

This is the accepted peer-reviewed version of the following article:

Atkinson, M.J.,* & Diedrichs, P.C. (in press). Assessing the Impact of Body Image Concerns on Functioning Across Life Domains: Development and Validation of the Body Image Life Disengagement Questionnaire (BILD-Q) Among British Adolescents. *Body Image*.

* Corresponding author. E-mail address: m.j.atkinson@bath.ac.uk

**Assessing the Impact of Body Image Concerns on Functioning Across Life Domains:
Development and Validation of the Body Image Life Disengagement Questionnaire
(BILD-Q) Among British Adolescents**

Melissa J. Atkinson^{1,2}, Phillippa C. Diedrichs²

¹Department of Psychology, University of Bath, UK

²Centre for Appearance Research, University of the West of England, UK

Correspondence to: Melissa Atkinson, Department of Psychology, University of Bath, Claverton Downs, Bath, BA27AY, UK; Email: m.j.atkinson@bath.ac.uk; phone: +441225386197.

Sources of Funding: This study was funded by a research grant from the Dove Self-Esteem Project, Unilever. The funders had no role in the study design, data collection and analysis, decision to publish, or manuscript preparation. They were permitted to review the manuscript and suggest changes, but the authors exclusively retained the final decision on content.

Disclosure Statement: PD is an independent research consultant to the Dove Self-Esteem Project global education initiative. The authors declare no other potential conflicts of interest in relation to this work.

Acknowledgments

We would like to thank the research associates who assisted with data collection: Kirsty M. Garbett, Jade Parnell, Hannah Jarman, Georgia Treneman-Evans. We would also like to thank Professor Mitchell Callan for assistance with testing of gender invariance in R.

Abstract

Assessing the impact of body image on engagement in a range of life domains is important; however, there is a lack of validated measures for adolescents. The current research developed the Body Image Life Disengagement Questionnaire (BILD-Q) and validated it among four samples of British adolescents. Study 1 ($N = 1707$; 11-13 years) indicated a 9-item unidimensional scale based on Exploratory Factor Analysis. In Study 2 ($N = 1403$; 11-13 years), Confirmatory Factor Analysis (CFA) showed an acceptable fit overall, but better among girls than boys. Further exploration with CFA in Study 3 ($N = 2034$; 13-14 years) showed a good to excellent fit overall, and acceptable among both boys and girls. The scale showed good internal consistency and test-retest reliability, and gender invariance indicated the scale can be used comparatively. In Study 4 ($N = 288$; 13-14 years), convergent validity was supported via expected relationships with body image and related constructs. Concurrent and predictive incremental validity were also evidenced via explaining significant unique variance in well-being. These studies provide support for the BILD-Q as a reliable and valid measure of broader impacts of body image among adolescents, which may benefit intervention evaluation and policy change efforts.

Keywords: body image; functioning; quality of life; validation; psychometrics; adolescents

1. Introduction

Body image is a multifaceted construct that encompasses affective, cognitive, perceptual and behavioural aspects related to one's body and appearance (Cash & Smolak, 2011). Extensive research has established multiple health correlates and consequences of negative body image (e.g., depression, anxiety, low self-esteem, substance abuse, unhealthy weight control practices and disordered eating; Bornioli et al., 2019; Neumark-Sztainer et al., 2006; Stice, 2002) and positive body image (e.g., health promoting behaviours and greater well-being; Andrew et al., 2015; Tylka & Piran, 2019; Tylka & Wood-Barcalow, 2015). There is some evidence that body image also impacts on other important life domains, such as academic and workplace performance and aspirations, and social functioning (Cash & Fleming, 2002; Halliwell et al., 2014). However, progress in rigorously measuring broader consequences of body image concerns has been hampered by a lack of available validated measures – particularly for adolescents – that focus specifically on impairment to engagement across various life activities, and that can be quickly and easily administered to inform research, practice and policy. The current research outlines the development and psychometric validation of a brief scale to assess the impact of body image concerns on adolescents' participation and engagement in a range of life domains, including education, socialising, sport participation, seeking healthcare, and self-assertion.

1.1. Assessing Broader Life Impacts of Body Image

Considering the significant physical and mental health consequences of body image concerns noted above, along with the high prevalence reported in many countries (Al Sabbah et al., 2009; Frederick et al., 2016; Frederick et al., 2019), and the persistence of stigma related to appearance (Puhl & Peterson, 2012), it is not surprising that body image is increasingly viewed as a significant public health, gender, and social justice issue (Atkinson et al., 2020; Bornioli et al., 2019). This has led to increased recognition across sectors that

body image requires attention and intervention, including from policy makers, governments, politicians, educators, public health officials, community and advocacy groups, members of civil society, and the private sector. This is particularly emphasised for young people, who are vulnerable to body image concerns during an intensive time of identity and bodily development (All Party Parliamentary Group on Body Image, 2012; British Youth Council, 2017). Extensive research has been and is being conducted to understand the development of body image concerns and evaluate evidence-based interventions among young people (Kusina & Exline, 2019; Yager et al., 2013). This has largely focused on body image and its mental and physical health-related predictors, correlates and consequences. Evaluating the impact of body image concerns on life domains *beyond* psychological and physical health functioning, however, could help to establish a more comprehensive understanding of the consequences of body image concerns e.g., how they may be limiting life experiences important for individual development and, ultimately, contribution to society. If broader functioning is indeed shown to be impacted, this understanding could provide useful leverage to advance advocacy efforts, funding agendas, and implementation of interventions (Atkinson et al., 2020; Austin, 2015).

1.2. Existing Measures

Currently, a small number of measures assess broader impacts with relevance to body image, including those related to quality of life, psychosocial functioning and behavioural avoidance. However, no measure exists that focuses *specifically* on a *range* of life domains beyond physical or mental health functioning, is specific to body image but has broad applicability across different appearance concerns, is validated for use with adolescents, *and* is brief and accessible to facilitate wide use.

For example, the *Body Image Quality of Life Inventory* (Cash & Fleming, 2002) assesses negative and positive impact of body image (using a 7-point bipolar scale from -3 to

+3) on various aspects of life, including sense of self, social functioning, sexuality, emotional well-being, eating, exercise, grooming, and overall satisfaction with life (e.g., example items include “My basic feelings about myself—feelings of personal adequacy and self-worth”, “My day-to-day emotions”, “My enjoyment of my sex life”, “My relationships with friends”). It was originally designed for clinical use, and later adapted and validated with university-aged men and women (Cash & Fleming, 2002; Cash & Grasso, 2005; Cash et al., 2004). Intended for and since used widely with adults aged 18 and over, it has not been adapted or validated for use with adolescents, although it has shown internal reliability as an outcome measure in an intervention trial for adolescents with body dysmorphic disorder (Krebs et al., 2017; Mataix-Cols et al., 2015). Notably, the BIQLI includes perceived impact on mental well-being (e.g., self-worth, emotions, happiness and confidence in everyday life) and health behaviours (e.g., control of eating and weight, physical exercise), as well as impact on broader social and work/school functioning, in an overall score. Therefore, it does not capture specific impacts on engagement in life activities beyond physical and mental health functioning. In addition, it is a relatively long measure at 19 items, as well as requiring payment and conditions for use, therefore potentially limiting easy access and widespread adoption.

Within the related field of eating disorders, researchers have assessed broader life impacts in different ways. For example, Stice and colleagues have assessed social functioning (selected items from the Social Adjustment Scale; Weissman & Bothwell, 1976) and healthcare use (purpose-built items) as outcomes of preventive interventions targeting risk factors for eating disorders (including but not limited to body dissatisfaction), among older adolescent and young adult females (Stice et al., 2015; Stice et al., 2011; Stice et al., 2006). While relevant, this is a fragmented approach using different measures that have not been

validated in this context beyond reporting of internal and test-retest reliability in efficacy trials with females, and are also not specific to body image.

The *Clinical Impairment Assessment* (Bohn et al., 2008) is a validated measure of psychosocial functioning related to eating disorders. This 16-item measure was developed and validated as a companion to the assessment of eating disorder features and designed to immediately follow such assessment (i.e., the Eating Disorder Examination Questionnaire; Fairburn & Beglin, 1994). It is therefore primarily used as a tool for assessing Cognitive Behavioural Therapy (CBT) treatment for eating disorders (Fairburn, 2008; Fairburn et al., 2009). It assesses impacts of eating disorder features (i.e., eating habits, exercise, feelings about weight or shape) on mood and self-perception, cognitive functioning, interpersonal functioning and work performance, in a global score. Similar to the BIQLI, it does not enable focused assessment of engagement in activities across a range of life domains separate to psychological functioning. Given its intended use alongside assessment of eating disorder psychopathology, it also does not encompass aspects of appearance other than weight and shape (e.g., skin, hair, muscularity), which limits its applicability for body image more broadly. Validation with adolescents is currently limited to an adapted and translated version among Fijian girls (Becker et al., 2010), although it has also shown good internal reliability in a trial of eating disorder prevention programmes among adolescent girls (Atkinson & Wade, 2015).

Finally, assessing behavioural avoidance related to the body could provide a useful index of broader impact. Avoidance is employed to escape or provide relief from potential (feared) or actual distress (Cash et al., 2005) and comprises both behavioural (e.g., physically avoiding situations or events) and cognitive strategies (e.g., avoiding thoughts and feelings). Avoidance has a paradoxical effect of reinforcing problems through lack of opportunity to disconfirm fears, and is therefore a key transdiagnostic maintenance process and target of

cognitive-behavioural intervention (Harvey et al. 2004; Kennerly et al., 2016). Behavioural avoidance is of particular relevance to assessing broader consequences of body image concerns, as it directly limits one's range of life experiences beneficial to development and later well-being. The Avoidance subscale of the *Body Image Coping Strategies Inventory* (BICSI; Cash et al., 2005) assesses the extent to which people manage body image threats or challenges using avoidance strategies. However, the 8 items include cognitive avoidance (e.g., "I try to ignore the situation and my feelings", "I try to tune out my thoughts and feelings") and is limited in its representation of life domains (i.e., only psychological, social, eating behaviour). The *Body Image Avoidance Questionnaire* (Rosen et al., 1991) does focus specifically on behavioural avoidance relevant to body image concerns (e.g., wearing baggy clothes, restricting food types, avoiding social situations, weighing). However, it is a relatively long measure at 19 items, and arguably also limited in scope (i.e., eating restraint, grooming and weighing, social situations) and most pertinent to weight and shape concerns (Lydecker, 2017).

Taken together, there exists an opportunity to provide a single measure specifically assessing broader consequences of body image concerns across life domains relevant to adolescents' lives, that draws on behavioural avoidance as a known contributor to impaired functioning, and: a) is broadly relevant to body image and appearance, beyond weight and shape concerns; b) does not conflate mental and physical health functioning with broader life activities; c) is validated and appropriate for use with adolescents; and d) is brief, easily accessed and free to use.

1.3. Origins of the Body Image Life Disengagement Questionnaire (BILD-Q)

In the context of their education initiative to improve body image and self-esteem among young people (the *Dove Self-Esteem Project*), Unilever brand *Dove* has conducted global surveys to assess body image among girls and women, including a focus on broader

life impacts. For example, in their most recent global study they assessed the prevalence, consequences and influences on body image concerns among representative online samples of girls in 13 countries (Dove, 2017). They were also interested in understanding the types of life activities that girls might hold back from engaging in due to concerns about their appearance. These life activities spanned multiple domains, including school/work attendance, sports participation, socialising, seeking healthcare, and asserting an opinion. Their findings indicated that up to 8 in 10 girls globally opt out of key life activities due to worrying about their looks (Dove, 2017). These findings signal ramifications for body image concerns on girls' development and potentially their future contributions to society. Nevertheless, this scale had not been validated psychometrically.

The aim of the current research, therefore, was to adapt and validate a brief measure based on the items used in the Dove (2017) report, which could be used to assess the impact of body image concerns on young people's engagement in key life areas. Below, we report on the process of development and psychometric validation across four studies with adolescents. Given the gendered nature of body image issues, we also aimed to assess gender differences with respect to factor structure, reliability and validity, and level of endorsement. Body image concerns are present across gender identities, although research has highlighted that the nature of concerns differs by gender, and that females generally endorse greater body dissatisfaction (see Murnen, 2011). We therefore expected that the psychometric properties would be upheld across gender, but that there would be higher endorsement among girls.

2. Study 1: Development and Exploratory Factor Analysis

The aims of Study 1 were to develop the BILD-Q based on an existing market research tool and explore its dimensionality, reliability, and mean scores in a sample of early adolescent girls and boys. We expected the BILD-Q to show a unidimensional structure, to be internally consistent and reliable over time among both boys and girls, and that girls would

report greater endorsement than boys.

2.1. Method

2.1.1. Participants and Procedure

Adolescent boys and girls ($N = 1707$; 11-13 years; 50.83% girls) from six secondary schools in southwest England completed self-report measures as part of a larger intervention trial (Diedrichs et al., 2015). The majority of students were White (78.1%), although reflected greater diversity than the general UK population (86% White; Office for National Statistics, 2016), born in the UK (88.4%), and spoke English at home (84.4%). The study was approved by the University ethics review board, with informed assent obtained from parents and students. All schools were public funded academies and had an average or below national average proportion of students claiming free school meals (proxy for socioeconomic status). Students completed questionnaires in a single session under supervision from their teachers and trained research staff. This procedure was repeated one week later, with data from the two assessment-only control schools used to provide a subsample for assessing test-retest reliability.

2.1.2. Measures

2.1.2.1. Life Disengagement. The BILD-Q was adapted from an existing market research interview question used by the *Dove Self-Esteem Project* as part of their global surveys of girls and women (Dove, 2016, 2017). This question asked “*Which, if any, of the following have you ever chosen NOT to do because you felt badly about how you looked?*”, with respondents requested to select any that applied from a list of 11 life activities. The question stem and life activities were developed on the basis of qualitative insights from 2-hour long focus group discussions and interviews conducted for Dove by a research agency with women aged 20-55 years ($n = 21$) and girls aged 11-13 years ($n = 6$) within each of the following countries; United Kingdom, Brazil, France, Indonesia, Japan and China (i.e., total

$N = 126$ women and $N = 36$ girls; Firefish, 2015).

To make these life activity items age appropriate for administration to early adolescents specifically and in school settings, we removed two items (“job interview”, “go to work”) and re-worded two others (“go on a date” to “talk/go out with someone from the opposite sex”; “give an opinion” to “give an opinion or stand up for myself”). We then generated a further seven items with the aim of capturing an even broader range of life activities across academic (school attendance and participation), social (spending time with friends and family), and engagement with healthcare services (going to a doctor or school nurse), guided by insights from our previous research conducted with adolescents (e.g., Atkinson & Wade, 2015; Halliwell & Diedrichs, 2014). We also desired to capture intentions as well as behaviours during a finite period, therefore we adapted the stem question to ask: *“In the PAST 2 WEEKS, how much have worries or feeling bad about the way you looked STOPPED you from doing any of the following things? (If you haven’t done any of these things recently, imagine how you THINK you would have felt)”*. The response scale was adapted from yes/no to a 4-point Likert-type scale (1=not at all, 2=a little bit, 3=quite a bit, 4=all the time) in order to increase measurement sensitivity and alignment with existing scales (e.g., the 4-point response scale of the CIA). This process resulted in an initial pool of 16 BILD-Q items (available on request). A mean is calculated using all items, with higher scores reflecting greater life disengagement.

To ensure suitability of these adaptations for adolescents, an international panel of five experts in the area of adolescent body image research reviewed the resulting scale for content and appropriateness of wording, with particular respect to girls and boys aged 11-14. No issues were raised, and all items were retained in their proposed form. In addition, no problems were observed or reported by students during administration of the scale or follow-up focus groups that were conducted as part of a pilot study to test procedures for the

intervention trial noted above (reported in Diedrichs et al., 2015), or during administration and focus groups for the intervention trial or those following (relevant to data for Study 2-4).

2.1.3. Statistical Analyses

All analyses were conducted in Stata version 14. We first examined missing data and normality of all 16 items. Next, initial conditions for exploratory factor analysis (EFA) were assessed, including inter-correlations and sampling adequacy (described below), and items removed where necessary. Finally, we conducted EFA on the remaining items using principle axis factoring with oblique (promax) rotation, which allows for any retained factors to be correlated. Factors were extracted based on criterion for parallel analysis (Horn, 1965), whereby retained factors needed to reflect eigenvalues larger than the eigenvalues averaged across 330 (30 x number of items) randomly generated samples with the same size and number of observations. Recommendations for removal of factors with <3 items or explaining <5% of variance due to likelihood of being uninterpretable were also followed (Guadagnoli & Velicer, 1988; Tabachnick & Fidell, 2013). Loading patterns were explored following extraction. All items loading with a value > 0.4 were considered acceptable, with at least four items loading > 0.6 ideal (Swami & Barron, 2019). Internal reliability (Cronbach's alpha), means and standard deviations were calculated for the whole sample and by gender. Gender difference in means was assessed statistically via independent samples t-test.

2.2. Results

2.2.1. Preliminary Analyses

Analysis of missing data indicated that 89 participants (5.2%) failed to complete any of the scale items, likely indicating those who did not have time to complete the measure within the broader questionnaire, and were therefore removed from further analysis. Of the remaining 1618 participants, 12.5% had at least one missing item. Examining patterns of missing data indicated the highest missing data for item 6 ("talk/go out with someone from

the opposite sex”; 4%), with remaining items ranging from 1.0-2.2%. As missing data was < 5% across individual items, we used the Expectation-Maximisation (EM) algorithm to impute missing values (Tabachnick & Fidell, 2013). All items were checked for normality, which were all within the recommended bounds of skewness < |3| and kurtosis < |10| (Kline, 2015).

2.2.2. Exploratory Factor Analysis (EFA)

We first explored conditions for EFA on the pool of 16 items. The sample size was adequate, exceeding the ideal 20:1 cases-to-item ratio recommended (Hogarty et al., 2005; Swami & Barron, 2019). We also found good sampling adequacy based on a Kaiser-Meyer-Olin (KMO) value close to 1.0 (KMO = 0.961), and factorable intercorrelations indicated by a significant Bartlett’s test of sphericity, $\chi^2(120) = 18433.08, p < .001$. Nevertheless, the correlation matrix indicated potential problems with multicollinearity, with determinant of the R-matrix < .00001 (Field, 2013). Inspection of the matrix revealed two similar items to be highly correlated (“Spending time with family” and “Spending time with friends”; $r = .82$) and therefore we combined these into one item reflecting the mean for further analysis. As multicollinearity was still present, we removed a further 3 items that had intercorrelations > .7 and overlapped conceptually with other retained items (“Participate in group activities at school”, “Be able to concentrate on what I’m doing” and “Eat with other people”). We also removed a further item due to a number of low correlations ($\leq .3$) with other items (“Complete homework”). The remaining 11 items showed acceptable multicollinearity according to the determinant of the correlation matrix (.002), good sampling adequacy (KMO = .929), and significant Bartlett’s test, $\chi^2(55) = 9092.542, p < .001$.

An EFA was conducted on these 11 items. Factor extraction based on parallel analysis initially indicated four factors for retention. However, factors 3 and 4 were removed due to each explaining <3% of variance, and factor 2 was removed due to all items loading <.4. This resulted in a single factor solution, explaining a total variance of 96.3% (eigenvalue = 5.65).

Factor loadings were all above .6, exceeding the minimum level required for retention. Nevertheless, inspection of the anti-image correlation matrix indicated three item-pairs with high correlations (correlated more with each other than with other items, $r > .3$), demonstrating potential statistical redundancy. After reviewing content of these items, we retained both items from two of the item pairs on the basis of conceptual importance, however, we decided to remove one item (“Try out for a team or club”) from the remaining pair due to potential ambiguity (i.e., “try out” may have led students to perceive this as only relating to sports teams/clubs and contributing to the overlap with the previous question “Do a sport or physical activity”). At this point we also removed the item related to talking or going out with the opposite sex, given the relatively high rate of missing data indicating it may be less appropriate for early adolescents and, in hindsight, that the wording of this item was problematic due to its heteronormative phrasing.

Thus, we conducted a final EFA on a 9-item version of the scale. Parallel analysis again indicated four factors, but three were removed for the same reasons noted above, resulting in a single factor solution (eigenvalue = 4.72; variance explained = 99.3%). When analysed separately by gender, the unidimensional nature of the scale was also upheld in both boys (eigenvalue = 5.4; variance explained = 97.7%) and girls (eigenvalue = 4.1; variance explained = 98.0%). Factor loadings are displayed in Table 1, with all $>.6$ and therefore showing acceptability for retention. This exploratory analysis therefore supported a unidimensional solution for the BILD-Q, in a 9-item configuration suitable for early adolescent samples.

2.2.3. Reliability and Descriptive Statistics

Internal consistency estimates and mean scores, combined and by gender, are reported in Table 1. The 9-item scale reflected good internal consistency based on Cronbach’s alpha. Item-total correlations were also in the acceptable range, all greater than .4 (overall = .64-.82;

boys = .68-.87, girls = .60-.78). Girls endorsed significantly higher scores than boys (see Notes for Table 1).

Test-retest reliability over one week was assessed using a subsample of $n = 290$ students allocated to the assessment-only control arm of the larger trial who had data at both timepoints (56.3% female). Reliability was supported, based on acceptable intraclass correlation coefficients (overall = .70; boys = .77; girls = .62), and paired-sample t-tests showing no significant change over time (overall: $t(289) = 1.60, p = .110$; boys: $t(127) = 1.75, p = .083$; girls: $t(161) = 0.66, p = .509$). Thus the 9-item BILD-Q showed internal consistency and stability over time for both boys and girls.

3. Study 2: Confirmatory Factor Analysis

The aim of Study 2 was to confirm the factor structure of the BILD-Q, by cross-validating in a second sample of early adolescents.

3.1. Method

3.1.1. Participants and Procedure

Adolescent boys and girls ($N = 1403$; 11-13 years; 54.74% boys) from a further 6 secondary schools in south England completed the measures as part of a second intervention trial (Diedrichs, Atkinson, Garbett, & Leckie, 2020). All schools were public funded academies or community schools and had an average or below national average proportion of students claiming free school meals. Students were predominantly White (83.2%), born in the UK (93.9%), and spoke English at home (90.9%). This study was approved by the University ethics review board, with full trial procedures described in Diedrichs et al., 2020. Procedure for data collection in schools was the same as for Study 1.

3.1.2. Measures

Life Disengagement. The 9-item configuration of the BILD-Q was administered, as described in Study 1.

3.2. Results

3.2.1. Preliminary Analysis

Complete missing data was observed for 62 participants (4.4%), again likely representing those who failed to reach this measure within the questionnaire during the time allocated, who were therefore removed from further analysis ($N = 1341$ remaining). Missing data across individual items ranged from 0.5-1.5%. As for Study 1, we used the Expectation-Maximisation (EM) algorithm to impute missing values. All items were checked for normality, with one item exceeding the recommended bounds of skewness $< |3|$ and kurtosis $< |10|$ and subsequently transformed using log transformation for further analysis.

3.2.2. Confirmatory Factor Analysis

Structural equation modelling in Stata version 14.0 was used to conduct a CFA to see how well the unidimensional model fit the data. Model fit was determined using several indices, using recommended index values for goodness-of-fit. Absolute fit is indicated by root mean square error of approximation (RMSEA) and standardised root-mean square residual (SRMR), and relative fit indicated by the Comparative Fit Index (CFI). Guidelines suggest that values close to or below .06 for RMSEA, .08 for SRMR, and .95 for CFI indicate a good fit (Hu & Bentler, 1999). RMSEA values up to .08 indicate acceptable fit, and up to .10 indicate marginal fit, and values above this are generally considered to be a poor fit (Kline, 2005). It has also been noted to avoid interpreting these cut-offs rigidly, but that these indices are preferred to use of Chi-square statistics, given that the latter are sensitive to sample size and therefore likely to indicate poor model fit in large samples such as in the current study (Swami & Barron, 2019).

All items were specified to load onto a single latent factor, informed by the unidimensional structure supported in Study 1. Fit indices are displayed in Table 2, and factor loadings in Table 1. This model provided a good fit to the data with respect to SRMR and

CFI, and an acceptable fit according to RMSEA. When conducted separately by gender, it can be seen that model fit was better for girls than for boys. Modification indices (correlated errors between items which indicate shared variance beyond that accounted for by the latent factor) and associated standardised expected parameter change (SEPC) were inspected to determine potential sources of misfit. In this sample, two of the largest modification indices, that were also associated with SEPC values $> .2$ (Sarlis et al., 2009), were identified for potential re-specification of the model. These represented item pairs that shared closely aligned source domains (i.e., “Go to the beach, pool, sauna” with “Go to social event, party, club” as social items; and “Go to school” and “Raise my hand in class” as academic items) and therefore conceivably shared overlap. While items from these pairs could have been removed to reduce potential statistical redundancy and improve model fit to excellent, we decided to retain all items given the conceptual desire to reflect a breadth of activities. Modelling the error variance for these two pairs in a re-specified model resulted in an improved model fit, $\chi^2(2)=84.80, p < .001$ (see Table 2 for fit indices).

3.2.3. Gender Invariance

We also tested measurement invariance of the BILD-Q with regard to gender, to establish whether it operated similarly in boys and girls. This is an important requirement if gender differences wish to be assessed using this tool. Invariance was assessed in terms of a) configurational (dimensional) invariance (i.e., similar factor structure; Model 1), b) metric (weak) invariance (i.e., similar factor loadings; Model 2), and c) scalar (strong) invariance (i.e., similar intercepts; Model 3), as well as (strict) invariance of between-group differences in means, variances and co-variances (Model 4), using Multigroup Confirmatory Factor Analysis (MGCFA). Based on recommendations by Cheung and Rensvold (2002), the CFI was used as a recommended goodness-of-fit index in preference to likelihood ratio tests, as it is independent of model complexity and sample size. They recommend a cut-off for change

in CFI of .01 as being indicative of a substantial decrease in model fit, thus values under this assumes invariance. Analysis was completed in R, due to the availability of the CFI within reporting options for MGCFA. Results of the model comparisons are found in Table 3. As can be seen, despite significant chi-square tests, the change in CFI was below .01 in all cases indicating invariance across gender.

3.2.4. Reliability and Descriptive Statistics

Internal consistency estimates and mean scores, combined and by gender, are reported in Table 1. The 9-item scale again reflected acceptable internal consistency based on Cronbach's alpha. Item-total correlations were also in the acceptable range, with all greater than .4 (overall = .61-.69; boys = .58-.69, girls =.58-.73). Girls again reported significantly higher scores than boys (see Notes for Table 1).

4. Study 3: Replicating Confirmatory Factor Analysis

The aim of Study 3 was to further confirm the factor structure of the BILD-Q, by cross-validating in a third sample of adolescents. In particular, we wanted to assess model fit among boys in a different sample given that a poor fit was indicated among boys in the Study 2 sample.

4.1. Method

4.1.1. Participants and Procedure

Adolescents ($N = 2034$; 13-14 years; 50.3% boys, 48.7% girls, 1.0% non-binary) from 20 secondary schools in south England completed the measures as part of a third large intervention trial (Atkinson, Stuijzand, Parnell, Treneman-Evans, & Diedrichs, 2021). A range of schools took part, including public funded academies, community schools, and private schools. Students were predominantly White (83.5%), born in the UK (91.1%), and spoke English at home (86.2%). This study was approved by the University ethics review board. Procedure for data collection in schools was the same as for previous studies.

4.1.2. Measures

4.1.2.1. Life Disengagement. The 9-item configuration of the BILD-Q was administered as for previous studies.

4.2. Results

4.2.1. Confirmatory Factor Analysis

Although the option to select non-binary identity was offered in Study 3, the low endorsement required us to exclude non-binary participants from CFA analysis. All items were again specified to load onto a single latent factor, as indicated by Study 1. Fit indices are displayed in Table 2, and factor loadings in Table 1. This model provided a good to excellent fit to the data with respect to SRMR, CFI, and RMSEA. Importantly, model fit was acceptable in both boys and girls in this sample. Modification indices indicated a similar pattern as in Study 2, therefore modified models were conducted using the same re-specification. This again resulted in an improved model fit, $\chi^2(2) = 196.34, p < .001$ (see Table 2 for modified fit indices).

4.2.2. Reliability and Descriptive Statistics

Internal consistency estimates and mean scores, combined and by boys, girls and non-binary gender identities, are reported in Table 1. Due to low endorsement, we were again not able to include non-binary participants in statistical analysis. The 9-item scale again reflected acceptable internal consistency based on Cronbach's alpha. Item-total correlations were also in the acceptable range, with all greater than .4 (overall = .55-.66; boys = .52-.67, girls = .51-.64). Consistent with Study 1 and 2, girls reported significantly higher scores than boys (see Notes for Table 1).

5. Study 4: Construct Validity

The aim of Study 4 was to provide further evidence of reliability and additionally examine construct (convergent and incremental) validity in a fourth sample of adolescents.

With respect to establishing convergent validity, we examined the association between the BILD-Q and constructs relevant to body image. We expected that the BILD-Q would be negatively correlated with body image measures framed in a positive direction (body esteem, body satisfaction), and positively correlated with a key negative risk factor for body image (internalisation of appearance ideals). We also predicted a moderate positive correlation with the Clinical Impairment Assessment (CIA) as a measure of psychosocial impairment due to eating disorder features, including feelings about weight and shape as well as eating and exercise. Finally, we also predicted that the BILD-Q would be related to broader outcomes of well-being; specifically, that it would be negatively correlated with overall mental well-being and positive affect, and positively correlated with negative affect.

With respect to examining incremental validity, we aimed to assess whether the BILD-Q offered additional information beyond assessing body image evaluation itself, and also whether it was distinct from the CIA as an existing and established measure of psychosocial impairment with some shared features. Given the aim of capturing the impacts of broader appearance-based concerns and across different life domains, we hypothesised that the BILD-Q would be uniquely associated with mental well-being, negative affect and positive affect.

5.1. Method

5.1.1. Participants and Procedure

Adolescent boys and girls ($N = 288$; 13-14 years; 50.7% boys) from a further 3 secondary schools in southwest England completed all measures as part of baseline data collection for the third intervention trial (Atkinson, Parnell, & Diedrichs, 2021). Students were predominantly White (87.5%), Born in the UK (95.5%), and spoke English at home (89.9%). This study was approved by the University ethics review board. Procedure for data collection was conducted as for Study 1 and 2, with all baseline data from the trial included

for all analysis except test-retest reliability, where baseline and post-intervention data was used from participants allocated to the control arm of the trial.

5.1.2. Measures

5.1.2.1. Life Disengagement. The BILD-Q was used in a 9-item version configuration as described above.

5.1.2.2. Body Image

5.1.2.2.1. Body Esteem. Global body image was assessed using the Appearance and Weight subscales of the Body Esteem Scale for Adolescents and Adults (Mendelson et al., 2001). This comprised 18 items (e.g., “I’m pretty happy about the way I look”) rated from 1 (Never) to 5 (Always). A mean score was calculated, with higher scores indicating higher body esteem. Cronbach’s alpha values indicated acceptable internal consistency (boys = .93; girls = .96).

5.1.2.2.2. Body Satisfaction. The Body Areas Satisfaction Subscale from the Multidimensional Body-Self Relations Questionnaire (Cash, 2000) was used to assess body satisfaction. It consists of 9 items assessing satisfaction with various body areas (e.g., face, lower torso, muscle tone). Participants were asked to indicate the extent of their satisfaction on a 5-point scale (1 = very unhappy to 5 = very happy). A mean score was calculated, with higher scores indicating greater overall satisfaction. Cronbach’s alpha values indicated acceptable internal consistency (boys = .87; girls = .87).

5.1.2.3. Internalisation of Appearance Ideals. Two subscales of the Sociocultural Attitudes Towards Appearance Questionnaire-3 (Thompson et al., 2004) were used to assess extent of personal subscription to societal appearance ideals: General internalisation (e.g., “I would like my body to look like the bodies of people who are on TV”), and Athletic internalisation (e.g., “I wish I looked as athletic as sports stars”). Participants rated each item from 1 (Totally disagree) to 5 (Totally agree), with mean scores indicating higher

internalisation. Cronbach's alpha values indicated acceptable internal consistency overall (boys = .95; girls = .96), and for individual subscales (General: boys = .94; girls = .95; Athletic: boys = .88; girls = .88).

5.1.2.4. Positive and Negative Affect. Frequency of positive and negative mood states were assessed using the Positive and Negative Affect Schedule for Children, Child Shortened Version (Ebesutani et al., 2012). Five positive (e.g., cheerful) and five negative (e.g., sad) affective states are rated on a scale from 1 ("Very slightly/Not at all") to 5 ("Extremely/Very much"). Mean scores were calculated separately for Positive Affect (PA) and Negative Affect (NA) subscales. Cronbach's alpha values indicated acceptable internal consistency for both (PA: boys = .89; girls = .90; NA: boys = .77; girls = .86).

5.1.2.5. Psychosocial Impairment Secondary to Eating Disorder Pathology. The Clinical Impairment Assessment (Bohn et al., 2008) is a 16-item self-report measure of impairment related eating habits, exercising, and feelings about weight and shape. The extent of impact on aspects of personal, social, and cognitive psychosocial functioning is rated on a scale from 1 (not at all) to 4 (a lot) over the prior 28 days. A mean score was calculated, with higher scores indicating greater impairment. Cronbach's alpha values indicated acceptable internal consistency in the current sample (boys = .94; girls = .96).

5.1.2.6. Well-being. The Warwick-Edinburgh Well-Being Scale (Tennant et al., 2007) measured general mental well-being. This comprises 14 items (e.g., "I've been feeling optimistic about the future" and "I've been feeling close to other people"). The frequency of experience for each item over the prior 14 days is rated from 1 (none of the time) to 5 (all of the time). A mean score is calculated with higher scores signifying higher wellbeing. Internal consistency in the current study was acceptable (boys = .86; girls = .88).

5.2. Results

5.2.1. Reliability

Cronbach's alphas for the 9-item BILD-Q again reflected good internal consistency (total = .82; boys = .77; girls = .80). The range of item-total correlations overall was slightly lower in this sample but still in the acceptable range (total = .57 - .71; boys = .42 - .76, girls = .52 - .73). Test-retest reliability over a 6-week period was assessed using a subsample of students allocated to the assessment-only control arm of the trial who had data at both timepoints ($n = 103$; 53.4% male). Reliability over time was supported, based on acceptable intraclass correlation coefficients (overall = .74; boys = .76; girls = .69), and paired-sample t -tests showing no significant change over time (overall: $t(102) = 1.22, p = .224$; boys: $t(51) = 0.68, p = .499$; girls: $t(50) = 1.02, p = .315$). Thus the 9-item BILD-Q showed internal consistency and stability over time for both boys and girls in this slightly older sample.

5.2.2. *Validity*

Construct validity was assessed by examining the correlation coefficient between scores on the BILD-Q with established measures of body image, a key risk factor for body image (internalisation of appearance ideals), and broader well-being. These are displayed in Table 4, along with means and standard deviations, by gender. As expected, life disengagement showed significant correlations with body image (body esteem and body satisfaction), internalisation of appearance ideals, mental well-being, and positive and negative affect, all in the expected directions. Among boys, the BILD-Q was significantly associated with internalisation of athletic appearance ideals only, and not general internalisation.

Incremental validity of the BILD-Q over body image (body esteem, body satisfaction) and psychosocial impairment due to eating, exercising and weight/shape concerns (CIA) was investigated using hierarchical multiple regression, separately for each outcome of mental well-being, negative affect, and positive affect. For each model, gender was entered at Step 1 to remove influence of gender differences, and body esteem and body satisfaction variables

were entered at Step 2 to ensure that we were testing contribution beyond body image evaluation itself. The CIA was then entered at step 3, and the BILD-Q at Step 4, with the change in R^2 indicating unique association with the outcome (Table 5). We first assessed concurrent validity (all variables assessed at baseline), and results showed the BILD-Q to explain small but significant additional unique variance in general mental well-being, negative affect, and positive affect. In contrast, the CIA was a significant predictor of negative affect only. Finally, we then examined predictive incremental validity, by assessing the contribution of baseline BILD-Q to these outcomes measured 6-weeks later. The results showed the BILD-Q to only explain unique additional variance in well-being, and not negative or positive affect. In contrast, the CIA was the only significant predictor of negative affect 6-weeks later.

6. Discussion

Body image concerns are increasingly recognised as a significant public health issue, given their high prevalence and negative association with a broad range of health outcomes (Al Sabbah et al., 2009; Bornioli et al., 2019). Evaluating the impact of such concerns on participation in important life events would contribute to a more comprehensive understanding of their consequences for adolescents, but requires a measure that simultaneously addresses identified limitations of existing measures. We developed the Body Image Life Disengagement Questionnaire (BILD-Q) to address this gap, and evaluated its psychometric properties across four samples of British adolescents. In general, the BILD-Q exhibited acceptable psychometric properties among adolescent boys and girls. Its predicted unidimensional factor structure was supported, and invariance across gender upheld. It also showed acceptable internal consistency and test-retest reliability over 1-week and 6-week periods, and evidence of construct (convergent and incremental) validity with respect to body image, appearance-ideal internalisation, and general mental well-being.

While psychometric properties were supported for both boys and girls, some gender differences were noted. As expected, girls consistently reported greater life disengagement than boys as measured by the BILD-Q. Although the gender gap appears to be closing (Karazsia et al., 2017; Lacroix et al., 2020), this nevertheless underscores the disproportionate way in which body image affects girls' lives. With regard to convergent validity, the BILD-Q was only related to internalisation of athletic ideals and not overall internalisation of appearance ideals among boys. This likely reflects the salience of muscularity and focus on function within societal appearance ideals for males (Murnen, 2011). Finally, while support for a unidimensional factor structure was observed among both boys and girls aged 13-14, model fit was poorer among boys aged 11-13. Although this improved with a re-specified model, this is likely due to limited relevance of certain aspects measured at this age (e.g., clothes shopping) and thus the BILD-Q may be less suitable for early adolescent boys in its current form. More formative research and validation may be necessary to capture different life aspects and ensure relevance for boys from early adolescence (McCabe & Ricciardelli, 2004).

Importantly, justification for developing and disseminating an additional measure was supported. The BILD-Q provided incremental value beyond assessing body image, and explained unique variance in predicting concurrent and later mental well-being over and above the Clinical Impairment Assessment (Bohn et al., 2008). The correlation between these two measures ($r = .6$) indicates overlap, as would be expected, but not redundancy. Additionally, the BILD-Q's broader framing of appearance concerns (i.e., "the way you look") is likely more relevant across a wider range of individuals and therefore more widely applicable. However, the CIA was a stronger predictor of negative affect than the BILD-Q. This supports the CIA in assessing clinical concerns in the context of eating disorders, of

which negative affect is a significant risk and maintenance factor (Jacobi & Fittig, 2010; Penessi & Wade, 2016).

6.1. Limitations

The strengths of this programme of studies includes the comprehensive assessment of psychometric properties, including factor structure, reliability, and validity, in four different adolescent samples. It additionally included investigation of gender variance, and used large samples, that were consistent with, or more diverse than, the UK population in terms of ethnicity (Office for National Statistics, 2011). However, we also acknowledge that these findings are likely only generalisable to early adolescent samples in high income Western countries. In addition, the original market research tool was developed in consultation with women and girls. Although all final items were reviewed by experts in adolescent body image including those with expertise in boys' concerns, and gender invariance was supported, we recognise that further consultation with adolescent boys would be worthwhile to ensure maximum relevance to boys' engagement in life activities. We also note that mean scores on the BILD-Q were low in these early adolescent samples. While we were able to establish psychometric properties, the risk of floor effects should be mentioned.

6.2. Conclusion

Overall, the BILD-Q shows evidence of being a reliable and valid tool to assess the impact of body image on engagement in activities across a range of life domains among adolescents (e.g., education, socialising, sports participation, seeking healthcare, and self-assertion). It may contribute to a more comprehensive understanding regarding the consequences of body image concerns and evaluation of interventions, and where relevant, underpin advocacy for intervention resourcing, implementation, and policy change. It is unidimensional and invariant across gender, significantly related to body image and related constructs, yet offers additional value beyond existing measures. While further validation in

different contexts and diverse samples is required, its brief form (nine items) and broad framing of appearance lends itself to easy and broad administration across research and practice settings.

References

- Al Sabbah, H., Vereecken, C. A., Elgar, F. J., Nansel, T., Aasvee, K., Abdeen, Z., Ojala, K., Ahluwalia, N., & Maes, L. (2009, Feb 6). Body weight dissatisfaction and communication with parents among adolescents in 24 countries: International cross-sectional survey. *BMC Public Health*, *9*, 52. <https://doi.org/10.1186/1471-2458-9-52>
- All Party Parliamentary Group on Body Image. (2012). *Reflections on body image*.
<http://www.ymca.co.uk/bodyimage/report>
- Andrew, R., Tiggemann, M., & Clark, L. (2015). The protective role of body appreciation against media-induced body dissatisfaction. *Body Image*, *15*, 98-104.
<https://doi.org/10.1016/j.bodyim.2015.07.005>
- Atkinson, M.J., Parnell, J., & Diedrichs, P.C. (2021). A pilot randomised controlled trial of a teacher-led mindfulness-based body image intervention for secondary schools.
Manuscript in preparation.
- Atkinson, M.J., Stuijzand, B.G, Parnell, J., Treneman-Evans, G., & Diedrichs, P.D. (2021). Effectiveness of teacher-led body image interventions in secondary schools: A 4-arm cluster randomised controlled trial. *Manuscript in preparation*.
- Atkinson, M. J., Stock, N. M., Alleva, J. M., Jankowski, G. S., Piran, N., Riley, S., Calogero, R., Clarke, A., Rumsey, N., Slater, A., Diedrichs, P. C., & Williamson, H. (2020). Looking to the future: Priorities for translating research to impact in the field of appearance and body image. *Body Image*, *32*, 53-61.
<https://doi.org/https://doi.org/10.1016/j.bodyim.2019.10.006>
- Atkinson, M. J., & Wade, T. D. (2015). Mindfulness-based prevention for eating disorders: A school-based cluster randomized controlled study. *International Journal of Eating Disorders*, *48*(7), 1024-1037. <https://doi.org/10.1002/eat.22416>

Austin, S. B. (2015). Accelerating progress in Eating Disorders prevention: A call for policy translation research and training. *Eating Disorders*, 24(1), 1-14.

<https://doi.org/10.1080/10640266.2015.1034056>

Becker, A. E., Thomas, J. J., Bainivualiku, A., Richards, L., Navara, K., Roberts, A. L., Gilman, S. E., Striegel-Moore, R. H., & Group, H. F. S. (2010). Adaptation and evaluation of the Clinical Impairment Assessment to assess disordered eating related distress in an adolescent female ethnic Fijian population. *International Journal of Eating Disorders*, 43(2), 179-186. <https://doi.org/10.1002/eat.20665>

Bennett, D. S., Ambrosini, P. J., Kudes, D., Metz, C., & Rabinovich, H. (2005). Gender differences in adolescent depression: Do symptoms differ for boys and girls? *Journal of Affective Disorders*, 89(1-3), 35-44. <https://doi.org/10.1016/j.jad.2005.05.020>

Bohn, K., Doll, H. A., Cooper, Z., O'Connor, M., Palmer, R. L., & Fairburn, C. G. (2008). The measurement of impairment due to eating disorder psychopathology. *Behaviour Research and Therapy*, 46(10), 1105-1110. <https://doi.org/10.1016/j.brat.2008.06.012>

Bornioli, A., Lewis-Smith, H., Smith, A., Slater, A., & Bray, I. (2019). Adolescent body dissatisfaction and disordered eating: Predictors of later risky health behaviours. *Social Science & Medicine*, 238, 112458.

<https://doi.org/10.1016/j.socscimed.2019.112458>

British Youth Council. (2017). *A body confident future*. <http://www.byc.org.uk/wp-content/uploads/2017/11/Youth-Select-Committee-A-Body-Confident-Future.pdf>

Bucchianeri, M. M., Arikian, A. J., Hannan, P. J., Eisenberg, M. E., & Neumark-Sztainer, D. (2013). Body dissatisfaction from adolescence to young adulthood: Findings from a 10-year longitudinal study. *Body Image*, 10(1), 1-7.

<https://doi.org/10.1016/j.bodyim.2012.09.001>

Cash, T. F. (2000). MBSRQ users' manual. Retrieved March, 16, 2006.

- Cash, T. F., & Fleming, E. C. (2002). The impact of body image experiences: development of the body image quality of life inventory. *International Journal of Eating Disorders*, 31(4), 455-460. <https://doi.org/10.1002/eat.10033>
- Cash, T. F., & Grasso, K. (2005). The norms and stability of new measures of the multidimensional body image construct. *Body Image*, 2(2), 199-203. <https://doi.org/10.1016/j.bodyim.2005.03.007>
- Cash, T. F., Jakatdar, T. A., & Williams, E. F. (2004). The Body Image Quality of Life Inventory: Further validation with college men and women. *Body Image*, 1(3), 279-287. [https://doi.org/10.1016/s1740-1445\(03\)00023-8](https://doi.org/10.1016/s1740-1445(03)00023-8)
- Cash, T. F., & Smolak, L. (2011). *Body image: A handbook of science, practice, and prevention*. Guilford Press.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. https://doi.org/10.1207/s15328007sem0902_5
- Coyne, S. M., Linder, J. R., Rasmussen, E. E., Nelson, D. A., & Birkbeck, V. (2016). Pretty as a princess: Longitudinal effects of engagement with Disney princesses on gender stereotypes, body esteem, and prosocial behavior in children. *Child Development*, 87(6), 1909-1925. <https://doi.org/10.1111/cdev.12569>
- Diedrichs, P. C., Atkinson, M. J., Garbett, K. M., & Leckie, G. (2020). Evaluating the "Dove Confident Me" Five-Session Body Image Intervention Delivered by Teachers in Schools: A Cluster Randomized Controlled Effectiveness Trial. *Journal of Adolescent Health*. Advance online publication. <https://doi.org/10.1016/j.jadohealth.2020.10.001>
- Diedrichs, P. C., Atkinson, M. J., Steer, R. J., Garbett, K. M., Rumsey, N., & Halliwell, E. (2015). Effectiveness of a brief school-based body image intervention 'Dove Confident Me: Single Session' when delivered by teachers and researchers: Results

- from a cluster randomised controlled trial. *Behaviour Research and Therapy*, 74, 94-104. <https://doi.org/10.1016/j.brat.2015.09.004>
- Doley, J. R., McLean, S. A., Griffiths, S., & Yager, Z. (2020). Designing body image and eating disorder prevention programs for boys and men: Theoretical, practical, and logistical considerations from boys, parents, teachers, and experts. *Psychology of Men & Masculinities*. Advance online publication. <https://doi.org/10.1037/men0000263>
- Dove. (2016). *The Dove Global Beauty and Confidence Report*. Unilever and Edelman Intelligence.
- Dove. (2017). *The 2017 Dove Global Girls Beauty and Confidence Report*. Unilever and Edelman Intelligence.
- Ebesutani, C., Regan, J., Smith, A., Reise, S., Higa-McMillan, C., & Chorpita, B. F. (2012). The 10-Item Positive and Negative Affect Schedule for Children, child and parent shortened versions: Application of item response theory for more efficient assessment. *Journal of Psychopathology and Behavioral Assessment*, 34(2), 191-203. <https://doi.org/10.1007/s10862-011-9273-2>
- Fairburn, C. G. (2008). *Cognitive behavior therapy and eating disorders*. Guilford Press.
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16(4), 363-370. [https://doi.org/10.1002/1098-108X\(199412\)16:4<363::AID-EAT2260160405>3.0.CO;2-%23](https://doi.org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-%23)
- Fairburn, C. G., Cooper, Z., Doll, H. A., O'Connor, M., Bohn, K., Hawker, D. M., Wales, J. A., & Palmer, R. L. (2009). Transdiagnostic cognitive-behavioral therapy for patients with eating disorders: A two-site trial with 60-week follow-up. *American Journal of Psychiatry*, 166 (3), 311-319. <https://doi.org/10.1176/appi.ajp.2008.08040608>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.

Firefish. (2015). *Care, Beauty, Happiness: Women and Girls Qualitative Formative Research*. Internal Report. Dove, Unilever: Unpublished.

Frederick, D. A., Garcia, J. R., Gesselman, A. N., Mark, K. P., Hatfield, E., & Bohrnstedt, G. (2019). The Happy American Body 2.0: Affective body satisfaction in two U.S. national Internet panel surveys. *Body Image, 32*, 70-84.
<https://doi.org/10.1016/j.bodyim.2019.11.003>

Frederick, D. A., Sandhu, G., Morse, P. J., & Swami, V. (2016). Correlates of appearance and weight satisfaction in a U.S. national sample: Personality, attachment style, television viewing, self-esteem, and life satisfaction. *Body Image, 17*, 191-203.
<https://doi.org/10.1016/j.bodyim.2016.04.001>

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103*(2), 265–275.
<http://dx.doi.org/10.1037/0033-2909.103.2.265>

Halliwell, E., & Diedrichs, P. C. (2014). Testing a dissonance body image intervention among young girls. *Health Psychology, 33*(2), 201-204.
<https://doi.org/10.1037/a0032585>

Halliwell, E., Diedrichs, P. C., & Orbach, S. (2014). *Costing the invisible: A review of the evidence examining the links between body image, aspirations, education and workplace confidence*. <http://eprints.uwe.ac.uk/24438>

Harvey, A. G., Watkins, E., & Mansell, W. (2004). *Cognitive behavioural processes across psychological disorders: A transdiagnostic approach to research and treatment*. Oxford University Press.

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size,

- communality, and overdetermination. *Educational and Psychological Measurement*, 65(2), 202-226. <https://doi.org/10.1177/0013164404267287>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185. <https://doi:10.1007/bf02289447>
- Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Jacobi, C., & Fittig, E. (2010). Psychosocial Risk Factors for Eating Disorders. In W. S. Agras (Ed.), *The Oxford Handbook of Eating Disorders*. Oxford University Press.
- Karazsia, B. T., Murnen, S. K., & Tylka, T. L. (2017). Is body dissatisfaction changing across time? A cross-temporal meta-analysis. *Psychological Bulletin*, 143(3), 293-320. <https://doi.org/10.1037/e510372016-001>
- Kennerley, H., Kirk, J., & Westbrook, D. (2016). *An introduction to cognitive behaviour therapy: Skills and applications*. Sage.
- Kline, R. (2005). *Principles and practice of structural equation modelling*. Guildford Press.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.
- Kostanski, M., & Gullone, E. (1998). Adolescent body image dissatisfaction: Relationships with self-esteem, anxiety, and depression controlling for body mass. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 39(2), 255-262. <https://doi.org/10.1017/s0021963097001807>
- Krebs, G., de la Cruz, L. F., Monzani, B., Bowyer, L., Anson, M., Cadman, J., Heyman, I., Turner, C., Veale, D., & Mataix-Cols, D. (2017). Long-term outcomes of cognitive-

- behavioral therapy for adolescent body dysmorphic disorder. *Behavior Therapy*, 48(4), 462-473. <https://doi.org/10.1016/j.beth.2017.01.001>
- Kusina, J. R., & Exline, J. J. (2019). Beyond Body Image: A Systematic Review of Classroom-Based Interventions Targeting Body Image of Adolescents. *Adolescent Research Review*, 4(3), 293-311. <https://doi.org/10.1007/s40894-019-00121-1>
- Lydecker J.A. (2017). Body Image Avoidance Questionnaire (BIAQ). In: Wade T. (eds) *Encyclopedia of Feeding and Eating Disorders*. Springer.
- Mataix-Cols, D., de la Cruz, L. F., Isomura, K., Anson, M., Turner, C., Monzani, B., Cadman, J., Bowyer, L., Heyman, I., & Veale, D. (2015). A pilot randomized controlled trial of cognitive-behavioral therapy for adolescents with body dysmorphic disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(11), 895-904. <https://doi.org/10.1016/j.jaac.2015.08.011>
- McCabe, M. P., & Ricciardelli, L. A. (2004). Body image dissatisfaction among males across the lifespan: A review of past literature. *Journal of Psychosomatic Research*, 56(6), 675-685. [https://doi.org/10.1016/s0022-3999\(03\)00129-6](https://doi.org/10.1016/s0022-3999(03)00129-6)
- Mendelson, B. K., Mendelson, M. J., & White, D. R. (2001). Body-Esteem Scale for Adolescents and Adults. *Journal of Personality Assessment*, 76(1), 90-106. https://doi.org/10.1207/S15327752JPA7601_6
- Murnen, S. K. (2011). Gender and body images. In T. F. Cash & L. Smolak (Eds.), *Body image: A handbook of science, practice, and prevention* (2nd ed., pp. 173-179). Guildford Press.
- Neumark-Sztainer, D., Paxton, S. J., Hannan, P. J., Haines, J., & Story, M. (2006). Does body satisfaction matter? Five-year longitudinal associations between body satisfaction and health behaviors in adolescent females and males. *Journal of Adolescent Health*, 39(2), 244-251. <https://doi.org/10.1016/j.jadohealth.2005.12.001>

- Office for National Statistics. (2016). National Records of Scotland; Northern Ireland Statistics and Research Agency (2016): 2011 Census aggregate data. UK Data Service. <http://dx.doi.org/10.5257/census/aggregate-2011-1>
- Tylka, T. L., & Piran, N. (2019). *Handbook of positive body image and embodiment: Constructs, protective factors, and interventions*. Oxford University Press.
- Puhl, R., & Peterson, J. (2012). Physical appearance and stigma. In T. F. Cash (Ed.), *Encyclopedia of body image and human appearance* (Vol. 2, pp. 588-594). Elsevier Academic Press.
- Rosen, J. C., Srebnik, D., Saltzberg, E., & Wendt, S. (1991). Development of a body image avoidance questionnaire. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3(1), 32-37. <https://doi.org/10.1037/1040-3590.3.1.32>
- Rumsey, N., & Harcourt, D. (Eds.). (2012). *Oxford handbook of the psychology of appearance*. Oxford University Press.
- Saris, W. E., Satorra, A., & Van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling*, 16(4), 561-582. <https://doi.org/10.1080/10705510903203433>
- Stice, E. (2002). Risk and maintenance factors for eating pathology: A meta-analytic review. *Psychological Bulletin*, 128(5), 825-848. <https://doi.org/10.1037//0033-2909.128.5.825>
- Stice, E., Rohde, P., Butryn, M., Menke, K. S., & Marti, C. N. (2015). Randomized controlled pilot trial of a novel dissonance-based group treatment for eating disorders. *Behaviour Research and Therapy*, 65, 67-75. <https://doi.org/10.1016/j.brat.2014.12.012>
- Stice, E., Rohde, P., Shaw, H., & Gau, J. (2011). An effectiveness trial of a selected dissonance-based eating disorder prevention program for female high school students:

Long-term effects. *Journal of Consulting and Clinical Psychology*, 79(4), 500-508.

<https://doi.org/10.1037/a0024351>

Stice, E., Shaw, H., Burton, E., & Wade, E. (2006). Dissonance and healthy weight eating disorder prevention programs: A randomized efficacy trial. *Journal of Consulting and Clinical Psychology*, 74(2), 263-275. <https://doi.org/10.1037/0022-006x.74.2.263>

Swami, V., & Barron, D. (2019). Translation and validation of body image instruments: Challenges, good practice guidelines, and reporting recommendations for test adaptation. *Body Image*, 31, 204-220. <https://doi.org/10.1016/j.bodyim.2018.08.014>

Tabachnick, B. G., & Fidell, L., S. (2013). *Using Multivariate Statistics* (Sixth ed.). Pearson.

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of life Outcomes*, 5(1), 63. <https://doi.org/10.1186/1477-7525-5-63>

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), 293-304.

<https://doi.org/10.1002/eat.10257>

Tylka, T. 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>

Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry*, 33(9), 1111-1115.

<https://doi.org/10.1001/archpsyc.1976.01770090101010>

Yager, Z., Diedrichs, P. C., Ricciardelli, L. A., & Halliwell, E. (2013). What works in secondary schools? A systematic review of classroom-based body image programs.

Body Image, 10(3), 271-281. <https://doi.org/10.1016/j.bodyim.2013.04.001>

Table 1

Factor Loadings, Descriptive Statistics, and Internal Consistency (Cronbach's Alpha) for the 9-item BILD-Q in Studies 1 (Exploratory), 2 and 3 (Confirmatory)

BILD-Q item	Study 1			Study 2			Study 3			Non-binary ^a (n=20)
	Overall (N=1618)	Boys (n=788)	Girls (n=830)	Overall (N=1341)	Boys (n=725)	Girls (n=616)	Overall (N=2034)	Boys (n=1024)	Girls (n=990)	
1. Go to the beach, pool, sauna	.61	.69	.54	.69	.69	.68	.61	.59	.57	-
2. Go to a social event, party, club	.73	.77	.69	.71	.63	.78	.72	.70	.72	-
3. Go shopping for clothes	.71	.74	.71	.65	.62	.68	.67	.74	.62	-
4. Do a physical activity/sport	.75	.82	.67	.71	.71	.70	.65	.63	.64	-
5. Give an opinion or stand up for myself	.73	.79	.68	.74	.74	.73	.66	.63	.66	-
6. Go to the doctor or school nurse	.58	.63	.53	.64	.68	.60	.66	.71	.61	-
7. Go to school	.82	.86	.77	.70	.66	.72	.61	.58	.61	-
8. Raise my hand in class	.72	.79	.66	.69	.72	.65	.67	.65	.66	-
9. Spend time with friends and family	.82	.87	.78	.69	.71	.67	.64	.72	.60	-
<i>M (SD)</i> ^b	1.52 (0.67)	1.47 (0.70)	1.57 (0.63)	1.31 (0.51)	1.24 (0.47)	1.40 (0.55)	1.41 (0.54)	1.29 (0.47)	1.54 (0.57)	1.77 (0.88)
Cronbach's alpha	.90	.93	.88	.89	.87	.89	.87	.87	.85	.89

^aNot included in statistical analyses due to insufficient numbers

^bScores for girls were significantly higher than for boys in Study 1 ($t = -2.95, p = .003$), Study 2 ($t = -5.52, p < .001$), and Study 3 ($t = -10.22, p < .001$)

Table 2

Fit Indices for Confirmatory Factor Analysis (CFA) of the 9-item BILD-Q in Studies 2 (aged 11-13) and 3 (aged 13-14)

	Study 2					Study 3				
	χ^2	df	CFI	RMSEA (90% CI)	SRMR	χ^2	df	CFI	RMSEA (90% CI)	SRMR
Original										
<i>Overall</i>	293.31	27	.948	.086 (.077-.095)	.034	422.55	27	.937	.087 (.080-.094)	.038
<i>Boys</i>	314.88	27	.898	.121 (.109-.133)	.047	194.03	27	.949	.079 (.069-.090)	.035
<i>Girls</i>	150.81	27	.948	.086 (.073-.100)	.035	219.04	27	.928	.087 (.077-.098)	.042
Modified										
<i>Overall</i>	208.51	25	.964	.074 (.065-.083)	.028	226.20	25	.968	.064 (.057-.072)	.029
<i>Boys</i>	229.89	25	.928	.106 (.094-.119)	.040	114.70	25	.972	.060 (.050-.072)	.027
<i>Girls</i>	129.79	25	.956	.082 (.069-.097)	.032	125.92	25	.962	.066 (.055-.078)	.034

Note. Study 2: *N* = 1341 (616 girls, 725 boys), Study 3: *N* = 2034 (990 girls, 1024 boys, 20 other – not included in CFA). BILD-Q = Body Image Life Disengagement Questionnaire, CFI = Comparative Fit Index, RMSEA = root mean square error of approximation, CI = confidence interval, SRMR = standardised root mean square residual.

Table 3*Study 2 Model Comparisons for Tests of Gender Invariance*

Model	$\chi^2 (\Delta \chi^2)$	<i>df</i> (Δdf)	<i>p</i> (Δp)	CFI (ΔCFI)
M1 Configural	481.83	54		.918
M2 Weak invariance (loadings)	506.29 (24.46)	62 (8)	.002	.915 (.005)
M3 Strong invariance (intercepts)	534.92 (28.63)	70 (8)	<.001	.911 (.004)
M4 Strict invariance (means)	563.81 (28.90)	71 (1)	<.001	.906 (.002)

Note. According to Cheung and Rensvold (2002) $\Delta CFI < 0.01$ implies that the invariance assumption still holds.

Table 4

Study 4 Correlational Matrix Between BILD-Q and Related Constructs, by Gender

	1	2	3	4	5	6	7	8	9	10
1. BILD-Q	-	-.62**	-.45**	.42**	.39**	.35**	-.49**	.55**	-.50**	.61**
2. BES	-.42**	-	.78**	-.64**	-.59**	-.60**	.67**	-.58**	.66**	-.74**
3. BASS	-.35**	.63**	-	-.43**	-.42**	-.39**	.68**	-.49**	.57**	-.66**
4. SATAQ_I	.15	-.41**	-.15	-	.87**	.97**	-.45**	.48**	-.33**	.57**
5. SATAQ_A	.17*	-.34**	-.11	.93**	-	.72**	-.42**	.44**	-.30**	.59**
6. SATAQ_G	.13	-.41**	-.17*	.97**	.80**	-	-.43**	.45**	-.33**	.51**
7. WEMWBS	-.43**	.52**	.53**	-.07	-.01	-.11	-	-.66**	.72**	-.59**
8. PANAS_NA	.38**	-.35**	-.34**	.17*	.12	.19*	-.47**	-	-.65**	.71**
9. PANAS_PA	-.28**	.42**	.35**	.02	.11	-.04	.62**	-.27**	-	-.53**
10. CIA	.62**	-.58**	-.33**	.29**	.29**	.26**	-.34**	.44**	-.23**	-
Range	1-4	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-4
<i>M (SD)</i> girls	1.47 (.50)	3.19 (.90)	3.29 (.73)	2.55 (.96)	2.63 (.99)	2.50 (1.05)	3.39 (.61)	2.36 (.99)	3.50 (.96)	1.63 (.72)
<i>M (SD)</i> boys	1.20 (.31)	3.61 (.72)	3.62 (.72)	2.04 (.90)	2.33 (1.06)	1.88 (.89)	3.60 (.54)	1.73 (.65)	3.72 (.89)	1.26 (.43)

* $p < .05$, ** $p < .01$

Note. Correlations above the diagonal are for girls, correlations below the diagonal are for boys. BILD-Q = Body Image Life Disengagement Questionnaire; BES = Body Esteem Scale; BASS = Body Areas Satisfaction Scale; SATAQ = Sociocultural Attitudes Towards Appearance Scale, I=Internalisation (Athletic and General subscales combined), A=Internalisation-Athletic, G=Internalisation-General; WEMWBS = Warwick-Edinburgh Mental Well-Being Scale; PANAS = Positive and Negative Affect Scale; CIA = Clinical Impairment Assessment

Table 5

Study 4 Results of Heirarchical Multiple Regression Showing Concurrent and Predictive Incremental Contribution of BILD-Q to Well-being Outcomes

	WEMWBS			PANAS NA			PANAS PA		
	ΔR^2	ΔF	β (t)	ΔR^2	ΔF	β (t)	ΔR^2	ΔF	β (t)
<i>Concurrent^a</i>									
Step 1:									
Gender	.048	12.88***	.01 (.21)	.128	37.90***	.14 (2.88)**	.026	7.01***	.04 (0.84)
Step 2:									
BES	.409	95.83***	.22 (.274)**	.230	45.84***	-.03 (-.37)	.308	59.68***	.30 (3.50)**
BASS			.36 (5.40)***			-.09 (-1.41)			.18 (2.41)*
Step 3:									
CIA	.007	3.40	-.03 (-.47)	.134	66.96***	.46 (6.58)***	.006	2.39	-.01 (-0.13)
Step 4:									
BILD-Q	.021	10.13**	-.20 (-3.18)**	.012	5.95*	.15 (2.44)*	.027	10.73**	-.23 (-3.28)**
<i>Predictive^b</i>									
Step 1:									
Gender	.050	5.29*	.02 (.27)	.121	11.86**	.22 (2.24)*	.046	4.17*	-.08 (-0.74)
Step 2:									
BES	.382	33.22***	.30 (2.12)*	.143	8.17**	-.06 (-.36)	.203	11.37***	.24 (1.31)
BASS			.25 (2.19)*			-.06 (-.48)			.15 (1.03)
Step 3:									
CIA	.001	0.22	.10 (.80)	.057	6.97*	.34 (2.20)*	.004	0.40	-.03 (-0.18)
Step 4:									
BILD-Q	.045	8.39**	-.34 (-2.90)**	.000	0.06	.04 (.24)	.006	0.62	-.13 (-0.78)

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

Note. BILD-Q = Body Image Life Disengagement Questionnaire; CIA = Clinical Impairment Assessment; WEMWBS = Warwick-Edinburgh Mental Well-Being Scale; BES = Body Esteem Scale; PANAS = Positive and Negative Affect Scale

^aWell-being and mood outcomes assessed at the same time, $N = 288$ (146 boys, 141 girls, 1 non-binary)

^bWell-being and mood outcomes assessed after 6 weeks, $n = 109$ (56 boys, 53 girls)

Appendix

Body Image Life Disengagement Questionnaire (BILD-Q)

Instructions:

In the PAST 2 WEEKS, how much have worries or feeling bad about the way you looked STOPPED you from doing any of the following things? (If you haven't done any of these things in the past 2 weeks, imagine how you THINK you would have felt.)

	Hasn't stopped me at all	Stopped me a little bit	Stopped me quite a bit	Stopped me all the time
1. Go to the beach or pool	1	2	3	4
2. Go to a social event, party or club	1	2	3	4
3. Go shopping for clothes	1	2	3	4
4. Do a physical activity/sport	1	2	3	4
5. Give an opinion or stand up for myself	1	2	3	4
6. Go to the doctor or school nurse	1	2	3	4
7. Go to school	1	2	3	4
8. Raise my hand in class	1	2	3	4
9. Spend time with friends and family	1	2	3	4

Scoring:

Calculate a mean of all items if at least 7 of 9 items have been rated: $(1+2+3+4+5+6+7+8+9)/9$

Higher scores indicate greater life disengagement