Original Research Article

Introducing Diagnostic Classification Modeling as an Unsupervised Method for Screening Probable Eating Disorders

Assessment
2025, Vol. 32(3) 405–416

DOI: 10.1177/10731911241247483
journals.sagepub.com/home/asm
Sage

Jihong Zhang[1], Shuqi Cui[2], Yinuo Xu[2], Tianxiang Cui[3], Wesley R. Barnhart[4], Feng Ji[5], Jason M. Nagata[6], and Jinbo He[2]

Abstract

Screening for eating disorders (EDs) is an essential part of the prevention and intervention of EDs. Traditional screening methods mostly rely on predefined cutoff scores which have limitations of generalizability and may produce biased results when the cutoff scores are used in populations where the instruments or cutoff scores have not been validated. Compared to the traditional cutoff score approach, the diagnostic classification modeling (DCM) approach can provide psychometric and classification information simultaneously and has been used for diagnosing mental disorders. In the present study, we introduce DCM as an innovative and alternative approach to screening individuals at risk of EDs. To illustrate the practical utility of DCM, we provide two examples: one involving the application of DCM to examine probable ED status from the 12-item Short form of the Eating Disorder Examination-Questionnaire (EDE-QS) to screen probable thinness-oriented EDs and the Muscularity-Oriented Eating Test (MOET) to screen probable muscularity-oriented EDs.

Keywords

diagnostic classification modeling, eating disorder, screening, prevalence

Eating disorders (EDs) are severe mental health conditions characterized by disturbances in eating behaviors and body image (American Psychiatric Association, 2013). According to an epidemiological systematic review, the prevalence of EDs significantly increased globally in recent decades (Galmiche et al., 2019). EDs have several major diagnoses, such as anorexia nervosa (AN), bulimia nervosa (BN), binge-eating disorder (BED), and other specified feeding and eating disorder (OSFED). Ample evidence suggests that EDs are related to varying adverse health consequences (e.g., cardiovascular diseases, osteoporosis, endocrine and metabolic disorders, impaired cognition, lower quality of life, increased risk of suicidality; Cardi et al., 2018; Himmerich et al., 2021; Smith et al., 2018; TODAY Study Group, 2022). Furthermore, individuals with EDs also have significantly elevated mortality rates (Arcelus et al., 2011). Given the high and increasing prevalence of EDs (Galmiche et al., 2019), but extremely low rates of treatment in individuals with EDs (Kazdin et al., 2017), continued efforts are needed to prevent EDs. The crucial first step of ED prevention is to identify those whose symptoms are of clinical concern or to screen probable ED cases with relatively short test lengths (Fitzsimmons-Craft et al., 2019).

Instruments for Screening Probable EDs

Probable EDs are mostly screened by using self-report instruments because of their cost-effectiveness and ease of implementation. For example, a recent review (Fitzsimmons-Craft et al., 2019) on screening probable EDs in the college-aged population summarized the four most widely used tools in this population: the Sick, Control, One stone, Fat, Food (SCOFF; Morgan et al., 1999), the Eating Disorder Examination-Questionnaire (EDE-Q; Fairburn & Beglin, 1994), the 26-item version of the Eating Attitudes Test (EAT-26; Garner et al., 1982), and the Stanford-Washington University Eating Disorder (SWED; Graham et al.,

[1]University of Arkansas, Arkansas, USA
[2]The Chinese University of Hong Kong, Shenzhen, Guangdong, China
[3]University of Macau, Macau, China
[4]Bowling Green State University, Ohio, USA
[5]University of Toronto, Ontario, Canada
[6]University of California, San Francisco, California, USA

Corresponding Author:
Jinbo He, Division of Applied Psychology, School of Humanities and Social Science, The Chinese University of Hong Kong, Shenzhen 518172, Guangdong, China.
Email: anlfhe@gmail.com; hejinbo@cuhk.edu.cn

2019). Furthermore, there are also other tools frequently used for such purposes, such as the Eating Disorder Diagnostic Scale (EDDS; Stice et al., 2000), the Screen for Disordered Eating (SDE; Maguen et al., 2018), the Eating Disorder Examination-Questionnaire Short Form (EDE-QS; Prnjak et al., 2020), and the more recently developed Brief Assessment of Stress and Eating (BASE; Forbush et al., 2022).

To distinguish probable ED cases from nonprobable ED cases, screening tools of EDs commonly rely on predefined cutoff scores, such as 2 for the SCOFF (Morgan et al., 1999), 20 for the EAT-26 (Garner et al., 1982), 2.3 for the EDE-Q (Mond et al., 2004), and 15 for the EDE-QS (Prnjak et al., 2020). Despite their widespread use, these cutoff score determination procedures present multiple limitations. First, methods used for cutoff determination like the receiver operating characteristic (ROC) analysis are considered a follow-up analysis of the scale development, assuming that the observed scores are 100% accurate reflections of the traits (e.g., Bray et al., 2015; Mond et al., 2004) which ignores the measurement errors of the observed scores. Consequently, using cutoff points as criteria for the severity of ED symptomatology without considering measurement errors may introduce bias in screening. Second, cutoff scores often developed in convenience samples may not be universally applicable. Generalizing these scores to other samples where the screening tools have not been validated can be inappropriate. For instance, the SCOFF (Morgan et al., 1999), a widely used tool for screening EDs (Kutz et al., 2020), has a recommended cutoff score of 2, developed based on UK women's clinical and nonclinical samples. While this score has shown high sensitivity in some studies, it has also presented low sensitivity and unlikely high positive rates in different populations, such as a low sensitivity of 53.7% identified in a general UK sample (Solmi et al., 2015), and a high positive rate of 48.8% in Vietnamese University women (Ko et al., 2015). Applying the cutoff score approach in different populations without validation can produce biased screening results. Adapting or validating cutoff scores for new populations involves substantial work, influenced by demographic factors and usually independent of the psychometric evaluation of the tool. Moreover, another potential biased source may lie in the generalizability of screening tools. The cutoff approach assumes the screening tools have adequate psychometric properties in the tested samples, an assumption that may be violated when the tools lack adequate psychometric properties in new samples. For instance, the cutoff score of 2 for the SCOFF was validated in samples of French women, but no psychometric properties were provided (Garcia et al., 2011), potentially leading to problems if certain items have poor item discrimination.

The Proposed Method

Given the aforementioned limitations of using cutoff scores for screening probable EDs (e.g., relying on self-reported ED status and ROC curve analysis to develop optimal cutoff scores; Morgan et al., 1999; Prnjak et al., 2020), the present study introduced an alternative approach, diagnostic classification modeling (DCM), as an unsupervised approach to distinguish probable ED cases from nonprobable cases. The "unsupervised" approach in the context of this study means screening probable EDs without predetermined cutoff thresholds. In DCM, ED status is considered a latent construct. Instead of using a predefined cutoff score to identify status, DCM uses the probability of being an ED case for each individual, considering their item responses. Unlike traditional cutoff scores, which lack information about the uncertainty surrounding them, DCM employs a maximum a posteriori probability (MAP) approach to identify the individual's estimated status. This MAP estimate represents the status that the individual is most likely to have based on available data. Thus, DCM's screening process does not rely on preset cutoff scores but does necessitate researchers' judgment, particularly in choosing the appropriate type of DCM for these data. In psychometrics, DCM has increasingly been used for diagnosing mental disorders of respondents (De La Torre et al., 2018). For example, DCM has been used for characterizing participants based on measures of alcohol-related problems and other psychological symptoms (Tan et al., 2022), identifying subtypes of mathematical learning disability (Ouyang et al., 2023), investigating the interactions of mental disorders (De La Torre et al., 2018), and diagnosing individuals with pathological gambling (Templin & Henson, 2006). However, the use of DCM for screening probable EDs has not been investigated.

There are multiple advantages of using DCM as an alternative screening method over the traditional cutoff score determination procedures for screening probable EDs. First, the ED status of each individual estimated by DCM (also known as *person parameters*) is conditional to the psychometric properties of the scale (also known as *item parameters*) rather than a separate follow-up analysis of scale development. From the perspective of psychometric modeling, the measurement model in DCM estimates the item qualities (e.g., item discrimination), while the structural model estimates the unobserved characteristics of individuals. In the estimation procedures of DCM, both the measurement errors of observed scores and the uncertainty of item qualities

are considered such that DCM-based screening may be less biased or have higher screening accuracy. Second, DCM identifies probable ED cases assuming that the ED status is captured by the commonality of observable responses of ED cases. Thus, DCM-based screening is unique to the samples to be screened and can also be flexibly adapted to capture the ED characteristics of other new samples. On the contrary, in cutoff score–based screening, the cutoff points developed in prior studies are directly used in new samples which reflects the problem of "one size fits all." Third and finally, DCM can provide psychometric information about the screening tools (e.g., item information, item difficulties, and item discrimination) while cutoff-score–based screening usually provides no such psychometric information in new samples. This is especially useful in samples that are distinct from the population in which cutoff scores were developed and thus DCM does not assume the psychometric properties are identical across these distinct populations. A detailed introduction to DCM (Section S1) is described in the Supplemental Materials.

The Present Study

The main purpose of this work was to provide an introduction to DCM to inform the screening of EDs with two empirical examples. We employed the dichotomized response DCM (items with responses coded with 0 = *No* and 1 = *Yes*) for psychometric evaluation and screening tool modification and then used polytomous response DCM (items with nonbinary ordinal responses; e.g., 1 = *never* and 5 = *always*) for screening probable ED cases. The rationale for using dichotomized response DCM for psychometric evaluation is that the model fit evaluation of dichotomized response DCM has been well established in simulation studies (e.g., Chen et al., 2013; Hu et al., 2016), while the model fit measures of polytomous response DCM are still in development and have not been well examined (Ma, 2020). Specifically, in the first example, the EDE-QS, which measures thinness-oriented ED psychopathology (e.g., dietary restriction, weight concerns, and compensatory behaviors to lose weight), was used. We first performed item selection and validity analysis with model fit indices, M_2, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Squared Residual (SRMSR), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC), using one dichotomized response DCM (log-linear cognitive diagnosis model [LCDM]; Henson et al., 2009) by transforming item responses from the original 4-point Likert-type scale into a 0–1 scale (responses equaling to or larger than 3 as 1 and less than 3 as 0). Next, we obtained the revised scale with the most informative items which we then used to screen the samples' ED status using a polytomous response DCM (general diagnostic model [GDM]; von Davier, 2008). This is because, compared to the dichotomized response DCM, polytomous response DCM can make use of more information on item responses for latent ED status estimation. For example, one item with a 4-point scale provides more diagnostic information about symptom severity than a dichotomized item with only *Yes* or *No* options. Previous literature also shows that the polytomous response DCM provides a more accurate estimation of latent attribute status (e.g., Chen & De La Torre, 2013; Ma, 2020; Templin et al., 2008). Finally, the screening results were validated by investigating the agreement reliability indicated by Cohen's Kappa coefficient (i.e., the extent to which the responses of 2 or more independent raters are concordant) between the proposed DCM method with the previous cutoff method.

Given ED psychopathology can also be driven by the pursuit of muscularity (e.g., rigid dietary rules and exercise routines for enhancing muscularity), which is distinct from thinness-oriented ED psychopathology (Murray et al., 2016), in the second example, we also demonstrated the use of DCM for screening probable muscularity-oriented EDs via the Muscularity-Oriented Eating Test (MOET; Murray et al., 2019). We concluded this article by outlining some of the challenges ahead and possible future research directions for the application of DCMs for ED assessment. In summary, this study poses two primary research questions for Example 1: (a) Does the DCM identify the most informative items, thereby reducing test length without compromising the reliability of screening results? (b) Can DCM be utilized for screening probable EDs with reliability and accuracy comparable to the predefined cut-off score method? In Example 2, the research questions were extended to the MOET, which assesses muscularity-oriented ED psychopathology: (a) Is DCM capable of detecting variations in muscularity-oriented EDs rates by sex? (b) Does DCM exhibit satisfactory discriminative ability without a prespecified cutoff score, as measured by the Area Under the Curve (AUC) statistic?

Example 1: Screening for the EDE-QS

Participants and Procedures

These data were from a project approved by the institutional review board of the Chinese University of Hong Kong, Shenzhen. The participants were from a university in Hunan province. The inclusion criteria were Chinese nationality, university students, and at least 18 years old. The data collection process involved a paper-and-pencil format. Two attention checks (e.g., "*please*

choose strongly agree for this item") were employed to ensure response quality, and passing both checks was necessary for a valid completion. Overall, 1,059 students met our inclusion criteria and provided written informed consent, but 247 were removed because of failure of attention checks ($N = 812$). The detailed sampling procedure can be found in previous publications using these data (He et al., 2022; He, Murray, et al., 2021), which had distinct research questions from the present study. In addition, in the present study, 32 participants with missing demographic information and/or responses on the EDE-QS were removed in post hoc data cleaning, leading to a final sample size of 780. The 780 participants were 18 to 24 years old ($M = 18.88$, $SD = 0.98$). Their body mass index (BMI) ranged from 14.80 to 36.70 kg/m^2 ($M = 21.10$, $SD = 3.41$). Of the 780 participants, 278 (35.64%) were males, and 502 (64.36%) were females. In addition, 709 (90.9%) were Han, while 71 (9.1%) were ethnic minorities.

Measures

12-Item EDE-QS. The Chinese version of the 12-item EDE-QS (Gideon et al., 2016; He, Sun, & Fan, 2021) was used for illustrative purposes. The scale consists of 12 items with response options ranging from 0 (0 days/not at all) to 3 (6–7 days/markedly). The sum of the 12 items results in a total score, and a higher total score indicates a higher level of thinness-oriented disordered eating. The EDE-QS indicated good internal consistency reliability ($\alpha = .89$), test–retest reliability, and convergent validity in Chinese young adults (He, Sun, & Fan, 2021). In the present study, the EDE-QS had a Cronbach's α of 0.88 and a McDonald's ω of 0.88.

Clinical Impairment Assessment 3.0. The Clinical Impairment Assessment 3.0 (CIA 3.0) is a 16-item self-report measure of the severity of psychosocial impairment due to ED pathology (Bohn & Fairburn, 2008). Each item is rated on a 4-point Likert-type scale, with the response options ranging from 0 (not at all) to 3 (a lot). The resulting global CIA impairment score ranges from 0 to 48, with a higher score indicative of higher psychosocial impairment. The Chinese translation of the CIA showed adequate internal consistency reliability, test–retest reliability, and convergent validity for Chinese adolescents and adults (He et al., 2022). In the present study, the CIA had a Cronbach's α of 0.93 and a McDonald's ω of 0.94.

Analysis Plan

As mentioned earlier, Example 1 has two purposes: (a) to filter out items containing the most information from the original scales and (b) to screen the ED status of participants. Figure 1 shows the proposed procedure for the item-selection purpose. First, items were dichotomized to 0/1, with item responses equaling 3 coded as 1 and those equaling to or lower than 2 coded as 0 (see Supplemental Materials S3 for more details on item dichotomization). Since models with dichotomized responses were used for item selection, EDs were statistically identified assuming that probable ED cases were expected to have a higher probability of selecting "1" than nonprobable cases. After these data were recoded, a DCM (*log-linear cognitive diagnostic model*) with a one-column Q-matrix (a unidimensional model) was employed to estimate model fit and item information. The rationale for using unidimensional DCM is that the EDE-QS was originally developed based on the unidimensional Rasch model (Gideon et al., 2016). We calculated absolute model fit indices to evaluate global model fit and item information to select the most informative items. Specifically, if the initial DCM showed an unacceptable model fit based on criteria, item-level Kullback-Leibler Information (KLI) for each item then was calculated to select items containing the highest information, and items with the lowest item-level KLI were dropped. Last, another DCM was fitted to the revised scale to determine whether the item-selection process was finished (i.e., acceptable model fit) or to be continued (i.e., unacceptable model fit). If the item-selection process was finished according to absolute model fit indices, the final scale was used for further diagnostic classification. Otherwise, the first two steps were repeated until the model achieved an acceptable fit.

For the second purpose of screening, the analysis was separated into two steps. First, we conducted a GDM for polytomous item responses (von Davier, 2008) to classify individuals. Second, the classification results were validated by calculating Cohen's Kappa coefficient between the GDM-estimated ED statuses with the 15-point cutoff score from the EDE-QS (Prnjak et al., 2020). In addition, the convergent validation was evaluated with the correlation between the diagnostic classification and CIA scores.

KLIs were implemented via *R* 4.2.1 with the *GDINA* package version 2.8.8 (Ma & De La Torre, 2022). Model estimation was performed using the *CDM* package version 8.1-12 (George et al., 2016). Reliability and correlation coefficients were calculated with the *psych* package version 2.2.9 (Revelle, 2022).

Results

Descriptive Statistics. Table 1 shows the difficulty and standard deviations of the 12 items with the original

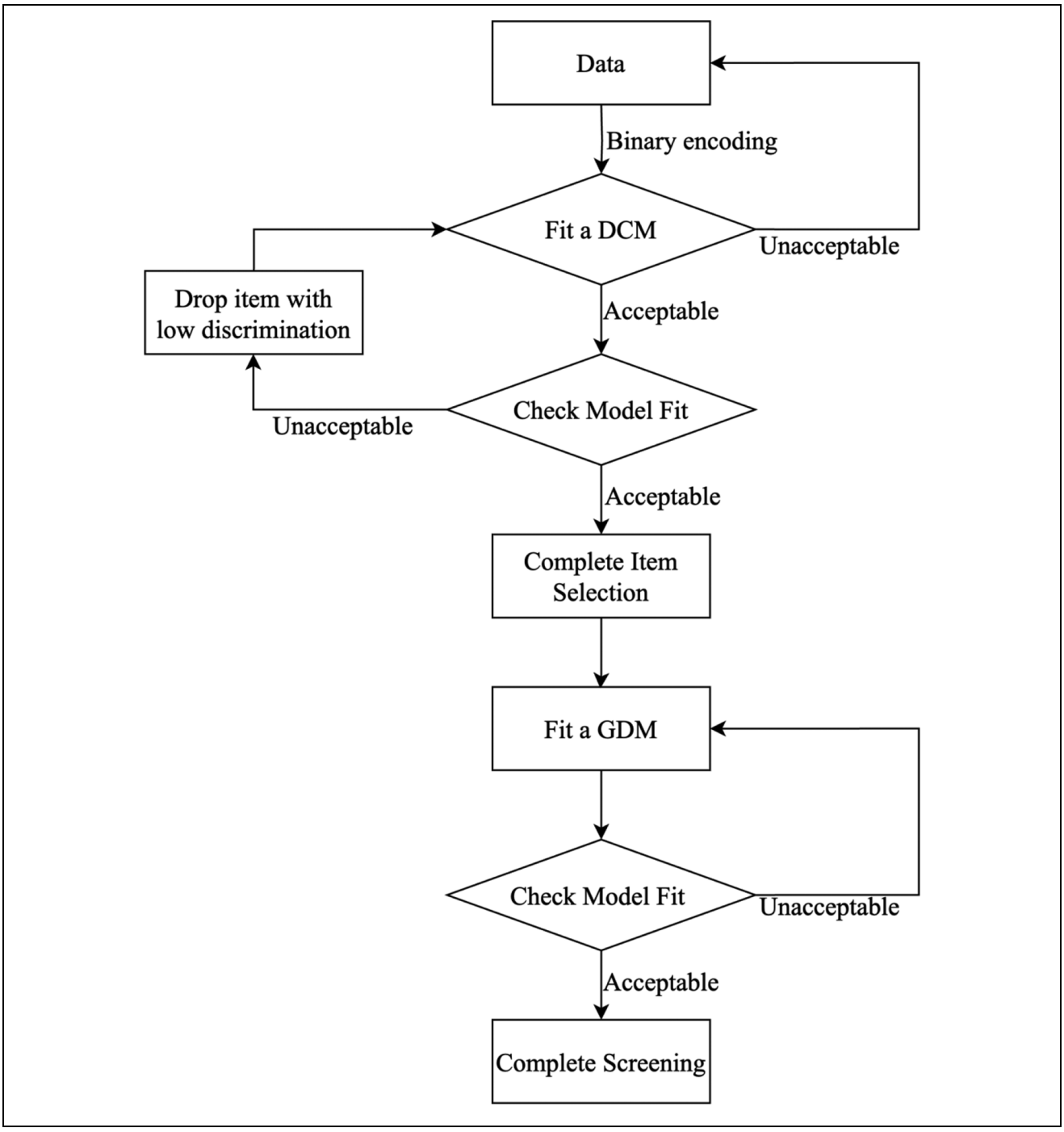

Figure 1. Diagram of Item Selection and Screening Using DCM.
Note. Binary encoding was based on the literature of eating disorder epidemiology; DCM: diagnostic classification model with dichotomized item responses; GDM: general diagnostic model for polytomous item responses.

scale and proportion of endorsement of selecting the highest level ("*6-7 days/markedly*"). To be specific, Item #12 ("*How dissatisfied have you been with your weight or shape?*") had the highest average score (M = 1.076, SD = 0.898), while Item #7 ("*Have you tried to control your weight or shape by making yourself sick [vomit] or taking laxatives?*") had the lowest average score (M = 0.127, SD = 0.437). Similar to the original scale, according to the proportion of endorsement of selecting 3, Item #6 ("*Have you have a strong desire to*

Table 1. Descriptive Statistics of the EDE-QS (N = 780) and the MOET (N = 761).

Scale	Item	*M*	*SD*	*Prop.*
EDE-QS	1. Have you been deliberately trying to limit the amount of food you eat to influence your weight or shape (whether or not you have succeeded)?	0.865	0.950	8.333
EDE-QS	2. Have you gone for long periods of time (e.g., 8 or more waking hours) without eating anything at all in order to influence your weight or shape?	0.317	0.632	1.667
EDE-QS	3. Has thinking about food, eating, or calories made it very difficult to concentrate on things you are interested in (such as working, following a conversation, or reading)?	0.328	0.625	1.282
EDE-QS	4. Has thinking about your weight or shape made it very difficult to concentrate on things you are interested in (such as working, following a conversation or reading)?	0.337	0.598	1.026
EDE-QS	5. Have you had a definite fear that you might gain weight?	0.928	0.994	11.282
EDE-QS	6. Have you had a strong desire to lose weight?	1.049	1.103	16.282
EDE-QS	7. Have you tried to control your weight or shape by making yourself sick (vomit) or taking laxatives?	0.127	0.437	0.897
EDE-QS	8. Have you exercised in a driven or compulsive way as a means of controlling your weight, shape, or body fat or to burn off calories?	0.554	0.847	4.872
EDE-QS	9. Have you had a sense of having lost control over your eating (at the time that you were eating)?	0.365	0.665	1.923
EDE-QS	10. On how many of these days (i.e., days on which you had a sense of having lost control over your eating) did you eat what other people would regard as an unusually large amount of food in one go?	0.401	0.638	1.026
EDE-QS	11. Has your weight or shape influenced how you think about (judge) yourself as a person?	0.712	0.753	1.923
EDE-QS	12. How dissatisfied have you been with your weight or shape?	1.076	0.898	7.821
MOET	1. I have recorded the macro-nutritional values of everything that I ate.	0.870	1.153	2.806
MOET	2. I have used meal-replacement supplements when I felt full.	0.722	0.844	0.383
MOET	3. What I ate has influenced how I think about myself as a person.	0.560	0.808	0.383
MOET	4. There are definite foods I have avoided eating due to the worry about how they might affect my shape or weight.	0.978	1.083	2.551
MOET	5. I have felt less anxious about eating out if I knew the macro-nutritional content of the food at the restaurant.	0.722	0.911	1.020
MOET	6. I have taken my own food out with me to social events in case the food on offer is inconsistent with my diet plan.	0.379	0.721	0.638
MOET	7. I cannot achieve my ideal body unless I exert complete control over everything I eat.	1.036	1.243	5.612
MOET	8. I have precooked several meals in advance to ensure that I do not deviate from my diet plan.	0.306	0.626	0.255
MOET	9. I have continued eating despite feeling full in an attempt to influence my muscularity.	0.472	0.764	0.510
MOET	10. I have felt anxious when I run out of protein-based supplements.	0.244	0.605	0.128
MOET	11. I have been deliberately trying to limit the overall volume of some foods so that my muscles look more defined.	0.374	0.716	0.255
MOET	12. If I broke any of my food rules, I attempted to make up for it with my next meal.	0.644	0.945	1.403
MOET	13. I have felt anxious about others knowing the rules I have around what I eat.	0.316	0.643	0.128
MOET	14. Other people do not seem to understand how important my food choices are to me.	0.513	0.821	0.765
MOET	15. Ensuring proper adherence to my dietary ideals is more important to me than adhering to a work schedule.	0.571	0.843	0.765

Note. EDE-QS = 12-item short form of the Eating Disorder Examination-Questionnaire; MOET = 15-item Muscularity-Oriented Eating Test; *M* = mean, *SD* = standard deviation, *Prob* = proportions of the responses of "6–7 days/markedly" and "always true" for the EDE-Q and the MOET, respectively.

lose weight?") had the highest percentage (16.28%), and Item #7 had the lowest percentage (0.90%), suggesting the consistency of the severity of EDs for the original scale and binary scale.

Item Selection. In total, five DCMs with dichotomized responses (Models 1a–e) were estimated to obtain the optimal DCM (Model 1e) with acceptable model fit (see Table S1 in Supplemental Materials). In addition, Model 1e has the best model fit with the lowest AIC and BIC values. One DCM with polytomous responses (Model 1f) was used to identify probable ED cases from nonprobable cases. Details of item selection were illustrated in the Supplemental Materials (Section S2).

Screening Results and Reliability. For the purpose of screening, we compared three GDMs (see Table S3 in Supplemental Materials). The information criteria revealed that the two-parameter logistic (2PL) general diagnostic model (2PL-GDM) with polytomous responses (Model 1f) yielded the best model fit among alternative models. Next, we fitted Model 1f to the revised eight-item EDE-QS for ED screening. The item parameters (category-level difficulty and slope) and item fit (RMSEA) for Model 1f were presented in Table S2. The overall positive rate of probable EDs was 12.31% (96 out of 780), with 23.96% males (23 out of 96) and 76.04% females (72 out of 96).

Furthermore, the results of comparing the diagnostic classification results under Model 1f to the results of the cutoff point of 15 (Prnjak et al., 2020) suggested that the agreement reliability (Cohen's $\kappa = = 0.706$, $p < .001$) and total agreement rate ($p = 93.46\%$) were high. Specifically, 74 out of 780 individuals had been identified as probable ED cases in both the proposed DCM screening method and the cutoff method, while 655 individuals had been identified as non-ED cases by both methods (see Section S4 in Supplemental Materials).

Finally, regarding the criterion validity of the proposed method, the correlation analysis suggested the screened probable ED status (1 = *ED case*; 0 = *non-ED case*) by DCM (Model 1f) had a positive point-biserial correlation with the total sum scores of the CIA ($r_{X_DY} = 0.33$, 95% $CI = [0.27, 0.40]$). The estimated probability of being screened as a probable ED case by the DCM was significantly correlated with the CIA scores ($r_{X_pY} = 0.36$, 95% $CI = [0.29, 0.42]$). The effect size of the correlation estimated by the method of Prnjak et al. (2020) for the EDE-QS with total CIA scores was close to the proposed DCM method ($r_{X_DY} = 0.39$, 95% $CI = [0.33, 0.45]$).

Example 2: Screening for the MOET

In example 2, we illustrate how DCM can be applied to screening probable muscularity-oriented EDs with the Chinese version of the MOET (He, Murray et al., 2021) of which the cutoff score has not been developed.

Participants and Procedures

Data used in Example 2 were the same as those used in Example 1 ($N = 812$). However, 51 participants with missing demographic information and/or item responses on the MOET were removed, leading to a final sample size of 761, aged 18 to 24 years ($M = 18.88$, $SD = 0.97$). Their BMI ranged from 14.80 to 34.90 kg/m^2 ($M = 21.10$, $SD = 3.39$). Of the whole sample, 90.8% ($N = 691$) were Han, while 9.2% ($N = 70$) were ethnic minorities. 35.7% ($N = 272$) were males and 64.3% ($N = 489$) were females.

Measures

The Muscularity-Oiented Eating Test (MOET) consists of 15 five-point Likert-type items, with response options ranging from 0 (*never true*) to 4 (*always true*) to assess muscularity-oriented disordered eating (Murray et al., 2019). The total scores were obtained by summing all items' scores, with a higher score indicating a higher degree of muscle-oriented disordered eating. The Chinese version of the MOET has been demonstrated to possess good internal consistency reliability (McDonald's $\omega = 0.90$–0.92) and concurrent validity among Chinese adult men and women (He et al., 2023; He, Murray et al., 2021). In the present study, the McDonald's ω and the Cronbach's α for the MOET were 0.89 and 0.89, respectively.

Analysis Plan

To investigate whether DCM can be used for screening probable muscularity-oriented disordered eating without a predefined cutoff, we designed the second study as follows: First, item selection was performed by applying DCM to the 15-item MOET to obtain a shorter version of the MOET. Next, we screened muscularity-oriented disordered eating using GDMs for ordinal responses and then estimated the positive rates of muscularity-oriented disordered eating by sex.

Results

Descriptive Statistics. Table 1 shows the average scores and standard deviations of the 15 items on the MOET and the proportion of endorsement of selecting the

highest level (4: always true). To be specific, Item 7 ("*I cannot achieve my ideal body unless I exert complete control over everything I eat*.") had the highest average score ($M = 1.036$, $SD = 1.243$), while Item #10 ("*I have felt anxious when I run out of protein-based supplements*.") had the lowest average score ($M = 0.244$, $SD = 0.605$). Similarly, according to the proportion of endorsement of selecting the highest option, Item 7 had the highest percentage (5.61%) while Item #10 ("*I have felt anxious when I run out of protein-based supplements*.") and Item #13 ("*I have felt anxious about others knowing the rules I have around what I eat*.") also had the lowest percentages (0.13%), suggesting the consistency of the distribution of muscularity-oriented EDs on the original scale compared to the dichotomized scale.

Item Selection. As shown in Table S1 in the Supplemental Materials, we applied two DCMs (Model 2a–b) to the 15-item MOET for the item selection and one GDM with ordinal responses (Model 2c) to improve the estimation accuracy. Both Model 2a and Model 2b showed the same correlations of screened ED status with CIA scores ($r_{CIA} = .152$), suggesting that the revision of the screening tool did not decrease its criterion validity. The initial DCM, Model 2a, showed an unacceptable model fit for M_2 and $SRMSR$, while the value $RMSEA_2$ indicated a close fit ($M_2(df = 89) = 114.7$, $p < .001$; $RMSEA_2 = 0.032$; $SRMSR = 0.110$). After investigating KLI rankings for items, we removed the item with the lowest KLI (Item 1). Then, Model 2b using the 14-item MOET achieved an acceptable model fit for the M_2 and $RMSEA_2$ statistics, while SRMSR indicated an unacceptable fit ($M_2(df = 76) = 61.6$, $p = 0.195$; $RMSEA_2 = 0.015$; $SRMSR = 0.116$). We still considered Model 2b as acceptable since it has been recommended that $SRMSR$ is more easily influenced by characteristics of data, and comparing the $SRMSR$ against a prespecified cutoff (e.g., .05) may not be appropriate (Ma, 2020). In addition, Model 2b showed lower values of AIC and BIC than Model 2a, suggesting a better model fit. Thus, the revised 14-item MOET was used for further screening.

Screening Results. After identifying the most informative items using Model 2b, we performed multiple GDMs to investigate which type of DCM had the best screening accuracy (see Table S3 in the Supplemental Materials). The 2PL-GDM for polytomous item responses (Model 2c) had the best model fit among alternative models and was eventually fitted to the 14-item MOET data. The item parameters (category-level difficulty and slope) and item fit (RMSEA) for Model 2c were presented in Table S4 (see Supplemental Materials). The overall positive rate of muscularity-oriented EDs is 4.07% (31 out of 761 participants). Regarding sex differences, the male samples had a higher positive rate of muscularity-oriented EDs than females ($PR_{male} = 5.26\%$, $PR_{female} = 3.43\%$). In addition, we calculated the AUC statistic with the total sum score of MOET as the predictor and the diagnostic classification as the outcome. As shown in Figure 1S (see Supplemental Materials), the MOET demonstrated excellent discriminatory power according to the AUC statistic (AUC = .995).

Discussion

Due to the importance of screening EDs, screening tools with predefined cutoffs are commonly used for identifying probable EDs. However, concerns remain about the suitability and generalizability of this traditional cutoff score–based approach. Thus, the present study examined a new screening perspective using a model-based method: DCM, which can estimate the probability of class membership of ED status. This work provides two empirical examples to illustrate how this model-based screening method can be used in real ED screening settings that do not have predefined cutoff scores.

Specifically, we first employed a baseline unidimensional DCM to the EDE-QS (thinness-oriented EDs) and MOET (muscularity-oriented EDs) to obtain item information. Next, according to item information rankings, we removed items with the lowest information and calculated model fit indices. If the model fit was unacceptable, we refitted the DCM after removing items with the lowest information and then performed item selection again until we obtained a shorter-length questionnaire with acceptable estimation accuracy. Then, we fitted polytomous response DCM to the revised questionnaires to obtain more accurate screening results. To evaluate the accuracy of results, we compared the DCM-based screening results for the EDE-QS to previous screening benchmarks (Prnjak et al., 2020) and found that the classification results from the DCM have good agreement with the method of Prnjak et al. (Cohen's $\kappa = 0.62$, $p < .001$), suggesting the two screening methods have relatively consistent screening results. The fair agreement between the two methods suggests that DCM could be an alternative method for screening ED status relative to traditional predefined cutoff methods. Compared to the predefined cutoff score method, the proposed method also has similar correlations with C-CIA scores (the proposed method: $r_{X_D Y} = 0.33$, 95% $CI = [0.27, 0.40]$; predefined cutoff method: $r_{X_D Y} = 0.39$, 95% $CI = [0.33, 0.45]$), indicating they have comparable criterion validity. For the MOET, we conducted the same procedure of item selection and screening and examined the positive rates of

muscularity-oriented EDs by sex. We found that the positive rates were consistent with those of muscle dysmorphia (e.g., around 1%−7%, with higher rates in males than in females; Compte et al., 2015; Lechner et al., 2019; Mitchison et al., 2022), which is closely related to (or a type of) muscularity-oriented EDs (Griffiths & Murray, 2017; Murray et al., 2010, 2017).

Implications and Recommendations

The DCM-based screening method can be an alternative way to screen for probable EDs and may be especially helpful in scenarios when a validated cutoff has not been developed for the measures (e.g., newly developed measures) and/or for certain populations (e.g., new samples). In such situations, researchers and clinicians can rely on DCM to dynamically differentiate participants/clients/patients with significantly higher risk of EDs from those with significantly lower risk of EDs. Furthermore, as demonstrated in the present work, DCM can also help shorten scale length and ease participant/client/patient burden without losing much information for screening. Specifically, DCM can select items with high discriminative power to classify test takers, such that the DCM-based screening method can provide a revised scale only containing items providing the most relevant information (high item discrimination index) and fitted to the characteristics of the target population (acceptable model fit).

To perform the DCM-based screening, this study suggests two measures of DCM-based screening: model fit evaluation and information measures. Acceptable model fit ensures the validity of the screening of DCM, while diagnostic information measures play a critical role in item selection. In addition, this study suggests using DCM with binary responses (e.g., LCDM or other types of DCM) for item selection, whereas DCM with ordinal responses (e.g., GDM; von Davier, 2008) should be used for screening. As mentioned in a previous section, DCMs with binary responses are more efficient because of the smaller number of parameters estimated and their model fit measures are well established (Ma, 2020), while DCM with ordinal responses has higher estimation accuracy in terms of attribute profiles estimation (Liu & Jiang, 2020; Templin et al., 2008; Tu et al., 2018).

Overall, DCM provides an alternative method of screening ED cases without predefined cutoff scores. For scales without predefined cutoff scores, we recommend that researchers perform DCM and set cutoff thresholds at the same time to cross-validate their results.

Limitations and Future Research Directions

There are some limitations and future research directions. First, in the present study, the samples used were nonclinical Chinese college students. The generalization of screening results to other populations in varied contexts needs more investigation. Second, in this study, due to the lack of self-reported ED information (i.e., by asking participants to respond with Yes or No to the question "*Do you currently suffer from an ED (anorexia nervosa, bulimia nervosa, binge-eating disorder, eating disorder not otherwise specified)?*"), future research may focus on investigating how to measure the predictive validity of the screening results of DCM (i.e., examining the association between results of DCM with information on ED status collected from clinical interviews with participants). Third, the current study does not investigate the estimation accuracy of screening information of DCM under different conditions (i.e., other types of DCMs and other ED screening scales such as SCOFF). This is important because, as a model-based screening method, the uncertainty of parameters estimated by DCM may affect the classification consistency and accuracy of ED status. For example, previous studies suggest that varied types of DCMs, sample size, confounders, number of attributes measured, and item quality may affect the parameter estimation (e.g., Bradshaw & Madison, 2016; Cui et al., 2012; Ma, 2020). Thus, further research may be needed to identify what and how factors affect the screening accuracy of DCM for ED questionnaires. Finally, the screening results of DCM cannot be overinterpreted as a diagnosis of ED. The diagnostic information of DCM can be considered self-reported instrument-screened EDs, which should be considered as a quick, statistical screening for EDs. It is suggested that practitioners combine DCM-based screening with other diagnosis tools (i.e., interviews) in research and clinical settings to improve the interpretability of results.

Conclusion

In conclusion, DCM may be useful in the screening of EDs. Future studies are needed to further test the performance of using DCM in different settings for screening EDs, such as different instruments (e.g., SCOFF), different EDs (e.g., avoidant/restrictive food intake disorder), and different populations (e.g., children and adolescents; sexual and gender minorities).

Author Contributions

Jihong Zhang: Conceptualization, Formal analysis, Writing—original draft, Writing—review & editing. Shuqi Cui: Writing—original draft, Writing—review & editing. Yinuo Xu: Writing—original draft, Writing—review & editing. Tianxiang Cui: Writing—original draft, Writing—review & editing. Wesley R. Barnhart: Writing—review & editing. Feng Ji: Writing—review & editing. Jason M. Nagata: Writing—

review & editing. Jinbo He: Conceptualization, Investigation, Supervision, Funding acquisition, Project administration, Writing—original draft, Writing—review & editing.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Jinbo He was supported by the National Natural Science Foundation of China (grant number: 7220 4208) and the Shenzhen College Stability Support Plan (深圳市高等院校稳定支持计划) from Shenzhen Science and Technology Innovation Commission (grant number: 2023SC0028).

Ethical Statement

These data were from a project approved by the institutional review board (Applied Psychology) of The Chinese University of Hong Kong, Shenzhen (no. EF20201024001).

ORCID iDs

Shuqi Cui https://orcid.org/0000-0002-4929-5167
Tianxiang Cui https://orcid.org/0000-0003-2826-1109
Wesley R. Barnhart https://orcid.org/0000-0002-9809-5225
Jinbo He https://orcid.org/0000-0002-2785-9371

Data Availability

These data and codes have been shared on the Open Science Framework (OSF) and can be accessed via https://osf.io/ajfzr/

Supplemental Material

Supplemental material for this article is available online.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders: A meta-analysis of 36 studies. *Archives of General Psychiatry*, *68*(7), 724–731. https://doi.org/10.1001/archgenpsychiatry.2011.74

Bohn, K., & Fairburn, C. G. (2008). Clinical impairment assessment questionnaire (CIA 3.0). In C. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders* (pp. 315–317). Guilford Press.

Bradshaw, L. P., & Madison, M. J. (2016). Invariance properties for general diagnostic classification models. *International Journal of Testing*, *16*(2), 99–118. https://doi.org/10.1080/15305058.2015.1107076

Bray, B. C., Lanza, S. T., & Tan, X. (2015). Eliminating bias in classify-analyze approaches for latent class analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(1), 1–11. https://doi.org/10.1080/10705511.2014.935265

Cardi, V., Tchanturia, K., & Treasure, J. (2018). Premorbid and illness-related social difficulties in eating disorders: An overview of the literature and treatment developments. *Current Neuropharmacology*, *16*(8), 1122–1130. https://doi.org/10.2174/1570159X16666180118100028

Chen, J., & De La Torre, J. (2018). Introducing the general polytomous diagnosis modeling framework. *Frontiers in Psychology*, *9*, Article 1474. https://doi.org/10.3389/fpsyg.2018.01474

Compte, E. J., Sepulveda, A. R., & Torrente, F. (2015). A two-stage epidemiological study of eating disorders and muscle dysmorphia in male university students in Buenos Aires. *International Journal of Eating Disorders*, *48*(8), 1092–1101. https://doi.org/10.1002/eat.22448

Cui, Y., Gierl, M. J., & Chang, H.-H. (2012). Estimating classification consistency and accuracy for cognitive diagnostic assessment. *Journal of Educational Measurement*, *49*(1), 19–38. https://doi.org/10.1111/j.1745-3984.2011.00158.x

De La Torre, J., van der Ark, L. A., & Rossi, G. (2018). Analysis of clinical data from a cognitive diagnosis modeling framework. *Measurement and Evaluation in Counseling and Development*, *51*(4), 281–296. https://doi.org/10.1080/07481756.2017.1327286

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

Fitzsimmons-Craft, E. E., Karam, A. M., Monterubio, G. E., Taylor, C. B., & Wilfley, D. E. (2019). Screening for eating disorders on college campuses: A review of the recent literature. *Current Psychiatry Reports*, *21*(10), Article 101. https://doi.org/10.1007/s11920-019-1093-1

Forbush, K. T., Richson, B. N., Swanson, T. J., Thomeczek, M. L., Negi, S., Johnson, S. N., Chapa, D. A. N., Morgan, R. W., O'Brien, C. J., Gould, S. R., Christensen, K. A., & Chen, Y. (2022). Screening for eating disorders across genders in college students: Initial validation of the brief assessment of stress and eating. *International Journal of Eating Disorders*, *55*(11), 1553–1564. https://doi.org/10.1002/eat.23815

Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating disorders over the 2000-2018 period: A systematic literature review. *The American Journal of Clinical Nutrition*, *109*(5), 1402–1413. https://doi.org/10.1093/ajcn/nqy342

Garcia, F. D., Grigioni, S., Allais, E., Houy-Durand, E., Thibaut, F., & Déchelotte, P. (2011). Detection of eating disorders in patients: Validity and reliability of the French version of the SCOFF questionnaire. *Clinical Nutrition*, *30*(2), 178–181. https://doi.org/10.1016/j.clnu.2010.09.007

Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The Eating Attitudes Test: Psychometric features

and clinical correlates. *Psychological Medicine*, *12*(4), 871–878. https://doi.org/10.1017/S0033291700049163

George, A. C., Robitzsch, A., Kiefer, T., Gross, J., & Uenlue, A. (2016). The R Package CDM for cognitive diagnosis models. *Journal of Statistical Software*, *74*(2), 1–24. https://doi.org/10.18637/jss.v074.i02

Gideon, N., Hawkes, N., Mond, J., Saunders, R., Tchanturia, K., & Serpell, L. (2016). Development and psychometric validation of the EDE-QS, a 12 item short form of the Eating Disorder Examination Questionnaire (EDE-Q). *PLOS ONE*, *11*(5), Article e0152744. https://doi.org/10.1371/journal.pone.0152744

Graham, A. K., Trockel, M., Weisman, H., Fitzsimmons-Craft, E. E., Balantekin, K. N., Wilfley, D. E., & Taylor, C. B. (2019). A screening tool for detecting eating disorder risk and diagnostic symptoms among college-age women. *Journal of American College Health*, *67*(4), 357–366. https://doi.org/10.1080/07448481.2018.1483936

Griffiths, S., & Murray, S. B. (2017). Muscle dysmorphia: Strategies for treating a muscularity-oriented eating disorder. L. K. Anderson, S. B. Murray & W. H. Kaye (Eds.), *The clinical handbook of complex and atypical eating disorders* (pp. 235–252). Oxford University Press.

He, J., Brytek-Matera, A., Cooper, M., Cui, S., & Chen, G. (2022). Chinese translation of the Clinical Impairment Assessment (CIA 3.0): Psychometric properties and measurement invariance across sex and age in adolescents, young adults, and adult men. *Eating Behaviors*, *45*, 101623. https://doi.org/10.1016/j.eatbeh.2022.101623

He, J., Cui, T., Barnhart, W. R., Cui, S., Xu, Y., Compte, E. J., Murray, S. B., & Nagata, J. M. (2023). Validation of the muscularity-oriented eating test in adult women in China. *International Journal of Eating Disorders*, *56*(6), 1207–1218. https://doi.org/10.1002/eat.23927

He, J., Murray, S., Compte, E. J., Song, J., & Nagata, J. M. (2021). The Muscularity-Oriented Eating Test, Drive for Muscularity Scale, and Muscle Dysmorphic Disorder Inventory among Chinese men: Confirmatory factor analyses. *International Journal of Environmental Research and Public Health*, *18*(21), 11690. https://doi.org/10.3390/ijerph182111690

He, J., Sun, S., & Fan, X. (2021). Validation of the 12-item short form of the Eating Disorder Examination Questionnaire in the Chinese context: Confirmatory factor analysis and Rasch analysis. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *26*(1), 201–209. https://doi.org/10.1007/s40519-019-00840-3

Henson, R. A., Templin, J. L., & Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika*, *74*(2), 191–210. https://doi.org/10.1007/s11336-008-9089-5

Himmerich, H., Kan, C., Au, K., & Treasure, J. (2021). Pharmacological treatment of eating disorders, comorbid mental health problems, malnutrition and physical health consequences. *Pharmacology & Therapeutics*, *217*, 107667. https://doi.org/10.1016/j.pharmthera.2020.107667

Hu, J., Miller, M. D., Huggins-Manley, A. C., & Chen, Y.-H. (2016). Evaluation of model fit in cognitive diagnosis models. *International Journal of Testing*, *16*(2), 119–141. https://doi.org/10.1080/15305058.2015.1133627

Kazdin, A. E., Fitzsimmons-Craft, E. E., & Wilfley, D. E. (2017). Addressing critical gaps in the treatment of eating disorders. *International Journal of Eating Disorders*, *50*(3), 170–189. https://doi.org/10.1002/eat.22670

Ko, N., Tam, D. M., Viet, N. K., Scheib, P., Wirsching, M., & Zeeck, A. (2015). Disordered eating behaviors in university students in Hanoi, Vietnam. *Journal of Eating Disorders*, *3*(1), 1–7. https://doi.org/10.1186/s40337-015-0054-2

Kutz, A. M., Marsh, A. G., Gunderson, C. G., Maguen, S., & Masheb, R. M. (2020). Eating disorder screening: A systematic review and meta-analysis of diagnostic test characteristics of the SCOFF. *Journal of General Internal Medicine*, *35*, 885–893. https://10.1007/s11606-019-05478-6

Lechner, T. E., Gill, E. M., Drees, M. J., Hamady, C. M., & Ludy, M. J. (2019). Prevalence of disordered eating and muscle dysmorphia in college students by predominant exercise type. *International Journal of Exercise Science*, *12*(4), 989–1000.

Liu, R., & Jiang, Z. (2020). A general diagnostic classification model for rating scales. *Behavior Research Methods*, *52*(1), 422–439. https://doi.org/10.3758/s13428-019-01239-9

Ma, W. (2020). Evaluating the fit of sequential G-DINA model using limited-information measures. *Applied Psychological Measurement*, *44*(3), 167–181. https://doi.org/10.1177/0146621619843829

Ma, W., & de la Torre, J. (2020). An empirical Q-matrix validation method for the sequential generalized DINA model. *The British Journal of Mathematical and Statistical Psychology*, *73*(1), 142–163. https://doi.org/10.1111/bmsp.12156

Maguen, S., Hebenstreit, C., Li, Y., Dinh, J. V., Donalson, R., Dalton, S., Rubin, E., & Masheb, R. (2018). Screen for disordered eating: Improving the accuracy of eating disorder screening in primary care. *General Hospital Psychiatry*, *50*, 20–25. https://doi.org/10.1016/j.genhosppsych.2017.09.004

Mitchison, D., Mond, J., Griffiths, S., Hay, P., Nagata, J. M., Bussey, K., Trompeter, N., Lonergan, A., & Murray, S. B. (2022). Prevalence of muscle dysmorphia in adolescents: Findings from the EveryBODY study. *Psychological Medicine*, *52*(14), 3142–3149. https://doi.org/10.1017/S0033291720005206

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, *42*(5), 551–567. https://doi.org/10.1016/S0005-7967(03)00161-X

Morgan, J. F., Reid, F., & Lacey, J. H. (1999). The SCOFF questionnaire: Assessment of a new screening tool for eating disorders. *BMJ*, *319*(7223), 1467–1468. https://doi.org/10.1136/bmj.319.7223.1467

Murray, S. B., Brown, T. A., Blashill, A. J., Compte, E. J., Lavender, J. M., Mitchison, D., Mond, J. M., Keel, P. K., & Nagata, J. M. (2019). The development and validation of the muscularity-oriented eating test: A novel measure of muscularity-oriented disordered eating. *International Journal of Eating Disorders*, *52*(12), 1389–1398. https://doi.org/10.1002/eat.23144

Murray, S. B., Griffiths, S., & Mond, J. M. (2016). Evolving eating disorder psychopathology: Conceptualising muscularity-oriented disordered eating. *The British Journal of Psychiatry*, *208*(5), 414–415. https://doi.org/10.1192/bjp.bp.115.168427

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

Murray, S. B., Rieger, E., Touyz, S. W., & De la Garza García Lic, Y. (2010). Muscle dysmorphia and the DSM-V conundrum: Where does it belong? A review paper. *International Journal of Eating Disorders*, *43*(6), 483–491. https://doi.org/10.1002/eat.20828

Ouyang, X., Zhang, X., Räsänen, P., Koponen, T., & Lerkkanen, M. (2023). Subtypes of mathematical learning disability and their antecedents: A cognitive diagnostic approach. *Child Development*, *94*(3), 633–647. https://doi.org/10.1111/cdev.13884

Prnjak, K., Mitchison, D., Griffiths, S., Mond, J., Gideon, N., Serpell, L., & Hay, P. (2020). Further development of the 12-item EDE-QS: Identifying a cut-off for screening purposes. *BMC Psychiatry*, *20*(1), Article 146. https://doi.org/10.1186/s12888-020-02565-5

Revelle, W. (2022). *psych: Procedures for personality and psychological research*. Northwestern University. https://CRAN.R-project.org/package = psych

Smith, A. R., Ortiz, S. N., Forrest, L. N., Velkoff, E. A., & Dodd, D. R. (2018). Which comes first? An examination of associations and shared risk factors for eating disorders and suicidality. *Current Psychiatry Reports*, *20*(9), Article 77. https://doi.org/10.1007/s11920-018-0931-x

Solmi, F., Hatch, S. L., Hotopf, M., Treasure, J., & Micali, N. (2015). Validation of the SCOFF questionnaire for eating disorders in a multiethnic general population sample. *International Journal of Eating Disorders*, *48*(3), 312–316. https://doi.org/10.1002/eat.22373

Stice, E., Telch, C. F., & Rizvi, S. L. (2000). Development and validation of the Eating Disorder Diagnostic Scale: A brief self-report measure of anorexia, bulimia, and binge-eating disorder. *Psychological Assessment*, *12*(2), 123–131. https://doi.org/10.1037/1040-3590.12.2.123

Tan, Z., de la Torre, J., Ma, W., Huh, D., Larimer, M. E., & Mun, E.-Y. (2022). A tutorial on cognitive diagnosis modeling for characterizing mental health symptom profiles using existing item responses. *Prevention Science*. https://doi.org/10.1007/s11121-022-01346-8

Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, *11*(3), 287–305. https://doi.org/10.1037/1082-989X.11.3.287

Templin, J., Henson, R., Rupp, A., Jang, E., & Ahmed, M. (2008). *Cognitive diagnosis models for nominal response data*. National Council for Measurement in Education.

TODAY Study Group. (2022). Longitudinal association of depressive symptoms, binge eating, and quality of life with cardiovascular risk factors in young adults with youth-onset type 2 diabetes: The TODAY2 study. *Diabetes Care*, *45*(5), 1073–1081. https://doi.org/10.2337/dc21-1995

Tu, D., Zheng, C., Cai, Y., Gao, X., & Wang, D. (2018). A polytomous model of cognitive diagnostic assessment for graded data. *International Journal of Testing*, *18*(3), 231–252. https://doi.org/10.1080/15305058.2017.1396465

von Davier, M. (2008). A general diagnostic model applied to language testing data. *British Journal of Mathematical and Statistical Psychology*, *61*(2), 287–307. https://doi.org/10.1348/000711007X193957