (1995). A review of body-image in anorexia nervosa and bulimia nervosa. *Journal of Psychiatric Research, 19,* 255-265.

Van den Berg, P., & Thompson, K. J. (2007). Self-schema and social comparison explanations of body dissatisfaction: A laboratory investigation. *Body Image, 4*, 29-38. DOI: 10.1016/j.nodyim.2006.12.004.

Waller, G., Turner, H. M., Tatham, M., Mountford, V. A., & Wade, T. D. (2019). *Brief cognitive behavioural therapy for non-underweight patients: CBT-T for eating disorders.* Hove, UK: Routledge.