Evaluating the Convergence of Muscle Appearance Attitude Measures

Guy Cafri
J. Kevin Thompson
University of South Florida

There has been growing interest in the assessment of a muscular appearance. Given the importance of assessing muscle appearance attitudes, the aim of this study was to explore the convergence of the Drive for Muscularity Scale, Somatomorphic Matrix, Contour Drawing Rating Scale, Male Figure Drawings, and the Muscularity Rating Scale. Participants were 76 men and 103 women who each completed four measures of body image that were meant to evaluate two distinct domains of appearance attitudes: muscularity and body fat. For both genders, data indicated low convergence among muscularity measures, but high convergence among body fat measures. For men, the Drive for Muscularity Body Image subscale exhibited the greatest concurrent validity with behaviors used to increase muscularity. The results are discussed in light of the selection of methods for the assessment of muscle appearance attitudes.

Keywords: men; body image; assessment; muscularity; validity

Body image research has traditionally focused on women and the dissatisfaction associated with being overweight (J. K. Thompson, Heinberg, Altabe, & Tantleff-Dunn, 1999). Researchers have become interested in a dimension of body image related to a muscular appearance (e.g., Cafri, Strauss, & Thompson, 2002). The importance of research on the assessment of the muscularity construct has predominantly been fueled by research evidence indicating that a significant number of men experience muscle dissatisfaction, and a variety of health risk behaviors are associated with this dissatisfaction (viz., steroid use, ephedrine use, and dieting behaviors; Cafri & Thompson, 2004; McCabe & Ricciardelli, 2004; Pope, Phillips, & Olivardia, 2000). Notably, evidence also suggests that a significant number of women report muscle dissatisfaction, with 57% of Psychology Today readers indicating muscle tone dissatisfaction (Garner, 1997). Directly related to the mushrooming of research in this area is the development of a significant number of methods to evaluate muscle appearance attitudes. Although several measures have been utilized, no study has examined the relationships among the measures, the extent to which they assess the same construct, and their relationship to behaviors related to the pursuit of muscularity. The current study will help determine which measures are appropriate in the assessment of a muscular body image.

The two most conventional methods of assessing appearance attitudes are Likert-type scale measures and contour drawn silhouette scales (Thompson et al., 1999). With respect to research examining a thin body ideal, the Body Dissatisfaction subscale of the Eating Disorders Inventory (EDI-BD; Garner, Olmstead, & Polivy, 1983) is a commonly used Likert-type scaled measure, while figures adapted from Stunkard, Sorensen, and Schulzinger (1983) have historically been the most widely used silhouette measure. For silhouette scales, participants typically select their self and ideal bodies from a range of scaled figures and the difference between these two ratings is interpreted as an index of dissatisfaction (Thompson, 1990).
Research examining a muscular ideal has adopted methods of assessment similar to those used in examining a thin ideal. For instance, McCreary and Sasse (2000) developed the Drive for Muscularity Scale, which uses a Likert-type scale format that is identical to the one used on the EDI-BD. More important, a recent exploratory factor analysis of the scale showed that items could be separated into Body Image and Behaviors subscales for men but not women, which suggest that in addition to using a broad index of the drive-for-muscularity, specific factors can be assessed as well (McCleary, Sasse, Saucier, & Dorsch, 2004). Lynch and Zellner (1999) developed a silhouette scale that has nine body figures arranged in a manner similar to those figures from Stunkard et al. (1983), however the figures vary in terms of increasing muscularity rather than adiposity. A similar silhouette measure titled the Muscularity Rating Scale was developed for women (Furnham, Titman, & Sleeman, 1994). Yet another silhouette rating measure is the Somatomorphic Matrix, which is a computer-based protocol that consists of separate figures for men and women (Gruber, Pope, Borowiecki, & Cohane, 1999). For each gender this procedure consists of a matrix of images that simultaneously vary in degree of muscularity and body fat, but scaled on two separate axes, thus allowing for a dissatisfaction index to be determined separately for muscularity and body fat. Given the methodological variability in the assessment of a muscular body image, it is important to determine the extent to which measures assessing a muscular body image converge or diverge within each gender.

An often-neglected theoretical point is that degree of muscularity is not the only determinant of a muscular appearance; body fat is also relevant because less body fat enhances an appearance of muscularity. Given this point, a comprehensive assessment of a muscular appearance should consist of measurement of muscularity and body fat (Cafri & Thompson, 2004). This line of reasoning supports the use of measures to evaluate body fat attitudes in conjunction with muscularity attitudes when the intention is to measure muscle appearance attitudes. The Body Fat Dissatisfaction index of the Somatomorphic Matrix would be one way to assess body fat dissatisfaction. Another way may be through the use of silhouette measures used to evaluate a thin appearance (e.g., Contour Drawn Rating Scale; M. A. Thompson & Gray, 1995). Again, it would be interesting to examine the extent to which these body fat measures converge given their methodological difference.

The current study had several objectives. The first goal was to examine the convergence among measures assessing a muscularity component of body image separately for men and women. Second, we sought to examine the convergence among body fat measures separately for men and women. Third, given the availability of a measure of behaviors related to the pursuit of muscularity for men (i.e., the Drive for Muscularity Behaviors subscale), there was interest in determining the extent to which each of the measures used to assess male body image uniquely predicted behaviors reflecting the pursuit of muscularity.

**METHOD**

**Participants**

Participants were 76 men and 103 women from the University of South Florida. Participants were sampled from an undergraduate psychology department participant pool. The students were responders to an advertisement offering course credit for participation in the study. Male participants’ ages ranged from 18 to 29 years ($M = 21.12$, $SD = 2.60$) and female participants’ ages ranged from 18 to 29 years ($M = 20.81$, $SD = 2.48$). Racial composition of male participants was 52.6% White, 18.4% African American, 17.1% Hispanic American, 7.9% Asian American, and 2.6% Native American. Racial composition of female participants was 44.7% White, 20.4% African American, 19.4% Hispanic American, 5.8% Asian American, 1.0% Native American, and 1.9% other.

**Materials**

**Somatomorphic Matrix.** This is a bidimensional computerized body image test that can assess satisfaction with respect to muscularity and body fat (Gruber et al., 1999). For each gender, the test consists of 100 images arranged in a $10 \times 10$ matrix, representing 10 degrees of adiposity and 10 degrees of muscularity. To answer an item, participants scroll through the library of images by using buttons on the screen to either add or subtract muscularity and body fat. Similar to other silhouette measures, the difference between ratings of one’s own body and ideal body is interpreted as an index of dissatisfaction. In support of the measure’s validity are data indicating that the images used correspond to particular fat-free mass index (FFMIs) and body fat percentages (BF%;s) (Gruber et al., 1999). This was done by photographing people with known FFMIs and BF% determined by skin-fold measurements with calipers, then having a graphic artist develop these into drawings (Gruber et al., 1999). Low test-retest values of satisfaction indexes have been found for the measure for college undergraduate men and women (Cafri, Roehrig, & Thompson, 2004). Specifically, for male ratings on the Somatomorphic Matrix, the reliabilities were $r_{xx} = .34$, for self-ideal muscularity discrepancy and $r_{xx} = .57$ for self-
ideal body fat discrepancy, and for women, self-ideal muscularity discrepancy was $r_{\alpha} = .35$ and self-ideal body fat discrepancy was $r_{\alpha} = .56$.

**Drive for Muscularity Scale.** This is a 15-item measure with a 6-point response format that assesses attitudes and behaviors related to a muscular appearance (McCreary & Sasse, 2000). For men, but not women, it was found that the scale could be factor analyzed into two lower order factors, body image and behaviors (McCreary et al., 2004). High internal consistencies have been found for both subscales (Body Image: Cronbach’s alpha = .88, Behavior: Cronbach’s alpha = .81; McCreary et al., 2004). The total score on the scale is also internally consistent for men (.83) and women (.78; McCreary & Sasse, 2000). In the current study, the internal consistencies of the total score for men (Cronbach’s alpha = .89) and women (Cronbach’s alpha = .81) were high, as were the body image (Cronbach’s alpha = .88) and behavior (Cronbach’s alpha = .86) factors for men. For the male portion of the current study, we used the Body Image subscale to test convergence with the other body image measures. In contrast, we used the Behaviors subscale to determine how well a muscular body image correlated with an outcome measure assessing relevant behaviors (steroid use, weightlifting, etc.). The items on the scale were coded in such a way that lower scores reflected higher drive for muscularity.

**Contour Drawing Rating Scale.** The scale contains separate silhouette scales for men and women. For each gender there are nine figures arranged in terms of increasing body fat on a 9-point scale (M. A. Thompson & Gray, 1995). Each participant rated the gender-specific set. The scale is well validated in that data indicate body fat differences between the figures are consistently discernable, and self-ratings are highly correlated with self-reported weight (M. A. Thompson & Gray, 1995). Furthermore, the scale has reliable test-retest ratings of current body size among college undergraduate men and women ($r = .78$, $n = 40$).

**Male Figure Drawings.** This is a silhouette scale containing nine male figures arranged in terms of increasing muscularity on a 9-point scale (Lynch & Zellner, 1999). The scale has acceptable test-retest reliabilities for current ($r = .88$) and ideal ($r = .71$) ratings (Lynch & Zellner, 1999).

**Muscularity Rating Scale.** This is a silhouette scale containing nine female figures arranged in terms of increasing muscularity on a 9-point scale (Furnham et al., 1994). No validity or reliability data are currently available.

**Procedure**

The measures were completed in classroom and lab settings. Male participants completed all the measures except the Muscularity Rating Scale (Furnham et al., 1994), and women completed all measures except the Male Figure Drawings (Lynch & Zellner, 1999). The order of the measures was counterbalanced to control for order effects. Given that there were four body image measures, four different orders were used. A process of random selection determined the order of the measures with the selection criterion that none of the four groups could have the same measure in the same rank order. For each silhouette-based measure, participants responded to two items: (a) Please choose the image that most closely resembles your own body and (b) Please choose the image that you would ideally like to have. To control for order effects within the silhouette measures, the self and ideal ratings items were counterbalanced as well. A subsample of 27 men completed the Drive for Muscularity Scale twice with an interval of 7 to 10 days to allow for an assessment of test-retest reliability.

**Analyses**

Correlations were used to assess convergence and test-retest reliability. Values of 0.70 or greater were interpreted as indicators of adequate test-retest reliability (Nunnally, 1978). On the muscularity dimension, the correlations to be assessed were among the Somatomorphic Matrix Muscularity index, the Drive for Muscularity Scale (Body Image subscale for men and Total Score for women), and either the Male Figure Drawings (for men) or Muscularity Rating Scale (for women). On the body fat dimension, the correlation assessed was between the Somatomorphic Matrix Body Fat index and the Contour Drawing Rating Scale.

In determining the extent to which each of the measures used to assess male body image uniquely predicted behaviors reflecting the pursuit of muscularity, the Drive for Muscularity Scale Behaviors subscale was simultaneously regressed onto the Drive for Muscularity Body Image subscale, Somatomorphic Matrix Muscularity Dissatisfaction index, and the Somatomorphic Matrix Body Fat Dissatisfaction index. Prior to running this analysis, outliers were screened (using studentized residuals and Cook’s D) and the assumptions of the linear regression model (normality, homoscedasticity, collinearity) were tested. To determine the contribution of each predictor, statistical significance was determined through a $t$ test of the regression coefficients, and standardized regression coefficients and squared semipartial correlations were examined.
RESULTS

Table 1 presents means and standard deviations for the four measures administered to men and women. One aspect of this table that is notable is that men had a lower mean score on the Drive for Muscularity than women, which signifies a higher drive for muscularity for men. A t-test comparing men and women indicates a statistically significant difference, $t(174) = 6.54, p < .05$. Table 2 presents the correlations for the measures administered to men. As can be seen in Table 2, there was low to moderate convergence among the body image measures of muscularity: the Drive for Muscularity Body Image subscale, the Somatomorphic Matrix Muscularity index, and the Male Figure Drawings. In contrast, on the body fat dimension, high convergence was observed between the Contour Drawing Rating Scale and the Somatomorphic Matrix Body Fat Dissatisfaction index.

A pattern of results emerged among women similar to which was found among men. Table 3 presents the correlations for the measures administered to women. On the muscularity dimension, relatively low convergence was observed among the measures: Drive for Muscularity total score, Somatomorphic Matrix Muscularity index, and the Muscularity Rating Scale. On the body fat dimension however, high convergence was observed between the Contour Drawing Rating Scale and the Somatomorphic Matrix Body Fat Dissatisfaction index.

To examine the extent to which each of the measures used to assess male body image uniquely predicted behaviors reflecting the pursuit of muscularity, the Drive for Muscularity Scale Body Image subscale (DMSBI), Somatomorphic Matrix Muscularity Dissatisfaction index (SMMUS), and the Somatomorphic Matrix Body Fat Dissatisfaction index (SMBF) were used to predict scores on the Drive for Muscularity Behaviors subscale (DMSBE). No severe violations of assumptions were found, except the reliabilities of the Somatomorphic Matrix indexes may be low enough to warrant some concern. The obtained $R^2$ value was .26 and the Wherry adjusted $R^2$ value was .22 (root $MSE = 6.54$). The DMSBI was a significant predictor of the DMSBE, $t(73) = 4.92, p < .05$, however neither the SMMUS, $t(73) = -1.69, p = .10$, nor the SMBF, $t(73) = -1.85, p = .07$, were significant. The standardized regression coefficient for DMSBI, SMMUS, and SMBF were .51, -.17, -.09, respectively. Squared semipartial correlations indicated that of the variance of DMSBE scores, 24.7% was uniquely accounted by the DMSBI, 3% was accounted for by the SMMUS, and .01% was accounted for by SMBF.

Test-retest reliabilities were computed for men on the Drive for Muscularity Scale. The overall measure was highly reliable (.93), as were the subscales of Body Image (.84) and Behaviors (.96).

DISCUSSION

The current study examined the convergence of several attitudinal measures of body image that could be used in the assessment of a muscular appearance, as well as the test-retest reliability of the Drive for Muscularity Scale. With respect to the convergence of measures, the results were quite surprising, but consistent across gender. On the body fat dimension, there was evidence of high convergence. In contrast, on the muscularity dimension it was found that none of the measures converged very highly. More important, for men, the Body Image subscale of the Drive for Muscularity Scale was the strongest predictor of behaviors related to the pursuit of muscularity, with the Somatomorphic Matrix Muscularity and Body Fat indexes contributing a statistically nonsignificant effect, however, the Muscularity index may account for a proportion of variance that is substantively meaningful. In addition, the overall indices of the Drive for Muscularity Scale demonstrated very high test-retest reliabilities.

The finding that the indices of body fat dissatisfaction converged highly and the muscularity measures converged to a small degree has important implications for future methods of assessment in this area of research. With respect to the two silhouette measures of body fat satisfac-
tion, the results suggest that either measure may be substituted for the other. Notably though, male responses were not separated into those who desired more body fat and less body fat (Cafri et al., 2002), which may have an impact on the calculated correlations. The finding of low convergence among measures assessing a muscularity dimension of muscle appearance attitudes suggests that each measure may be assessing a distinct domain of body image. For instance, for men, the lack of association between muscularity ratings on the silhouette measures and the Drive For Muscularity Body Image subscale may have to do with the fact that the latter is a more direct measure of muscle satisfaction than the former. Specifically, in the silhouette scales, a satisfaction score is achieved by subtracting self and ideal ratings without the participant’s knowledge that his or her degree of satisfaction is being evaluated, which makes this method more latent in its assessment. Moreover, it could be argued that the silhouette scales assess exclusively a satisfaction component of body image, whereas the Body Image subscale of the Drive For Muscularity assesses satisfaction and other attitudes (e.g., confidence in appearance). For women, the explanation would be the same except that the behavior items of the Drive for Muscularity Scale were included, which might have further contributed to decreased convergence.

Given the low correlation among the muscularity measures for men, we wanted to determine which of the measures had the highest concurrent validity with a criterion of interest to determine which measure was assessing a muscular body image to the greatest extent. The results of the regression analyses provided a means of answering this question. The results indicated that the Drive for Muscularity exhibited the greatest concurrent validity. The Somatomorphic Matrix Muscularity and Body Fat indexes did not exhibit statistically significant prediction after controlling for the Drive for Muscularity Body Image subscale, although the Muscularity index did appear to be a practically significant predictor of behaviors (i.e., it accounted for a substantively meaningful proportion of variance). Such findings are not surprising given that the Dissatisfaction index is equivalent to a single item (i.e., self-ratings are subtracted from ideal ratings to form a single score), therefore the likelihood that it can account for a large proportion of variance is limited. These findings indicate that the Drive for Muscularity Body Image subscale might be the measure best suited for the analysis of behaviors related to the pursuit of a muscular ideal.

It is noteworthy to mention that if the Somatomorphic Matrix is to be used, low reliability of difference scores are a problem (Cafri, Roehrig, & Thompson, 2004). Therefore, an alternative method of analyzing the difference score might be considered, whereby the difference score is not used, and instead self, ideal, self-by-ideal interaction, and any appropriate higher order terms are used as predictors in a regression equation (e.g., Edwards & Parry, 1993). The validity of such an approach has been discussed extensively by Edwards (2001) and would eliminate the problem of low difference-score reliability.

Several shortcomings of the current study should be noted. First, although some validation has taken place with the Drive for Muscularity Scale as a whole (e.g., convergent validity with self-esteem and divergent validity with the drive for thinness; McCreary & Sasse, 2000), the Be-

### TABLE 2
Correlation Matrix of Measures Used With Men

<table>
<thead>
<tr>
<th>Scales</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Male figure drawings</td>
<td>—</td>
<td>.15</td>
<td>.12</td>
<td>.05</td>
<td>.37*</td>
<td>.12</td>
<td>.29*</td>
</tr>
<tr>
<td>2. Contour Drawn Rating Scale</td>
<td>—</td>
<td>—</td>
<td>−.02</td>
<td>.59*</td>
<td>.10</td>
<td>−.20</td>
<td>−.05</td>
</tr>
<tr>
<td>3. Somatomorphic matrix–muscularity</td>
<td>—</td>
<td>—</td>
<td>.05</td>
<td>.19</td>
<td>−.10</td>
<td>.08</td>
<td></td>
</tr>
<tr>
<td>4. Somatomorphic matrix–body fat</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.09</td>
<td>—</td>
<td>—</td>
<td>.02</td>
</tr>
<tr>
<td>5. Drive for muscularity–body image</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.48*</td>
<td>.85*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Drive for muscularity–behaviors</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.86*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Drive for muscularity–total</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Correlations based on *n* = 79. *p* < .05 (two-tailed).

### TABLE 3
Correlation Matrix of Measures Used With Women

<table>
<thead>
<tr>
<th>Scales</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Muscularity Rating Scale</td>
<td>—</td>
<td>.26*</td>
<td>.18</td>
<td>.30*</td>
<td>.35*</td>
</tr>
<tr>
<td>2. Contour Drawn Rating Scale</td>
<td>—</td>
<td>−.02</td>
<td>.58*</td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>3. Somatomorphic matrix–muscularity</td>
<td>—</td>
<td>−.16</td>
<td>.20*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Somatomorphic matrix–body fat</td>
<td>—</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Drive for Muscularity–total</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Correlations based on *n* = 101. *p* < .05 (two-tailed).
haviors subscale has not been extensively validated among groups of people that should show differences on this measure. Specifically, there is no research evidence to suggest that athletes, bodybuilders, or those who have clinically significant muscle-related disturbance (viz., muscle dysmorphia; Pope, Gruber, Choi, Olivardia, & Phillips, 1997) score higher on this measure. Given the lack of validation of the Drive for Muscularity Behaviors subscale, its use as an outcome measure may be questioned to some extent. One particular shortcoming of the Behaviors subscale is the lack of items assessing body fat reduction methods, which might explain the weak predictive power of the Body Fat index of the Somatomorphic Matrix. Another source of concern is the presence of low reliabilities of the Somatomorphic Matrix indexes in conjunction with moderate intercorrelations among the predictors. Such a situation may have lead to inconsistent estimates of the regression coefficients (i.e., the coefficients will be biased, however the direction of the bias will be difficult to predict; Bollen, 1989). Given these limitations, the results related to the predictive utility of the body image measures should be treated tentatively. Finally, the extreme masculinity of the figures on the Muscularity Rating Scale may have elicited an aversive response in women that may have influenced its convergence with other measures. The growing interest in the assessment of muscle appearance attitudes, particularly among men, makes it imperative that researchers attend to the methodological issues related to assessment of this construct. Accurate measurement of muscle appearance attitudes is a necessary first step prior to the assessment of this construct’s relationship to relevant outcomes of interest, which include but are not limited to symptoms of body dysmorphic disorder, steroid use, and dieting behaviors.

REFERENCES


Guy Cafri is a doctoral student in the psychology program at the University of South Florida who has published four articles on body image issues in boys and men.

J. Kevin Thompson, Ph.D., is a professor of psychology at the University of South Florida who has edited, coedited, authored, or coauthored five books in the area of body image and eating disorders.