3D visualisation of psychometric judgements of the ideal male body

.

Authors: Sophie Mohamed[a], Robin S.S. Kramer[b], Tracey Thornborrow[b], Thomas V. Pollet[c], Martin J. Tovée[c], Piers L. Cornelissen[c]

[a]Department of Clinical Health, Aberdeen Royal Infirmary, Foresterhill Rd, Cornhill Rd, Aberdeen AB25 2ZN

[b]School of Psychology, College of Social Science, University of Lincoln, Lincolnshire, LN6 7TS, UK

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

CORRESPONDENCE TO:

Martin J. Tovée: martin.j.tovee@northumbria.ac.uk

Declarations of interest: none

Abstract

Psychological concerns are frequently indexed by psychometric questionnaires but the mental representations that they seek to quantify are difficult to visualise. We used a set of questionnaires designed to measure men's concept of their bodies including: the Drive for Muscularity Scale (DMS; McCreary & Sasse, 2000), the Perceived Sociocultural Pressures Scale (PSPS; Stice et al., 1996), the Body Appreciation Scale (BAS-2; Tylka & Wood-Barcalow, 2015), and the Sociocultural Attitudes Towards Appearance Questionnaire-3 (SATAQ-3; Thompson et al., 2004). We combined their use with an interactive 3D modelling programme to allow men to create computer-generated representations of their ideal bodies. We used a principal component analysis to extract those shape components of our participants' CGI ideal bodies that were predicted by the questionnaires and reconstructed the body shapes that these questionnaires were capturing. Moving from the lowest to the highest score on both the DMS and SATAQ corresponded with changes in muscularity, particularly muscle mass and definition. This approach allows us to demonstrate the actual body features that are being captured by a particular questionnaire.

***Keywords*:** drive for muscularity; male body image; muscularity; fat; BMI; male body ideals

1. Introduction

The ideal body for most men in Western society is characterised by both a high degree of upper body muscularity and a low degree of body fat, the latter enhancing the salience of the former (Cafri et al., 2005; Leit, Gray, & Pope, 2002; McCreary, Hildebrandt, Heinberg, Boroughs, & Thompson, 2007a; Murray et al., 2017; Ridgeway & Tylka, 2005). Several lines of evidence have led to this conclusion about men's ideal body shape. For example, Ridgeway and Tylka (2005) applied the Consensual Qualitative Research method (CQR; Hill, Thompson, & Nutt-Williams, 1997) to open-ended interview questions which were given to male undergraduates. They found that most men expressed greatest interest in muscular definition and leanness in and around the arms, chest and particularly the abdomen. Analysis of participants' responses suggested that the ideal male body derived from five key components – overall body muscularity, overall body leanness, being tall, V-shape torso, and a muscled abdominal region. Undesirable body characteristics included fat, short stature, and low body fat coupled with low muscle tone leading to small girth (Ridgeway & Tylka, 2005). Comparable results have been reported by others, including Grogan and Richards (2002) and Pope et al. (2000). Qualitative data like these have led to the development of quantitative psychometric tools, such as the Drive for Muscularity Scale (DMS; McCreary, 2007b; McCreary & Sasse, 2000), the two subscales of which, and the total score have shown acceptable levels of internal consistency, test-retest reliability, and construct validity (McPherson, McCarthy, McCreary, & McMilland, 2010).

From a longitudinal and cultural perspective, the tendency towards increasing leanness and muscle bulk has been systematically mirrored in the physiques of male models (Frederick, Fessler, & Haselton, 2005; Lanzieri & Cook, 2013), movie stars (Pope, Phillips, & Olivardia, 2000), computer game characters (Martins, Williams, Ratan, & Harrison, 2011), and action play dolls (Baghurst, Hollander, Nardella, & Haff, 2006). The body shapes for the latter two categories have even accrued levels of muscularity that are biologically implausible (Martins et al., 2011).

Here, we wanted to find direct, rather than indirect, evidence for what the ideal male body actually looks like. Is it possible to give three-dimensional form to the ideal body residing in a man's head? Arguably, the evidence from questionnaires and psychometric data, however valid, reliable, and repeatable, does not constitute a direct visualisation of this ideal. Rather, it simply provides a verbal description of it. Similarly, the longitudinal cultural data can be thought of as a visual proxy for men's ideals, at least as far as the creative expressions of film directors, toy makers, model agencies, and computer games manufacturers are concerned. Perhaps the closest approximation to what we seek is to be found in participants' responses to variants of the somatomorphic matrix (Pope, Gruber, Mangweth, Bureau, Decol, Jouvent, et al., 2000). In an early incarnation of this, Cafri and Thompson (2004) constructed a set of line-drawn images intended to represent combinations of increasing adiposity and muscularity throughout the matrix. More recent versions (e.g., Arkenau et al., 2020) have adopted computer generated imagery (CGI) to achieve the same result. Typically, participants might then be asked to select their actual ("How do you actually look?"), felt ("How do you feel you look?"), and ideal bodies ("How would you like to look?") from the 2D image array. However, we would argue that this approach may be too restrictive and is limited in its ecological validity through the use of poor imagery (Gardner, Stark, Jackson, & Friedman, 1999; Ralph-Nearman & Filik, 2018), separate considerations of adiposity and muscularity (Ralph-Nearman & Filik, 2018; Talbot, Smith, & Cass, 2019), and/or the presentation of partial body size and shape information as a result of figure orientation (Crossley, Cornelissen, and Tovée, 2012; Gardner, Friedman, & Jackson, 1998). The only information that changes in the matrix is covariation in adiposity and muscularity of the body as a whole, with no consideration of specific areas of the body that may be particularly pertinent to men's body ideals (Ridgeway & Tylka, 2005) or ethnic differences in patterns of body development (Abe et al., 2012; Shiwaku et al., 2004; Silva et al., 2010). Therefore, this is the only information a participant has in order to make a match between the ideal body shape/size in his head and the stimuli presented to him. As a consequence, we must ask whether there are other features about his mental image that he would like to express, given the opportunity.

1.1 Current study

In the current study, we used a 3D modelling environment to allow male participants to recreate their ideal body sizes and shapes. The degrees of freedom available to them were not infinite, as they would be, if we had asked them to sculpt their ideal body from a lump of clay. Instead, we allowed participants to change the shape of an initially thin and an initially fat version of a baseline CGI model by manipulating any of the 18 shape morphing slider controls available to them. The BMI of these CGI bodies was calculated based on their relative size and shape and compared to the participants' actual BMI (a measure of body dissatisfaction). We then submitted the CGI bodies to a principal component analysis (PCA) to extract uncorrelated 3D shape features from participants' models. We then sought statistical dependencies between the factor scores for these visualisable principal components (PCs) and participants' performance on a number of standardised psychometric tasks that assessed their drive for muscularity, attitudes toward and appreciation for their body, and sociocultural influences. In a third step, we created direct visualisations for a subset of these psychometric tasks to show how body features for ideal body shape change from the minimum to the maximum scale scores.

2. Methods

The study was granted ethical approval by the University of Lincoln School of Psychology Research Ethics Committee on the 12th December 2017 (PSY1718346).

2.1 Sample size calculation and participants

We are not aware of any investigation that has used the same approach as we have followed in the current study. The nearest equivalent we know of is a study by Smith et al., (2007) in which 2D photographs of women were decomposed into 4 body-shape features using PCA. These features were then mapped onto an observer's behavioural ratings of the images. A multiple regression model of these judgements, using the 4 PCs as explanatory variables, accounted for 70% of the variance in ratings. Here, we estimated how many participants would be needed to estimate a new multiple regression model with 3, 4, or 5 independent variables for the null hypothesis:

$$H_0: R^2_{Y.B} = 0$$

For a power of .9 and an alpha of .05, a sample size of 11 to 14 would be required for models with 3 to 5 explanatory variables in order to account for 70% of variance (G*Power v3.1.9.7).

Conservatively, if the multiple regression models only accounted for 35% of the variance, i.e., half of that described in Smith et al. (2007), we would need a sample size of 31 to 37 participants. To offset attrition and possible data loss, for the current study we recruited 42 men from the University of Lincoln staff, students, and the general population near Lincoln. All participants self-identified as 'White British', except for one participant who identified as 'White British/Norwegian. We used email invitations, social media sites, and posters to recruit this sample. Inclusion criteria for participants were: (a) aged 18 to 45 years inclusive, and (b) self-identification as male.

2.2 Psychometric Measures

2.2.1. Drive for muscularity: We used the Drive for Muscularity Scale (DMS; McCreary & Sasse, 2000) to measure an individual's motivation and preoccupation with increasing their muscularity levels, as well as their engagement in relevant behaviours. The scale consists of 15 items rated on a 6-point response scale (1 = *never*, 6 = *always*) which are summed to give a total DMS score. The questionnaire has two individual subscales – the Muscularity Behaviours subscale (MB) and the Muscularity-oriented Body Image subscale (MBI). The MB consists of 8 items assessing an individual's engagement in behaviours that can lead to increased muscle mass, such as the consumption of protein shakes and a high-calorie diet. The MBI comprises 7 items that measure an individual's satisfaction with their body shape and desire to increase their own muscle mass.

2.2.2. Positive body image: We used the Body Appreciation Scale-2 (BAS-2; Tylka & Wood-Barcalow, 2015) to measure men's levels of positive body image and body appreciation. This 10-item scale is rated on a 5-point response scale (1 = *never*, 5 = *always*) and a total score across all items is calculated, with higher scores representing a greater overall body appreciation.

2.2.3. Sociocultural influences: We used the Sociocultural Attitudes Towards Appearance Questionnaire-3 (SATAQ-3; Thompson, van den Berg, Roehrig, Guarda & Heinberg, 2004) to measure the influence of the media on sociocultural factors such as appearance standards and ideals. This 30-item questionnaire is rated on a 5-point response scale (1 = *disagree strongly*, 5 = *agree strongly*), and consists of four individual subscales – internalisation-general, internalisation-athlete, information, and pressures. The internalisation-general subscale measures the degree to which unrealistic body image ideals from mass media are accepted, while the internalisation-athlete subscale focuses specifically on acceptance of an athletic muscular body ideal. The information subscale assesses the influence of various forms of media as a source of information for ideas of attractiveness. Finally, the pressures subscale measures direct pressures from media sources to achieve sociocultural body ideals.

2.2.4. Perceived sociocultural pressures: We administered the Perceived Sociocultural Pressures Scale (PSPS; Stice, Nemeroff, & Shaw, 1996) to assess an individual's perceived pressure for thinness from external factors, with separate subscales relating to family members, friends, intimate partners, and the media. The 20-item questionnaire is rated on a 5-point response scale (1 = *none*, 5 = *a lot*) that is summed to form a total PSPS score as well as individual subscale scores, with higher scores representing greater perceived pressures to be thin.

2.3 Procedure

At the start of each session, participants were asked to provide personal demographic information, including their age, nationality, occupation, years in full-time education, and engagement in behaviours to change their current body weight or muscle mass.

Body measurements were then taken from each participant, including height (cm), weight (kg), chest circumference (cm), hip circumference (cm), and waist circumference (cm), using a digital weighing scale and tape measure. Participants were required to remove any footwear and bulky clothing prior to measurement. Individuals were also given the opportunity to take the body measurements themselves in order to accommodate any potential religious, cultural, or other concerns. Participants were then presented with a 3D Genesis 2 male avatar on a computer screen using the Daz Studio 4.10 software. The researcher demonstrated the full range of body size/shape variation that was possible using a set of 18 preselected Genesis 2 body morphs for manipulating different regions of the avatar. The following body morphs were made available to participants to create their ideal body perceptions: *body size, body tone, bodybuilder details, bodybuilder size, emaciated, portly, weight, shoulder size, shoulder width, upper arms size, chest width, pectorals heavy, lats size, glutes size, glutes width, hip size, thigh size*, and *thigh tone*. Therefore, participants were given the opportunity to modify both specific areas and more general whole-body aspects of the avatars. Participants were then shown either a ‘larger’ (BMI = 35.95) or ‘thinner’ (BMI = 13.04) male avatar and were asked to orally describe how they would like to alter the 3D body to create their ideal body size/shape. The researcher then adjusted the body morphs in the software, based on the instructions given by the participant, to visualize these body perceptions. Importantly, the developing model could be rotated through 360° so that the participant was able to see clearly the impact of any shape changes they sought. The ideal body task was run twice in a row, starting from either the ‘larger’ or the ‘thinner’ 3D avatar. The initial body type (‘larger’ or ‘thinner’ body) was randomised across participants, and no time limit was set for creating the models. Figure 1 illustrates the high degree of variability in muscularity and adiposity that could be achieved with the stimulus set up.

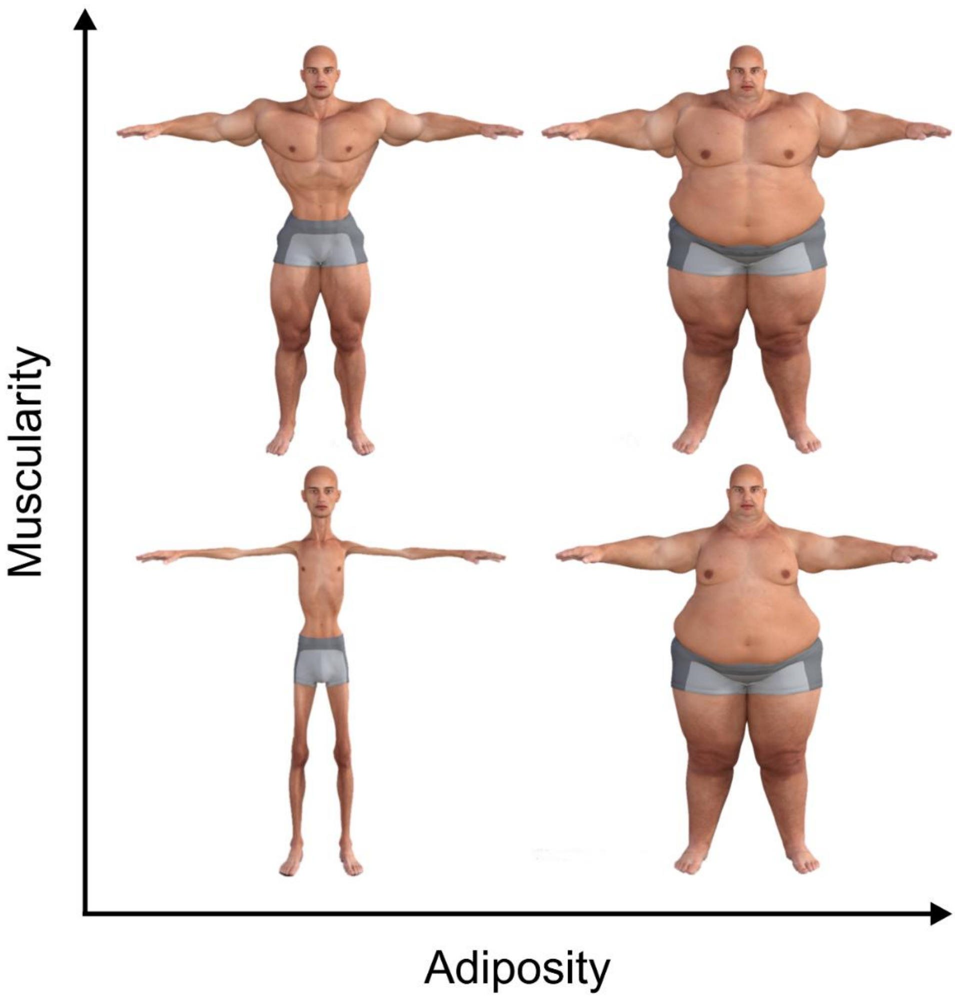


Figure 1. Illustrates the high degree of variability in adiposity and muscularity that could be achieved with the stimulus set up.

Finally, participants were asked to complete the following four psychometric measures: BAS-2, DMS, SATAQ-3 and the PSPS. The data collection procedure took approximately 30-45 minutes per participant.

2.4 Estimating BMI of ideal model

First, each ideal body created by participants in Daz Studio was adjusted in the software to match the actual height of the participant in the real world. Circumference measurements of the chest, waist, and hips of each avatar were then recorded using the 'Measuremetrics' function in Daz Studio. Average measurements were calculated from the two bodies created for the ideal body task. The height, waist and hip circumferences were then used to determine the ideal BMI of the CGI bodies, using the formula for BMI below (Cornelissen et al., 2015; Cornelissen, McCarty, Cornelissen, & Tovée, 2017; Groves et al., 2019). The ideal BMI measurements were then compared with participants' actual BMI values.

$$BMI = 6.8195 + (0.21302\times hip) + (0.22509\times waist) - (0.13991\times height) + (0.06781\times age) - (0.00101\times age^2)$$

2.5 Extraction of 3D shape features from Daz ideal models

In order to be able to relate men's responses in the psychometric tasks to the ideal body shapes that they created within the Daz modelling environment, we had to extract a set of uncorrelated 3D shape features from the Daz models they produced. To do this, we used customised MATLAB software. Each body shape created by a participant was converted to a set of 85,253 three-dimensional coordinates. These were subsequently unwrapped into a vector, or long row, of 255,759 numbers (i.e., 85,253 × 3, because each point has an x, y, and z coordinate). Given that each participant provided two shapes regarding their ideal body, the two corresponding vectors were averaged to produce a single 255,759 number vector representing the average of the two body shapes.

Next, all 42 vectors (one per participant) were entered into a PCA. This method derives a multidimensional space in which to locate the body shapes, based on the eigenvectors of a PCA-decomposition of these shapes. Following this approach, which has been used extensively in the literature regarding face (e.g., Burton, Kramer, Ritchie, & Jenkins, 2016; Kramer, Young, & Burton, 2018; Kramer, Young, Day, & Burton, 2017) and body images (e.g., Collins, Zhang, Miller, Wang, & Zhou, 2010; Ruto, Lee, & Buxton, 2006), allowed us to identify the few, important dimensions along which our body shapes varied in a low-dimensional space (in comparison with the 255,759-dimensional space in which they originally appeared).

In order to identify those PCs which were important in explaining shape variation in our sample without significant loss of information, and in line with previous research using this approach (e.g., Kramer et al., 2017), we chose to keep the earliest components that cumulatively explained at least 95% of the variance in the 3D coordinate information. (In PCA, derived components are ordered from highest to lowest in terms of the amount of variance explained.) For our PCA, only the first four components were required in order to explain 95.4% of the variance, and as such, the remaining components were discarded. Mathematically, not only is it possible to reduce the dimensionality of the body shape data into, in this case, 4 PCs. It is also possible to reverse this process and reconstruct the 3D shape components which are encoded in the lower-dimensional space of the PC scores, by projecting back from them into 3D Cartesian space. By doing this, we can illustrate the shape features that are separately encoded by each of the 4 PCs, as shown in Figure 2. To ensure that this illustration emphasises what the features look like qualitatively, we reconstructed the body shape features encoded by each PC at factor scores +5 and -5 SD.

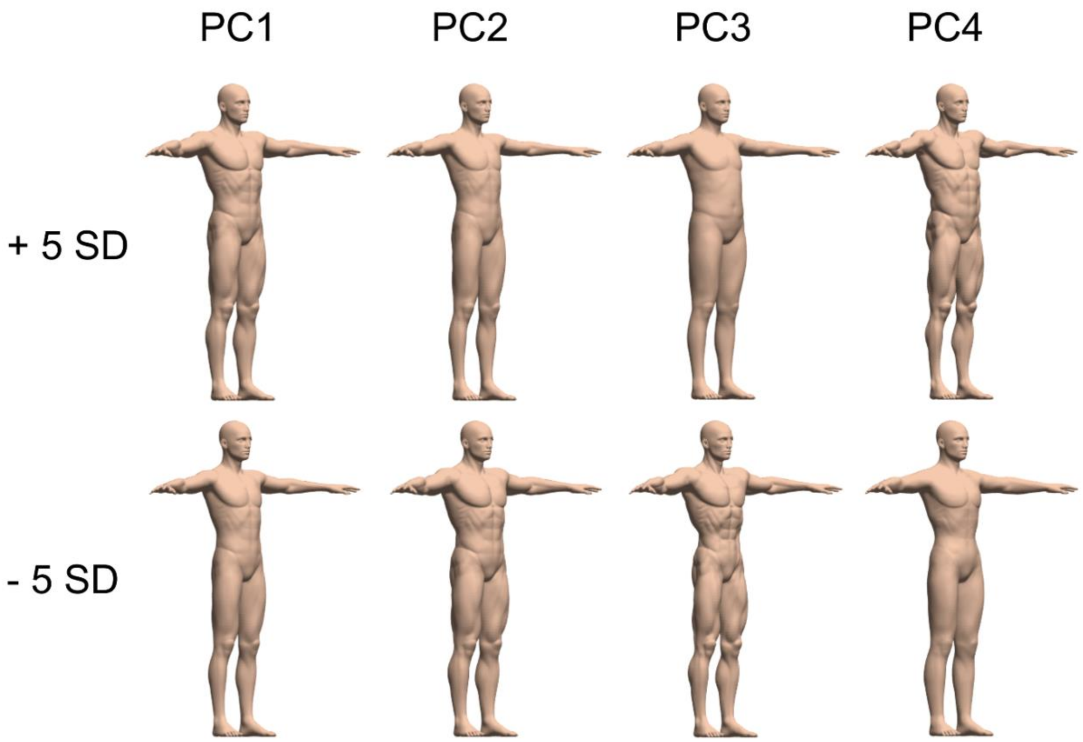


Figure 2. Shows the body shape variability associated with the 4 principal components. Each column corresponds to an individual principal component. The top row shows the principal components all set to +5 SD. The bottom row shows the principle components all set to -5 SD.

PC1 appears to represent a change in the bulk of the bodies. There is some change in muscle definition, but relatively little change in shape. It is more a change in the body circumference. PC2 represents a change in overall musculature, particularly in the upper torso creating the traditional V-shape male torso. PC3 also represents a change in the body's muscularity, but more directly focussed on the development of the V-shaped torso with a narrower waist and a broader chest and shoulders. PC4 indexes a more complex change in multiple physical parameters. This includes an increase in muscle definition, particularly on the torso, and a thickening of the torso as part of an increase in muscle mass.

2.6 Summary of analysis path

- Descriptive statistics of participants' characteristics, anthropometric measures, psychometric task performance including Cronbach's alpha, and Pearson correlations between psychometric tasks.
- Multiple regression analysis of participants' ideal BMI predicted by their actual BMI, and psychometric task performance.
- Correlation analysis to establish which of the 4 PC factor scores, extracted from the ideal Daz models, are associated with participants' psychometric task performance.
- Visualization of the range of body shape variation that is captured between the minimum and maximum scores on the DMS and SATAQ Global psychometric tests.
- Content validity analysis. Tests how well the body shape changes that are reconstructed from our PCA + regression approach, at the lowest and highest DMS/SATAQ Global scores, map onto the equivalent changes for the 10 highest and 10 lowest scoring raw Daz ideal bodies.
- Cross-validation. Tests how well the reconstructed bodies from the PCA + regression approach for the DMS and SATAQ Global scores predict the raw ideal Daz body shapes. Moreover, we used a leave-one-out (LOO) cross-validation technique to ask how well this approach would predict new data from a novel participant sample.

3. Results

3.1 Univariate statistics

Table 1 shows the age, biometric measurements, educational and occupational status, intention towards body weight change, and media usage for the 42 male participants.

Table 1: *A summary of participant characteristics.*

	Mean	*SD*
Chronological age (years)	23.64	5.29
Years in full time education (years)	15.44	3.18
Height (cm)	179.65	8.30
Weight (kg)	79.09	14.62
Waist circumference (cm)	86.56	9.63
Hip circumference (cm)	98.91	7.09
Chest circumference (cm)	98.38	9.39

BMI (kg/m^2)	24.41	3.52
Hours television watched /week	14.90	7.30
	N	%
Occupational status		
Full/part time employment	17	40.48
Student	25	69.52
Intention to change body weight		
None	12	28.57
Decrease weight	13	30.95
Increase weight	7	16.67
Increase muscle mass	10	23.81

Table 2 shows psychometric task performance. Cronbach's alpha shows good reliability across all psychometric tasks. According to normative data for the DMS (Hughes, Dean, & Allen, 2016) 2/42 participants fell into the first quartile, 4/42 the second quartile, 15/24 the third quartile, and 21/42 the fourth quartile. Based on normative data for the SATAQ reported by Karazsia and Crowther (2008), we calculated the number of participants whose sub-scale scores fell into three groups: (a) less than -1SD below the mean, (b) between -1SD below and +1SD above the mean, and (c) higher than + 1SD above the mean. The respective numbers of participants within each of these groups for the SATAQ Internalization-General subscale were 3, 16, and 23; for the Internalization-Athletic subscale 2, 17, & 23; for the Pressures subscale 2, 23, & 17, and for the Attitudes subscale 8, 24, and 10. While these results suggest that our sample of men represented the full range of possible psychometric responses, clearly there was a substantial proportion of participants with high drive for muscularity associated with high levels of internalized societal pressure for an ideal body.

Table 2: *Participants' performance in the psychometric tasks.*

Psychometric task	*Mean*	*SD*	Range: Actual	Range: Potential	Cronbach alpha
BAS-2	38.50	6.55	27 – 50	10 – 50	.90
PSPS	44.71	13.48	22 – 76	20 – 100	.89
DMS	48.69	15.03	23 – 90	15 – 90	.90
SATAQ-Internalisation (general)	29.17	8.58	11 – 44	9 – 45	.91
SATAQ-Internalisation (athlete)	19.31	3.86	8 – 25	5 – 25	.87
SATAQ-Pressures	21.83	6.65	9 – 34	7 – 35	.89
SATAQ-Information	23.38	8.72	9 – 38	9 – 45	.76

Finally, Table 3 shows the Pearson correlations between the psychometric tasks. Table 3 clearly illustrates strong, positive associations between tasks that tap societal pressures for thinness and muscularity, the degree to which these factors are internalised, and an individual's preoccupation with and engagement in behaviours that increase muscularity. In addition, as all of these tendencies increased, so the association with participants' levels of positive body image and body appreciation were significantly reduced.

Table 3: *Pearson correlations between the psychometric tasks, across the whole sample.*

	BAS-2	DMS	PSPS	SATAQ-Int(gen)	SATAQ-Int(ath)	SATAQ-Pressures
DMS	-.37*	-				
PSPS	-.18	.38*	-			
SATAQ-Int(gen)	-.32*	.61***	.38*	-		
SATAQ-Int(ath)	-.42**	.54***	.18	.62***	-	
SATAQ-Pressures	-.41**	.54***	.57***	.75***	.58***	-
SATAQ-Information	-.26	.50***	.39*	.72***	.43**	.63***

* = $p < .05$; ** = $p < .01$; *** = $p < .001$

NB: SATAQ-Int(gen) = SATAQ-Internalisation (general); SATAQ-Int(ath) = SATAQ-Internalisation (athletic).

3.2 Multivariate statistics

3.2.1 Comparison of ideal BMI with actual BMI

Given that participants were asked to model the body size/shape that they would like to have, we first wanted to understand the relationship between men's actual and ideal BMI, while controlling for any influences of participants' psychometric performance. Given the substantial correlations between the psychometric tasks shown in Table 3, we sought to include a selection procedure in the model which would avoid potential problems with multicolinearity. Since stepwise selection algorithms are known to lead to biases in parameter estimation (Grafen & Hails, 2002; Hurvich & Tsai 1990; Steyerberg, Eijkemans, & Habbema, 1999), we used PROC GLMSELECT in SAS v9.4 (SAS Institute, North Carolina, USA) to run adaptive LASSO (least absolute shrinkage and selection operator) regression for variable selection (Efron, Hastie, Johnstone & Tibshirani, 2004; Osborne, Presnell & Turlach, 2000; Tibshirani, 1996). LASSO and stepwise regression differ in their criteria for retaining predictors in the final model, and LASSO has been shown to produce more stable results. The LASSO algorithm selects an optimal value for t, the tuning or shrinkage parameter which, in our case, minimized the Schwarz Bayesian information criterion (SBIC) for model fitting. We included actual BMI, BAS-2, PSPS, DMS, SATAQ-Internalisation (general), SATAQ-Internalisation (athletic), SATAQ-Pressures, and SATAQ-Information as explanatory variables at the start of the selection procedure. By the end of selection, the optimal subset of variables chosen to model ideal BMI had a SBIC value of 75.7. We then used PROC GLM (SAS v9.4) to build an ordinary least squares multiple regression model of ideal BMI predicted by actual BMI, BAS-2, and PSPS. This best fit model explained 37.9% of the variance in ideal BMI. We found statistically significant effects of actual BMI ($F(1,38) = 6.51$, $p = .015$, $\beta = .23$, 95% CI [.048, .42], partial $\eta^2 = .15$), BAS-2 ($F(1,38) = 5.85$, $p = .021$, $\beta = -.12$, 95% CI [-.22, -.020], partial $\eta^2 = .13$), and PSPS ($F(1,38) = 5.42$, $p = .025$, $\beta = .056$, 95% CI [.0073, .10], partial $\eta^2 = .12$). This model outcome is illustrated in Figure 3. It shows a scatterplot of ideal BMI predicted from the regression model as a function of actual BMI. The regression line is for the entire sample, and shows a tendency for participants' ideal body size to compress toward a narrow range from ~19 to ~22. This means that individuals with a lower BMI sought larger bodies, and vice versa, men with higher BMIs sought smaller bodies. Moreover, the tendency to seek larger bodies increased with lower BAS-2 scores and higher PSPS scores, independent of a man's actual BMI.

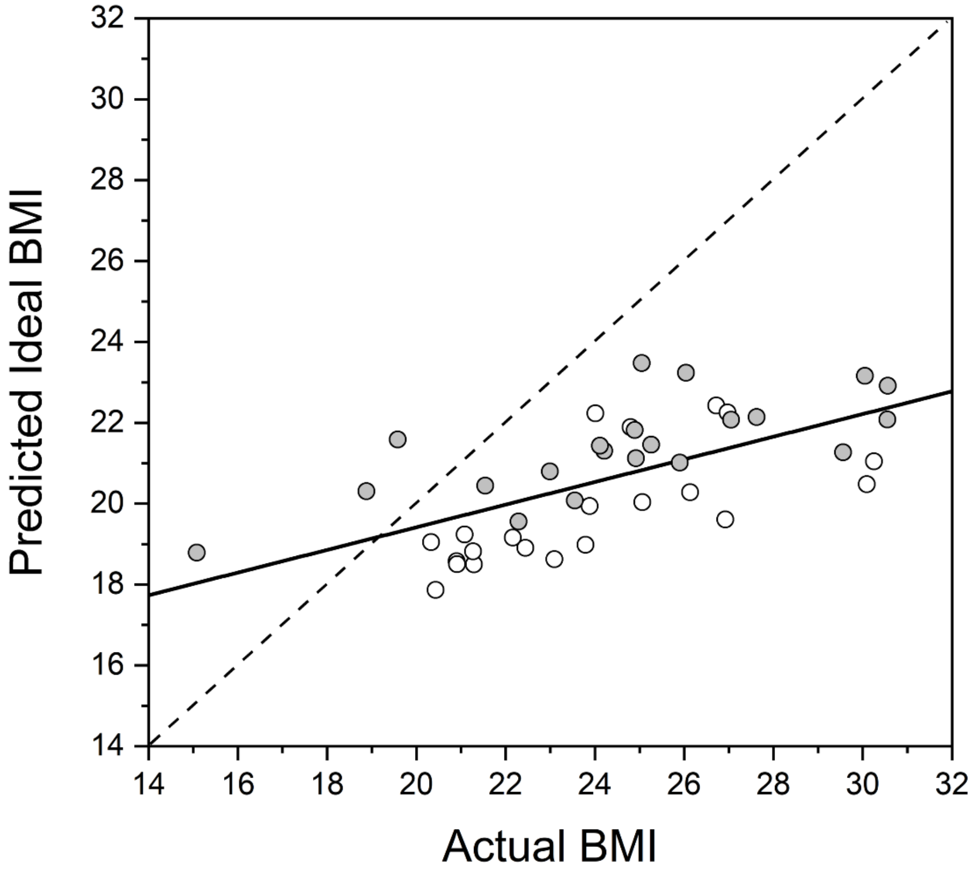


Figure 3. A scatterplot of model-predicted ideal BMI as a function of actual BMI. The dashed line represents the line of equivalence where actual and ideal BMI values are identical. The data points are colour coded according to whether the participant scored at or below the median for BAS-2 (gray) or above the median (white). The regression line is for the entire sample, ignoring BAS-2 status.

3.2.2 Correlations between 3D shape features and psychometric performance

Next, we wanted to address the central question of this study – to what extent do the 3D shape features extracted from the ideal models map onto psychometric task performance? We answered this question initially by calculating the Pearson correlations between the 4 PCs and BAS-2, DMS, PSPS and SATAQ Global scores. We also calculated the total r-square for each psychometric task associated with the PCs, which corresponds to the sum of the four squared Pearson correlation coefficients. This gives a sense of the overall strength of association between the psychometric tasks and the shape variation in the ideal Daz models. Table 4 shows that the BAS-2 was statistically significantly associated with PC2, DMS with PC2 and PC3, PSPS with PC1, and SATAQ Global with PC3.

Table 4: *Pearson correlations between PC scores and psychometric tasks*

	PC1	PC2	PC3	PC4	R-square Total
BAS	-.075	.39**	.083	-.030	.17
DMS	.29	-.33*	-.43**	-.13	.39
PSPS	.39**	-.19	.079	.081	.20
SATAQ Global	.27	-.21	-.36*	-.098	.26

*$p < .05$; **$p < .01$; ***$p < .001$

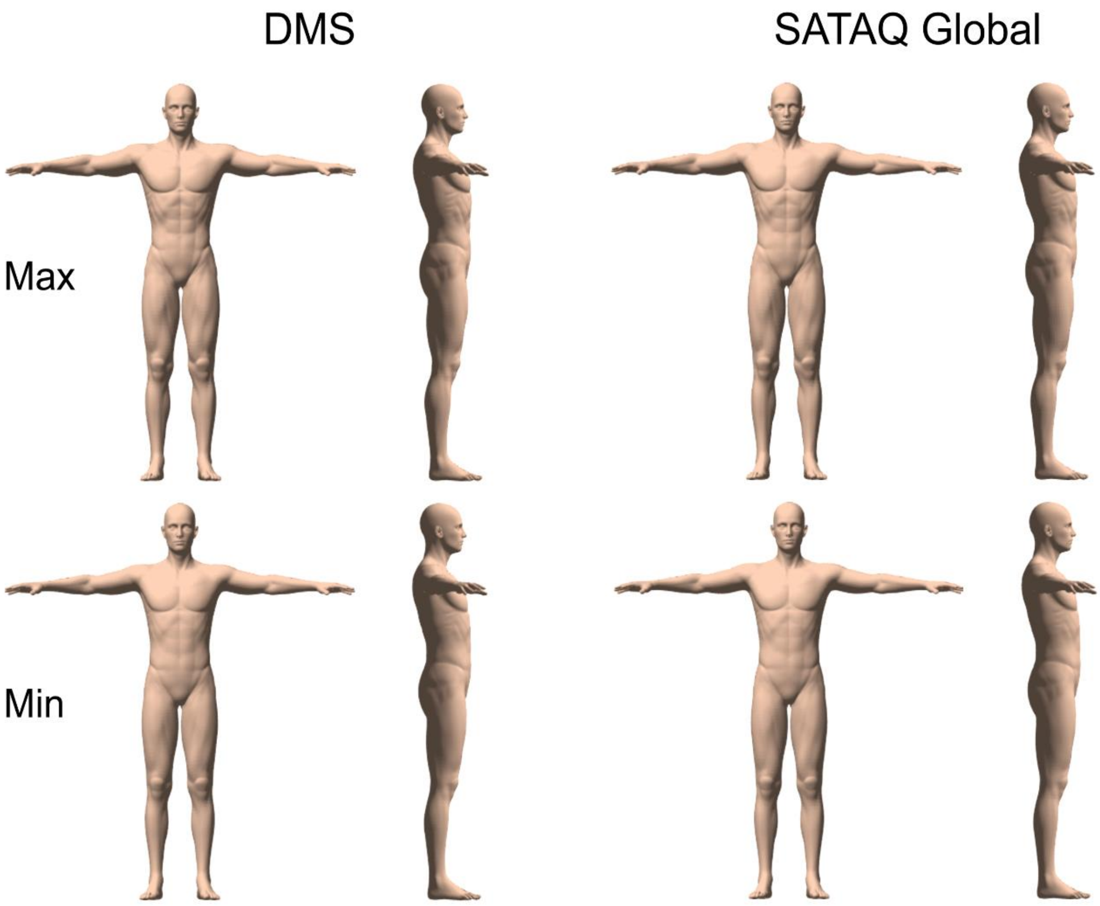
There are three main observations to be made from this analysis. First, the 3D shape features extracted by the PCA of the ideal models were either only moderately associated with these psychometric tasks or not at all (e.g., PC4); there is not a simple one-to-one mapping. Second, associations between different shape features (e.g., PC2 and PC3) can be captured by the same psychometric task (e.g., DMS). Finally, it is clear particularly for PC2 and PC3 that variation in the *same* 3D shape feature can be associated simultaneously with different psychometric tasks (e.g., BAS-2, DMS, and SATAQ Global). Given the strength of inter-correlation between the psychometric tasks, as illustrated in Table 3, this is perhaps not surprising.

3.2.3 Visualising the relationships between 3D shape, the DMS, and PSPS

Inspection of Table 4 gives very little insight into how systematic variation along the range of a particular psychometric task may relate to body shape change, especially if the same psychometric task is associated with variation in more than one uncorrelated shape feature, as is the case with the DMS which was significantly correlated with PC2 and PC3. Therefore, as a final step in our analysis, we asked the following questions: (a) how do separate shape features change in relation to a particular psychometric task, in isolation? and (b) if we carry out a linear recombination of these varying shape features, what do the resultant changes in body shape look like? In short, what body shape corresponds to a particular score on, e.g., the DMS? Given the highest r-square totals in Table 4 were for DMS and SATAQ Global, we focused this analysis on just these two tasks. This analysis strategy is viable precisely because the four PCs are statistically uncorrelated with each other. Therefore, it is legitimate to quantify the relationship between DMS/SATAQ Global by running four separate regression analyses, one for each PC. If, in theory, there was residual correlation between the PC shape features, we would have to run a multivariate multiple regression analysis.

To carry out this analysis, we used PROC REG (SAS v9.4) to run 4 ordinary least squares regression models each for the DMS and, separately, the SATAQ Global. This gave a total of 8 regression models. In each case, we used PC1, PC2, PC3, or PC4 as the outcome measure and DMS or SATAQ Global as the predictor variable. We then used the intercepts and regression coefficients in each case to compute the predicted PC score for the highest and lowest possible scores on each of the psychometric tasks. For example, the minimum and maximum scores on the DMS are 15 and 90 respectively. The regression of PC3 on DMS produced an intercept of 1.379 and a regression weight, β, of -0.0283. Therefore, the PC3 factor scores corresponding to the bottom and top of the DMS scale are 0.955 (i.e., $1.379 - 0.0283\times15$) and -1.168 (i.e., $1.379 - 0.0283\times90$) respectively.

This reconstruction strategy, which we will refer to as "PCA + regression", rendered the predicted PC values at the lower and upper ranges for each psychometric task, and these values could then be used to select appropriately scaled body shape features for linear recombination. Figure 4 shows the results of recombining the PC features that contribute significantly to the DMS and SATAQ Global. Visual inspection of Figure 4 suggests that the outcomes are very similar for both psychometric tasks, and capture changes in muscularity, particularly with respect to muscle mass and definition. The strong similarity between the two outcomes is attributable to the majority contribution of PC3 in both cases (see Figure 2 and Table 4).



	R-square	Scale minimum (z-score)	Scale maximum (z-score)
PC1	0.083	-0.65	0.79
PC2	0.11	0.74	-0.91
PC3	0.18	0.95	-1.17
PC4	0.017	0.29	-0.36

	R-square	Scale minimum (z-score)	Scale maximum (z-score)
PC1	0.073	-0.72	0.64
PC2	0.045	0.56	-0.49
PC3	0.13	0.96	-0.85
PC4	0.0095	0.26	-0.23

Figure 4. Shows bodies reconstructed using the "PCA + regression" approach. They illustrate the ideal body shape changes captured by the DMS (columns 1 and 2) and, separately, the SATAQ (Global) (columns 3 and 4). The top row shows reconstructions at the top of the scale for each psychometric task (maximum values for the DMS and SATAQ Global are 90 and 150, respectively). The bottom row shows reconstructions for the bottom of each scale (minimum values for the DMS and SATAQ Global are 15 and 30, respectively). Each table shows the r-squared for the variance in each PC explained by the DMS (left) and SATAQ Global (right) scores. In addition, the z-score settings for each PC are shown for the reconstructions.

3.2.4 Content validity

We wanted to determine whether the DMS and SATAQ models that our "PCA + regression" strategy predicted, and which are illustrated in Figure 4, correctly reflected the body shape changes that participants created in relation to their performance on the psychometric tasks themselves. In other words, we wanted to assess the content validity of the PCA models for DMS and SATAQ. A useful definition for content validity is: "… if the items of a test can be shown to reflect all aspects of the subject being tested, then it is per se valid, given that the instructions are clear. This is not simply face validity, which is related to the appearance of the test items …" (Kline, 2015). To do this, we averaged together the raw Daz models created by participants who had the 10 lowest and 10 highest scores on the DMS, and separately, the 10 lowest and 10 highest scores on the SATAQ. This gave us estimates of the averaged, raw Daz models corresponding to both the bottom and top of the DMS and SATAQ scales, respectively.

We then calculated the differences between the low and high scoring models, separately for the DMS and SATAQ, and plotted these differences on the 3D reference body, as is shown in the first and third columns of Figure 5. This is possible because every model represented a polygonal mesh with a pattern of nodes and edges that is topologically identical across different models. Therefore, the Euclidian distance between corresponding nodes from two models, located on the elbow for example, can be calculated. Consequently, we can visualize where on the body the largest shape changes occur when moving from the bottom to the top of each psychometric scale, by colour coding these differences in Euclidian distances. Small to large differences between low and high DMS/SATAQ scores are shown on the colour scale from black, through red and yellow, to white. We then went through exactly the same procedure for the PCA models for the DMS and SATAQ. The corresponding low versus high difference for the PCA models are shown in columns two and four of Figure 5. Note that in all examples, we excluded the heads, hands, and feet from these calculations. This was because the model manipulations had minimal effects on these body parts. Therefore, to include them in the comparisons would inflate the goodness of fit.

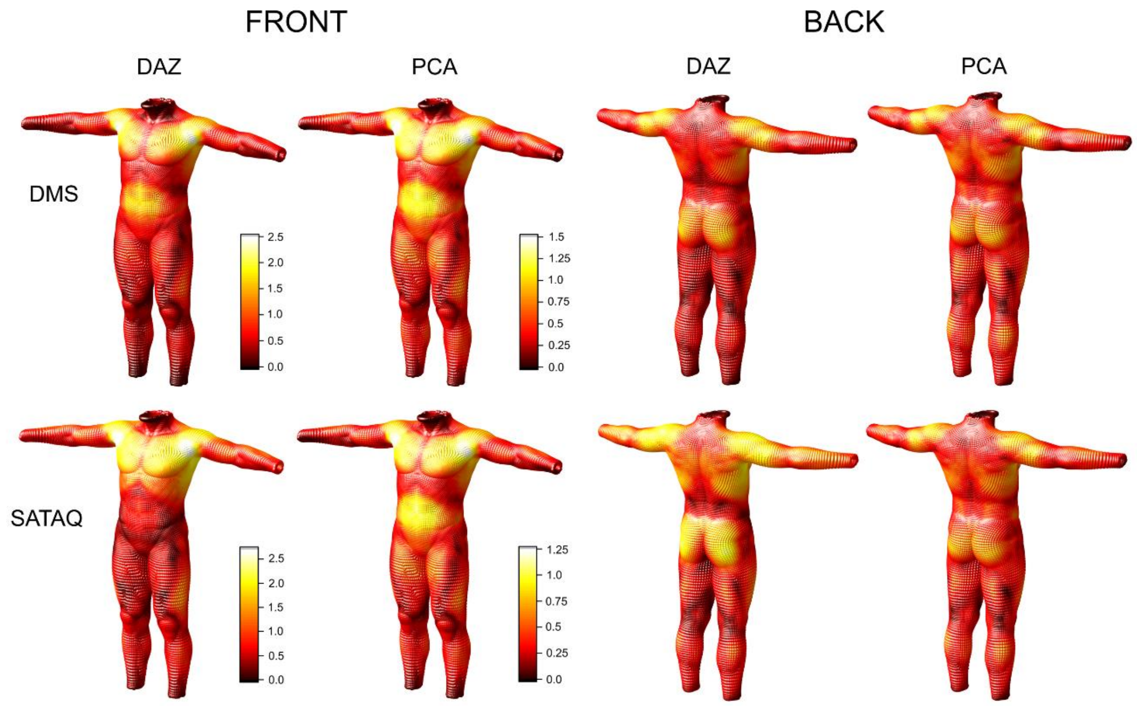


Figure 5. Each mannequin shows the colour coded difference between the averages of the 10 lowest and 10 highest DMS (upper row) and SATAQ Global (bottom row) body shapes. The smallest to largest differences (centimetres) are indexed by the colour scale changes from dark brown, through red and yellow, to white. The first two columns show the front view differences, and the second two columns show the back views. The comparisons for the raw Daz ideal models are shown in columns 1 and 3. The comparisons for the "PCA + regression" models are shown in columns 2 and 4.

Visual inspection shows that the largest differences between the highest and lowest DMS/SATAQ scores in the Daz averages (i.e., columns 1 and 3) are located in the shoulders, the pectoralis muscles, the latissimus dorsi, the abdominal muscles, and the gluteus maximus. Moreover, these localized differences are well captured by the PCA model for DMS (columns 2 and 4, upper row), to the extent that the Pearson correlation between the raw Daz differences and the PCA differences for the DMS was $r = 0.94$. For the SATAQ, the raw Daz averages show largest changes to the shoulders, the pectoralis muscles, latissimus dorsi and gluteus maximus. While the PCA model for SATAQ (columns 2 and 4, bottom row) captures the shoulder and pectoral muscle changes, it also appears to suggest abdominal muscle changes which are not so salient in the raw Daz averages. Moreover, the PCA model for the SATAQ appears to under-estimate the extent of change to the gluteus maximus compared to the raw Daz average. These discrepancies likely explain the lower Pearson correlation between the raw Daz differences and the PCA differences for the SATAQ, which was $r = 0.69$. Based on these results, we argue that the DMS model in particular showed good content validity.

3.2.5 Cross-validation

Cross-validation includes a variety of model validation techniques to assess how the outcome of a statistical analysis is likely to generalize to an independent data set of the same kind. To this end, a model is first trained on a known dataset and then its ability to predict the same outcome(s) from a new dataset is tested. In essence, cross-validation tests a model's ability to predict new data that was not used in estimating it. Here we used leave-one-out (LOO) cross-validation which is a particular case of leave-n-out cross-validation, where $n = 1$ (Celisse, 2014; James, Witten, Hastie, & Tibshirani, 2014; Stone, 1974).

In our case, we wanted to ask how well the DMS and SATAQ PCA models, that were calculated from all 42 participants (see Figure 4), predicted the raw Daz ideal models that participants created. To do this, we first took an individual’s score on the DMS/SATAQ and used our “PCA + regression” approach to calculate what factor scores this corresponded to for each of the 4 PCs, using the full PCA model (i.e., with all 42 participants). We then took the 4 PC scores for each individual, and reconstructed their ideal body based on their DMS/SATAQ score. Next, we calculated the Euclidian distance between every corresponding pair of points on each individual’s ideal Daz model and the DMS/SATAQ reconstruction. Finally, we averaged the 42 differences for every corresponding pair of points, colour coded them, and plotted them on a representative body. This can be seen in the first two columns of Figure 6. Small to large differences are illustrated by the colour scale changes from blue, through green, to yellow. It is clear that the differences between the ideal Daz and DMS/SATAQ models are greatest in the arm, the abdomen and particularly the shoulder areas. However, even the largest differences are only of the order of ~1.7cm, which we argue suggests good model fits.

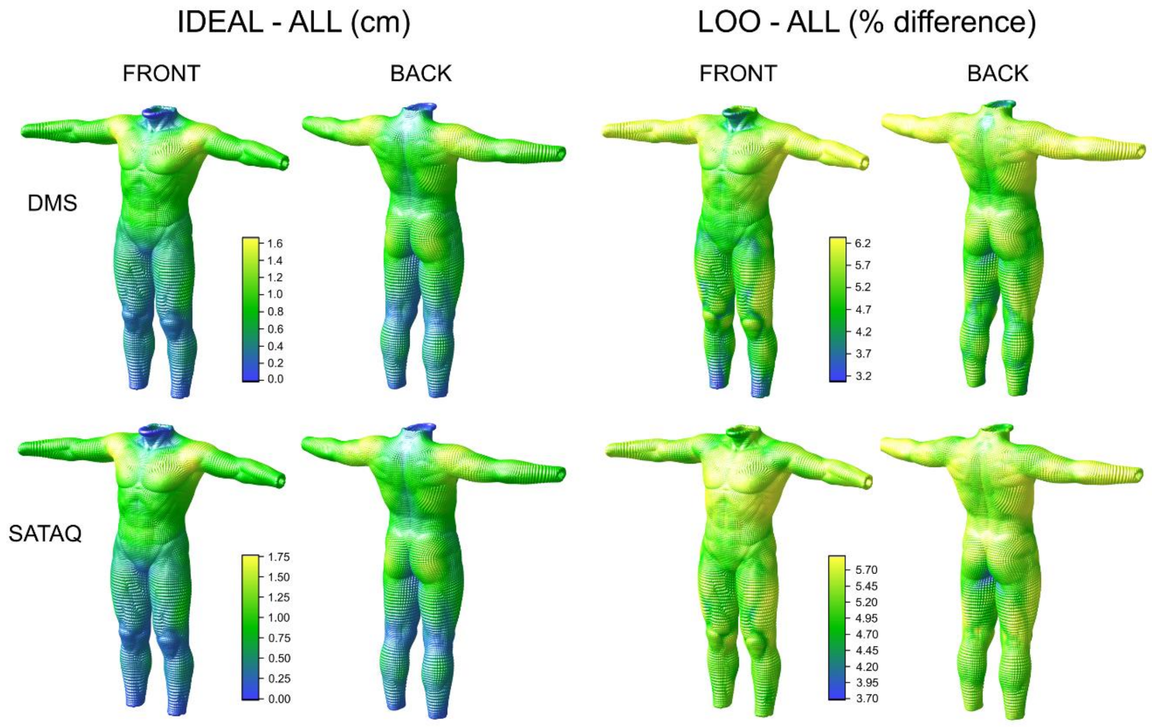


Figure 6. Columns 1 and 2 show the mean differences, averaged across 42 participants, between the raw ideal Daz models and the full PCA + regression model using all 42 participants. Data for the DMS appear in the upper row. Data for the SATAQ Global appear in the bottom row. Smaller to larger differences (centimetres) at each point on the body are colour coded from blue, through green, to yellow. Columns 3 and 4 compare the average percentage error between the LOO PCA + regression models and the full PCA + regression model which included all 42 participants.

To assess, how well the PCA models generalized, we re-iterated the above procedure for 42 LOO PCA models. Specifically, we left out the first of 42 participants and re-calculated separate PCA models for the DMS and SATAQ, and then calculated the differences between this first knock-out model and the full PCA model including all 42 participants. We then re-iterated this procedure, leaving out the second, third, fourth, and so on participant, each time comparing the full PCA model with the LOO model. We then calculated the average percentage error between the full model and all 42 LOO models, and plotted these differences as shown in columns 3 and 4 of Figure 6. Again, the fact that the percentage error between the LOO and full models never exceeds ~6%, on average, suggests that the PCA/SATAQ models would generalize to the UK White male population at large.

4. Discussion

In this study, we examined whether ideal male body size and shape attributes, usually captured by psychometric questionnaires, can be visualised directly in 3D. To explore this idea, we asked male participants to create their ideal body size and shape by manipulating a 3D CGI model using 18 preselected body morphs. This approach allowed participants to capture and visualise appearance ideals relating to specific areas of a 3D body, as opposed to whole body size and shape variation that is often relied upon when using linear and matrix-style 2D figure scales. We also asked participants to fill in standard psychometric questionnaires that assessed their drive for muscularity, attitudes toward and appreciation for their body, and sociocultural influences. We carried out a principal component analysis to extract four uncorrelated 3D shape components that accounted for approximately 95% of the shape variability in the CGI models and we used a combination of correlation and multiple regression to identify the statistical associations between the psychometric questionnaires and the 4 shape components.

Visual inspection suggested that PC1 and particularly PC2 were characterised by changes in the overall bulk of the CGI models. PC3 very clearly captured parallel changes in muscle bulk and definition. PC4 captured a visually striking “trade-off” between very clearly defined muscle of relatively low mass at one extreme, compared to extensive muscle mass especially in the arms and thighs, but with relatively poor muscle definition. The correlation analysis showed that the BAS-2 was statistically significantly associated with PC2, DMS with PC2 and PC3, PSPS with PC1, and SATAQ Global with PC3. This is important because it suggests that there is no simple one-to-one mapping between the shape features extracted from the ideal bodies (indexed by the PCs) and our psychometric measures; the process is Daedalian. Nevertheless, as a proof of concept, the fact that significant correlations exist at all suggests that some mapping is possible.

Another problem is that by presenting the outcomes separately for each PC, while statistically legitimate (they are perfectly uncorrelated with each other), trying to interpret the outcomes across PCs is challenging because the 3D shape features are not easily combined in the mind's eye. It is also important to point out that PC4, in particular, was not associated with any of the behavioural tasks we administered in this study. This may not be surprising because there is no necessary reason why a blind feature decomposition should generate components that are either biologically or psychologically plausible. It is also feasible that PC4 happened to capture a complex trade-off between muscle bulk and muscle definition of the kind that may never have occurred to the designers of the DMS or SATAQ. However, by the same token, had we made a different selection of psychometric tasks, variance in PC4 may well have been captured.

To resolve these difficulties with interpretation, for our final analysis we focused our attention on the DMS and the SATAQ (Global) because these two tasks showed the strongest overall associations with the PCs. We asked the following: (a) how do separate shape features change in relation to each psychometric task? and (b) what body shape corresponds to a particular score on the DMS and the SATAQ questionnaires? As the reconstructions in Figure 4 show, using the "PCA + regression" approach, we found that moving from the lowest to the highest score on both of these tasks corresponds with changes in muscularity, particularly muscle definition. However, perhaps the most interesting feature of this result was that while these effects are, qualitatively, visible on inspection, both are nevertheless quite subtle quantitatively. An obvious, if mundane, explanation for this might be that the CGI modelling method we have used is simply not capable of generating more exaggerated effects that succeed in capturing men's desired ideal. However, the potential shape variation of which our stimulus set up was capable demonstrates that this is unlikely to be the case (see Figure 1). An alternative possibility might have been that we did not observe substantial variability in psychometric task performance, but this was also not the case. Therefore, we suggest that there may be another intriguing possibility. In the DMS, for example, participants respond on a never-to-always scale to verbal statements like “I wish that I were more muscular”, “I lift weights to build up muscle”, “Other people think I work out with weights too often”, and “I think about taking anabolic steroids”. It is possible that phrases like this might conjure up images in the mind not only of the participant, but also of the test interpreter, that are quite extreme precisely because they exist in an imaginary world, and are not necessarily constrained by the physical world. In a similar fashion, it could be that the exaggerated, unattainable muscle patterns represented in action figures and computer game characters say more about the imaginations of their creators than about a direct translation of men’s body ideals into physical form. Therefore, assuming that the method we have developed here is approximately correct and accurate, then what we may have achieved, and what may therefore be of value, is to give some quantitative calibration of the real world extent of men’s ideal body shapes as described by the DMS and SATAQ. An interesting corollary of this is that what may appear to be statistically large effect sizes in the psychometric domain, may amount to quite modest changes when translated into the physical world of anatomy, at least as far as any statistical central tendency is concerned. In addition, although beyond the scope of the current paper, if future research gives rise to validated and normed “visualized” versions of the DMS and SATAQ scales, we could reverse engineer the shape of a given man’s ideal body shape by combining psychometric scores to re-create it.

Another useful outcome from the present study is that we have achieved a finer degree of anatomical precision than can be inferred from the DMS and SATAQ psychometric tasks. In the DMS, the statements most specifically related to the desired locations for changes in muscle mass are: “I think that my arms are not muscular enough”, “I think that my chest is not muscular enough”, and “I think that my legs are not muscular enough”. Similarly, statements from the SATAQ include: “I’ve felt pressure from TV or magazines to look muscular”, “I would like my body to look like the models who appear in magazines”, and “I try to look like sports athletes”. By comparison, as the analysis of content validity showed in Figure 5, we can localize these desires more specifically to the shoulders, the pectoralis muscles, the latissimus dorsi, the abdominal muscles, and the gluteus maximus.

4.1 Limitations and future directions

In this study, the Daz avatars were White in their ethnic appearance and derivation. Distinct anthropometric differences between different ethnic groups have been shown, and some of these differences might be incorporated into ideal body shapes within different ethnic groups, assuming the stimulus set up is suitable. For example, ethnic differences in body weight distribution and the association between BMI and percentage body fat have been established, as well as different patterns of skeletal muscle mass across the lifespan (Abe et al., 2012; Shiwaku et al., 2004; Silva et al., 2010). Clearly, future studies would need to incorporate shape morph controls that allow for such inter-group variation to be included in the resultant models. In a similar vein, it might be valuable to ask how far participants might either want to, or be prepared to, change their body size/shape; what would be the upper limit for what is both desirable and achievable, in the participant's view? Another limitation of this study is that there was no measure of participants' socioeconomic status, despite previous evidence of it being a contributing factor to individual levels of body dissatisfaction, as measured by psychometric questionnaires (Story, French, Resnick, & Blum, 1995; Swami et al., 2010). Future studies might make a virtue of potential ethnographic differences and use substantially larger participant samples to determine whether there may be meaningfully different ideals between different ethnic groups, and what socio-cultural attributes might drive these.

In this study, all our participants carried out the different tasks in the same order and were not counter-balanced. As a result, it is possible that the concerns indexed on the psychometric measures, which were answered last, were magnified by the focus on body image. Future studies could address this concern although it might be argued that any study which has recruited participants for a body perception experiment may cause those participants to be thinking more about their bodies than usual.

This study used psychometric measures which focussed primarily on muscularity and bulk, and future studies may want to look at measures targeting other alternative physical dimensions and psychological concerns such as leanness and thinspiration where the drive is to lose physical bulk and become thin but toned. A drive for muscularity is not qualitatively the same as the drive for thinness which is associated with body dissatisfaction and disordered eating (McCreary & Sasse, 2000). Men with body image disorders seem to show a split between those who strive for thinness and those who seek to increase muscularity (Adams, Turner, & Bucks, 2005; Barlett, Vowels, & Saucier, 2008; McCabe & Ricciardelli, 2004).

Finally, this approach may be of particular interest in visualising the ideal body shape of specific groups, such as body builders or steroid abusers whose ideals may be far more extreme than those seen here, and may indeed use the full range of what is available in 3D body shape. More specifically, it is possible that running similar studies in such groups might reveal effects of measurement equivalence and range restriction. In the former case, a comparison between non-athletes and body builders might show that both groups use the same range of scores on the DMS task, but if visualized, there might be a step change such that levels of muscularity sought by body builders might substantially exceed those sought by non-athletes. It may also be very useful to apply similar techniques to clinical groups, such as people with an eating disorder or body dysmorphia, to achieve a more complete understanding of their concerns and the body size and shape they are attempting to achieve. This method could also be used to visualise body ideals prior to their treatment to provide a starting point on which to focus therapy and allow a targeted intervention on a patient's specific body concerns. Furthermore, the periodic application of the psychometric questionnaires and visualisation procedure would allow a way of visualising progress during therapy.

Acknowledgements

PLC dedicates this paper to the memory of Douglas Leonard Cornelissen (17/1/1931 – 26/3/2021) who had no problem with his body image, at all.

References

Abe, T., Bemben, M. G., Kondo, M., Kawakami, Y., & Fukunaga, T. (2012). Comparison of skeletal muscle mass to fat-free mass ratios among different ethnic groups. *The Journal of Nutrition, Health & Aging*, *16*(6), 534-538. https://doi.org/10.1007/s12603-012-0015-2

Adams, G., Turner, H., & Bucks, R. (2005). The experience of body dissatisfaction in men. *Body Image*, 2, 271–283. https://doi: 10.1016/j.bodyim.2005.05.004.

Arkenau, R., Vocks, S., Taube, C.O., Waldorf, M., & Hartmann, A.S. (2020). The Body Image Matrix of Thinness and Muscularity—Male Bodies: Development and validation of a new figure rating scale for body image in men. *Journal of Clinical Psychology*, *76*, 1283–1292. https://doi.org/10.1002/jclp.22933

Baghurst, T., Hollander, D. B., Nardella, B., & Haff, G. G. (2006). Change in sociocultural ideal meal physique. An examination of past and present action figures. *Body Image*, 3, 87–91. https://doi.org/10.1016/j.bodyim.2005.11.001

Barlett, C. P., Vowels, C. L., & Saucier, D. A. (2008). Meta-analyses of the effects of media images on men's body-image concerns. Journal of Social & Clinical Psychology, 27, 279–310. https://doi: 10.1521/jscp.2008.27.3.279

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science*, *40*(1), 202-223. https://doi.org/10.1111/cogs.12231

Cafri, G., & Thompson, J. K. (2004). Measuring male body image: A review of the current methodology. *Psychology of Men & Masculinity*, *5*, 18–29. https://doi.org/10.1037/1524-9220.5.1.18

Cafri, G., Thompson, J. K., Ricciardelli, L., McCabe, M., Smolak, L., & Yesalis, C. (2005). Pursuit of the muscular ideal: Physical and psychological consequences and putative risk factors. *Clinical Psychology Review*, *25*, 215–239. https://doi.org/10.1016/j.cpr.2004.09.003

Celisse, A. (2014) Optimal cross-validation in density estimation with the L^2-loss. *Annals of Statistics, 42*(5), 1879-1910. (October 2014). https://doi.org/10.1214/14-AOS1240

Collins, M., Zhang, J., Miller, P., Wang, H., & Zhou, H. (2010). EigenBody: Analysis of body shape for gender from noisy images. In M. J. Leahy & M. -L. O'Connell (Eds.), *14th International Machine Vision and Image Processing Conference* (pp. 131-144). Cambridge Scholars Publishing.

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

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

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

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

Frederick, D. A., Fessler, D. M. T., & Hasselton, M. G. (2005). Do representations of male muscularity differ in men's and women's magazines. *Body Image*, *2*, 81–86. https://doi.org/10.1016/j.bodyim.2004.12.002

Gardner, R. M., Friedman, B. N., & Jackson, N. A. (1998). Methodological concerns when using silhouettes to measure body image. *Perceptual and Motor Skills, 86*(2), 387-395. https://doi.org/10.2466/pms.1998.86.2.387

Gardner, R. M., Stark, K. I. M., Jackson, N. A., & Friedman, B. N. (1999). Development and validation of two new scales for assessment of body-image. *Perceptual and Motor Skills, 89*(3), 981-993. https://doi.org/10.2466/pms.1999.89.3.981

Grafen, A., & Hails, R. (2002). *Modern statistics for the life sciences*. Oxford University Press.

Grogan, S., & Richards, H. (2002). Body image: Focus groups with boys and men. *Men and Masculinities*, *4*(3), 219-232. https://doi.org/10.1177/1097184X02004003001

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

Hill, C. E., Thompson, B. J., & Nutt-Williams, E. (1997). A guide to conducting consensual qualitative research. *Counseling Psychologist*, *25*, 517–572. https://doi.org/10.1177/0011000097254001

Hughes, E. K., Dean, C., & Allen, J. S. (2016). Measures of eating disorder symptoms, drive for muscularity, and muscle dysmorphia: Norms and typologies of Australian men. *Australian Journal of Psychology*, *68*, 270–280. https://10.1111/ajpy.12105

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

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2014) *An introduction to statistical learning: With applications in R*. Springer Publishing Company, Incorporated. https://doi.org/10.1007/978-1-4614-7138-7

Karazsia, B. T., & Crowther, J. H. (2008). Psychological and behavioral correlates of the SATAQ-3 with males. *Body Image*, *5*(1), 109-115. https://doi.org/10.1016/j.bodyim.2007.08.004

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

Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition, 172*, 46-58. https://doi.org/10.1016/j.cognition.2017.12.005

Kramer, R. S. S., Young, A. W., Day, M. G., & Burton, A. M. (2017). Robust social categorization emerges from learning the identities of very few faces. *Psychological Review, 124*(2), 115-129. https://doi.org/10.1037/rev0000048

Lanzieri, N., & Cook, B. J. (2013). Examination of muscularity and body fat depictions in magazines that target heterosexual and gay men. *Body Image, 10*, 251–254. https://doi.org/10.1016/j.bodyim.2012.12.003

Leit, R. A., Gray, J. J., & Pope, H. G. (2002). The media's representation of the ideal male body: A cause for muscle dysmorphia? *International Journal of Eating Disorders*, *31*, 334–338. https://doi.org/10.1002/eat.10019

Martins, N., Williams, D. C., Ratan, R. A., & Harrison, K. (2011). Virtual muscularity: A content analysis of male video game characters. *Body Image*, *8*, 43–51. https://doi.org/10.1016/j.bodyim.2010.10.002

McCabe, M. P., and Ricciardelli, L. A. (2004). Body image dissatisfaction among males across the lifespan: a review of past literature. *Journal of Psychosomatic Research,* 56, 675–685. https://doi: 10.1016/S0022-3999(03)00129-6

McCreary, D. R., & Sasse, D. K. (2000). An exploration of the drive for muscularity in adolescent boys and girls. *Journal of American College Health*, *48*(6), 297-304. https://doi.org/10.1080/07448480009596271

McCreary, D. R., Hildebrandt, T. B., Heinberg, L. J., Boroughs, M., & Thompson, J. K. (2007a). A review of body image influences on men's fitness goals and supplement use. *American Journal of Men's Health*, *1*, 307–316. https://doi.org/10.1177/1557988306309408

McCreary, D. R. (2007b). The Drive for Muscularity Scale: Description, Psychometrics, and Research Findings. In *The muscular ideal: Psychological, social, and medical perspectives.* (pp. 87-106). American Psychological Association.

McPherson, K.E., McCarthy, P., McCreary, D.R., & McMilland, S. (2010). Psychometric evaluation of the Drive for Muscularity Scale in a community-based sample of Scottish men participating in an organized sporting event. *Body Image*, *7*(4), 386-371. https://doi.org/10.1016/j.bodyim.2010.06.001

Murray, S.B., Nagatab, J.M., Griffith, S., Calzod, J.P., Brown, T.A., Mitchison, D., Blashill, A.J., & Mond, J.M. (2017). The enigma of male eating disorders: A critical review and synthesis. *Clinical Psychology Review*, *57*, 1-11. https://doi.org/10.1016/j.cpr.2017.08.001

Osborne, M. R., Presnell, B., & Turlach, B. A. (2000). A new approach to variable selection in least squares problems. *IMA Journal of Numerical Analysis*, 20, 389–404.

Pope, H. G., Gruber, A. J., Mangweth, B., Bureau, B., deCol, C., Jouvent, R., & Hudson, J. I. (2000). Body image perception among men in three countries. *American Journal of Psychiatry*, *157*, 1297–1301. https://doi.org/10.1176/appi.ajp.157.8.1297

Pope, H. G., Phillips, K. A., & Olivardia, R. (2000). *The Adonis complex: The secret crisis of male body obsession.* The Free Press.

Ralph-Nearman, C. & Filik, R. (2018). New body scales reveal body dissatisfaction, thin-ideal, and muscularity-ideal in males. *American Journal of Men's Health, 12(*4), 740-750. https://doi.org/10.1177/1557988318763516

Ridgeway, R. T., & Tylka, T. L. (2005). College men's perceptions of ideal body composition and shape. *Psychology of Men and Masculinity*, *6*, 209–220. https://doi.org/10.1037/1524-9220.6.3.209

Ruto, A., Lee, M., & Buxton, B. (2006, September). *Comparing principal and independent modes of variation in 3D human torso shape using PCA and ICA*. Paper presented at the ICA Research Network International Workshop, Liverpool, UK.

Shiwaku, K., Anuurad, E., Enkhmaa, B., Kitajima, K., & Yamane, Y. (2004). Appropriate BMI for Asian populations. *The Lancet*, *363*(9414), 1077. https://doi.org/10.1016/S0140-6736(04)15856-X

Silva, A. M., Shen, W., Heo, M., Gallagher, D., Wang, Z., Sardinha, L. B., & Heymsfield, S. B. (2010). Ethnicity-related skeletal muscle differences across the lifespan. *American Journal of Human Biology: The Official Journal of the Human Biology Association*, *22*(1), 76-82. https://doi.org/10.1002/ajhb.20956

Smith, K.L., Tovée, M.J., Hancock, P., Cox, M. & Cornelissen, P.L. (2007). An analysis of body shape attractiveness based on image statistics: evidence for a dissociation between expressions of preference and shape discrimination. *Visual Cognition, 15*(7), 1-27. https://doi.org/10.1080/13506280601029515

Stice, E., Nemeroff, C., & Shaw, H. E. (1996). Perceived Sociocultural Pressure Scale (PSPS). APA PsycTests. https://doi.org/10.1037/t04547-000

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

Stice, E., Nemeroff, C., & Shaw, H. E. (1996) Test of the dual pathway model of bulimia nervosa: Evidence for dietary restraint and affect regulation mechanisms. *Journal of Social and Clinical Psychology,* 15(3), 340-363. https://doi.org/10.1521/jscp.1996.15.3.340

Stone, M. (1974). An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), 44-47.

Story, M., French, S. A., Resnick, M. D., & Blum, R. W. (1995). Ethnic/racial and socioeconomic differences in dieting behaviors and body image perceptions in adolescents. *International Journal of Eating Disorders*, *18*(2), 173-179. https://doi.org/10.1002/1098-108X(199509)18:2<173::AID-EAT2260180210>3.0.CO;2-Q

Swami, V., Frederick, D. A., Aavik, T., Alcalay, L., Allik, J., Anderson, D., ... & Danel, D. (2010). The attractive female body weight and female body dissatisfaction in 26 countries across 10 world regions: Results of the International Body Project I. *Personality and Social Psychology Bulletin*, *36*(3), 309-325. https://doi.org/10.1177/0146167209359702

Talbot, D., Smith, E., & Cass, J. (2019). Male body dissatisfaction, eating disorder symptoms, body composition, and attentional bias to body stimuli evaluated using visual search. *Journal of Experimental Psychopathology*, *10*(2), 1-13. https://doi.org/10.1177/2043808719848292

Thompson, J.K., van den Berg, P., Roehrig, M., Guarda, A.S., & Heinberg, L.J. (2004). The Sociocultural Attitudes Towards Appearance Scale-3 (SATAQ-3): Development and validation. *International Journal of Eating Disorders*, *35*(3), 293-302. https://doi.org/10.1002/eat.10257

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

Tylka, F. L., & Wood-Barcalow, N. L. (2015) The Body Appreciation Scale-2: Item refinement and psychometric evaluation. *Body Image*, 12, 53-67. https://doi.org/10.1016/j.bodyim.2014.09.006