



Exploring differential item functioning on eating disorder measures by food security status

Shannon M. O'Connor¹ · Vivienne M. Hazzard² · Hana F. Zickgraf³

Received: 3 June 2021 / Accepted: 6 August 2021 / Published online: 23 August 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

Purpose Food insecurity is associated with elevated eating disorder (ED) pathology, yet commonly used ED measures may not fully capture ED pathology in the context of food insecurity. The present study used differential item functioning (DIF) analyses to explore whether item endorsement on two commonly used ED questionnaires differed by food security status.

Methods Participants were 634 cisgender women recruited through Amazon's Mechanical Turk. DIF was explored for items on the Short Eating Disorder Examination Questionnaire (S-EDE-Q) and the Eating Disorder Diagnostic Scale for DSM-5 (EDDS-5). DIF analyses used a hybrid ordinal logistic regression/item response theory approach, with the presence of both statistical ($p < .01$) and clinical significance (pseudo $\Delta R^2 \geq .035$) indicating DIF.

Results There was no evidence of clinically significant DIF within the S-EDE-Q. Two items on the EDDS-5 exhibited statistically and clinically significant DIF, with moderate effect sizes. Specifically, compared to food-secure participants, food-insecure participants were more likely to report (1) eating large amounts of food when not physically hungry and (2) feeling disgusted, depressed, or guilty about overeating at lower levels of overall ED pathology but less likely to report these experiences at higher levels of overall ED pathology.

Conclusions Findings highlight a potential need to adapt ED measures to fully capture ED pathology in food-insecure populations.

Level of evidence Level III, well-designed cohort study.

Keywords Food insecurity · Differential item functioning · Eating disorders · Assessment · Measurement

Introduction

Food security is the state of having access to a sufficient quality and quantity of food needed for healthy living [1]. In 2018, 11% of the US population lacked such access, with the prevalence of food insecurity increasing dramatically during the 2020 economic downturn [2, 3]. Despite being partly defined by limited access, food insecurity in the United States and other developed countries is a risk factor for weight gain and obesity, likely due to reliance on

relatively inexpensive, energy-dense, highly processed food [4–6]. Food insecurity is also associated with psychological morbidity, including—strikingly—as much as a six-fold greater prevalence of eating disorder (ED) pathology compared to non-food-insecure individuals [7, 8].

Despite this strong association between food insecurity and ED pathology highlighting a critical need for additional research understanding the impact of food insecurity on ED pathology, existing ED measures were developed based on samples of individuals who were likely predominantly food secure. Indeed, most commonly used ED measures were largely developed in undergraduate and other convenience samples that may have over-represented female, White, and/or middle-class participants [e.g., 9–11]. Several authors have begun to address issues of diversity and ED stereotypes in measurement by examining the measurement of ED symptoms in Black or Hispanic respondents, and in men, but these studies still relied on undergraduate samples, and were not designed or recruited to be socioeconomically diverse

✉ Shannon M. O'Connor
oconnors@montclair.edu

¹ Department of Psychology, Montclair State University, Dickson Hall 119, 1 Normal Ave, Montclair, NJ 07043, USA

² Center for Bio-Behavioral Research, Sanford Research, Fargo, ND, USA

³ Department of Psychology, University of South Alabama, Mobile, AL, USA

[12–15]. Notably, there may be differences in the presentation of ED pathology in food-insecure populations. For instance, commonly used ED pathology measures typically inquire about dietary restraint due to weight/shape concerns, whereas “other-reason” dietary restraint is also linked to ED pathology [16]

People with food insecurity and ED pathology may be under-diagnosed and under-treated in medical settings. Clinicians and patients are both influenced by stereotypes that the prototypical patient with ED pathology is young, thin, white, and affluent [17]. This leads to the under-diagnosis and under-representation in treatment settings of people of color, men, people with overweight and obesity, and people who do not restrict for reasons related to weight and shape [18–22]. The use of validated screening measures in health-care settings may help reduce disparities in identifying and treating ED pathology in food-insecure populations. However, given that existing measures (1) limit their assessment of dietary restraint to focus on weight and shape concerns and (2) were developed in mostly food-secure samples, there is a need to understand whether existing ED symptom measures function appropriately in samples with food insecurity.

The underlying assumption of self-report trait and symptom measures is that they measure the same construct in the same way across groups and populations [23]. This assumption is the basis of research that compares levels of ED pathology between groups of interest: given the same level of actual ED pathology, members of different subgroups should endorse items at equivalent rates. If the probability of endorsing a specific item differs by group (e.g., food security status), this is evidence of differential item functioning (DIF). In psychological self-report instruments, DIF is commonly interpreted as group membership influencing the way participants understand or interpret an item [e.g., 14]. However, it may also be the case that DIF in ED measures suggests that ED symptoms themselves are differentially likely to appear in relation to underlying measured ED pathology across food security status. The aim of the present study was to assess for DIF in two commonly used measures of ED symptomatology in a sample of women with high/marginal food security and low/very low food security.

Methods

Participants

Participants included cisgender women who were recruited to participate in an online study about the “impact of food availability on eating and feeding behaviors” using Amazon’s Mechanical Turk (MTurk) [24]. Participants were compensated \$5 for participation in the study, which took an average of 28.49 min ($SD = 22.31$) to complete. The present

study is a secondary analysis of data collected between July 2019 and January 2020 for a study of intergenerational effects of food insecurity on home and parenting factors. Thus, approximately half of the participants were recruited to have experienced food insecurity during childhood, and the other half denied experiencing food insecurity during childhood. An additional inclusion criterion was having at least one child between the ages of six and eleven. Participants were women given the higher prevalence of ED pathology in women [25] and the increased likelihood that mothers would be more directly involved in the feeding of their child. To ensure quality responses, MTurk worker qualifications/requirements to participate included a HIT approval rate of greater than 95% (# of approved HITs that a worker has completed) and number of HITs approved greater than 100 (# of HITs that a worker has successfully completed since registering with MTurk). These requirements are consistent with other studies using MTurk samples (i.e., $\geq 90\%$ approval rate, ≥ 100 approved HITs; [26]). Participants also needed to reside within the United States. Prior to completing the online questionnaire, participants provided online informed consent. All study procedures were approved by the University of Chicago’s Institutional Review Board and participants were compensated via MTurk.

Data were collected from 805 women. Six quality checks required the participant to select the one sentence that does not make semantic sense (e.g., “Planes yell on the dream”) from a set of four syntactically correct sentences (e.g., “Boats are sailing on the lake”). The average completion time of the survey did not differ based on number of quality checks missed, with the exception of one comparison [missing 0 quality checks ($M = 26.47$ min, $SD = 11.92$) and 4 quality checks ($M = 17.84$ min, $SD = 10.29$)].¹ The internal consistency of well-validated, reliable measures that included items that are reverse scored was explored with differing numbers of quality checks missed. Cronbach’s alphas decreased when including individuals who missed 4 or more quality checks. Thus, participants who answered incorrectly at least four of six quality check questions were excluded ($n = 171$), resulting in an analytic sample of 634 women. The mean age of participants was 34.75 years ($SD = 7.40$, range 22–72) and the mean body mass index (BMI; calculated using self-reported height and weight) was 25.16 ($SD = 6.34$). Participants were economically diverse with 18.0% reporting a household income $< \$30,000$, 40.8% between $\$30,000$ and $\$59,999$, 23% between $\$60,000$ and $\$89,999$, and 18.3% $\geq \$90,000$. Participants predominantly identified as non-Hispanic (82.0%) and White (77.4%),

¹ Using the outlier labeling rule with a multiplier of 2.2 [27], $n = 27$ participants were removed from these analyses. All these participants spent more time on the survey (i.e., > 65 min).

followed by participants who identified as African American (18.3%), Asian (3.8%), American Indian/Alaskan Native (3.2%), Native Hawaiian/Pacific Islander (1.1%), and those who preferred not to provide their race or ethnicity (1.3% and 2.4%, respectively).

Measures

Food security

Food security was assessed using a self-report version of the 18-item United States Department of Agriculture (USDA) Food Security Survey Module [28]. Items assess eating patterns and anxiety about accessing food over the past 12 months (e.g., “In the last 12 months, were you ever hungry but didn't eat because there wasn't enough money for food?”). A total score was created from the sum of affirmative responses. Total scores can range from 0 to 18, with 0 indicating high food security ($n=203$), 1–2 indicating marginal food security ($n=37$), 3–7 indicating low food security ($n=117$), and 8–18 indicating very low food security ($n=277$) [28]. For analyses, high and marginal food security were combined to represent food security ($n=240$), and low and very low food security were combined to represent food insecurity ($n=394$). Very good internal consistency was observed in the current sample (Cronbach's $\alpha=0.84$).

Eating disorder pathology

ED pathology was assessed with the Short-Eating Disorder Examination Questionnaire (S-EDE-Q) [9, 29] and the Eating Disorder Diagnostic Scale for DSM-5 (EDDS-5) [11]. The S-EDE-Q is a 7-item, self-report questionnaire (adapted from the original 28-item EDE-Q; [30]) aimed at assessing ED attitudes and behaviors over the past 28 days. Items are rated on a 7-point ordered rating scale. The S-EDE-Q yields a global score (calculated as the mean of all seven items) reflecting participants' overall level of dietary restraint, weight/shape overvaluation, and body dissatisfaction [9, 10, 29]. In the current sample, Cronbach's α for the global S-EDE-Q was 0.93.

The revised version of the Eating Disorder Diagnostic Scale for DSM-5 (EDDS-5) [11] is a 23-item self-report measure that assesses diagnostic criteria for anorexia nervosa, bulimia nervosa, and binge eating disorder by asking the respondent about body image, eating habits, and compensatory behaviors over the last 3–6 months. It includes Likert, dichotomous (Yes/No), and frequency response options. Items 1–17, which can be summed to yield a symptom composite score, were used for the present analyses. Research with the EDDS-4 (the DSM-IV version of the same assessment) suggests good internal consistency and criterion validity with interview-based measures of disordered

eating [11]. In the current sample, Cronbach's α for the EDDS-5 symptom composite was 0.86.

Data analysis

The `lordif` package in R [31] was used to conduct differential item functioning (DIF) analyses examining whether endorsement of items on the S-EDE-Q and EDDS-5 differed across food-secure and food-insecure groups after conditioning on overall ED pathology as measured by each respective scale. There were no missing data on the S-EDE-Q or EDDS-5 among participants with food security data. The `lordif` package uses a hybrid ordinal logistic regression/item response theory (IRT) approach to test for DIF. Instead of conditioning on an observed composite score, this approach uses IRT-based estimates of the latent variable (θ)—in this case, overall ED pathology—as the matching criterion. Predictors of item responses were entered into ordinal logistic regression models in a three-step approach: the IRT-based estimate of the latent variable representing overall ED pathology was entered into the model first (Model 1), the group variable representing food security status was added to the model next (Model 2), and, finally, an interaction term between the two variables ($\theta_{\text{overall ED pathology}} \times \text{food security status}$) was added to the model (Model 3).

Uniform and non-uniform DIF were examined. Uniform DIF suggests the possibility of consistent item-level bias by group membership in the same direction across all levels of the underlying trait (i.e., analogous to confounding), whereas non-uniform DIF suggests the possibility of item-level bias by group membership that differs by level of the underlying trait (i.e., analogous to moderation [32]). Statistically significant uniform DIF was indicated by a significant likelihood ratio χ^2 difference test comparing model fit between Models 1 and 2, and statistically significant non-uniform DIF was indicated by a significant likelihood ratio χ^2 difference test comparing model fit between Models 2 and 3. To account for the high number of statistical tests performed, criteria for a statistically significant likelihood ratio χ^2 difference test was set at $\alpha=0.01$, following standard practice for DIF analyses as established by Zumbo [33]. To determine clinical significance of the DIF results, the Jodoin and Gierl's [34] effect size guidelines were used, such that changes in pseudo R^2 between each step were characterized as follows: negligible (pseudo $\Delta R^2 < 0.035$), moderate (pseudo $\Delta R^2 \geq 0.035$ and < 0.070), or large (pseudo $\Delta R^2 \geq 0.070$). McFadden's pseudo R^2 was used, as it is considered one of the most generally applicable versions of pseudo R^2 for logistic regression due to its relative independence from the base rate of the outcome [35]. Both statistical ($p < 0.01$) and clinical significance (pseudo $\Delta R^2 \geq 0.035$) must be demonstrated for evidence of uniform or non-uniform DIF [34, 36]. Item characteristic curves—depicting the likelihood of

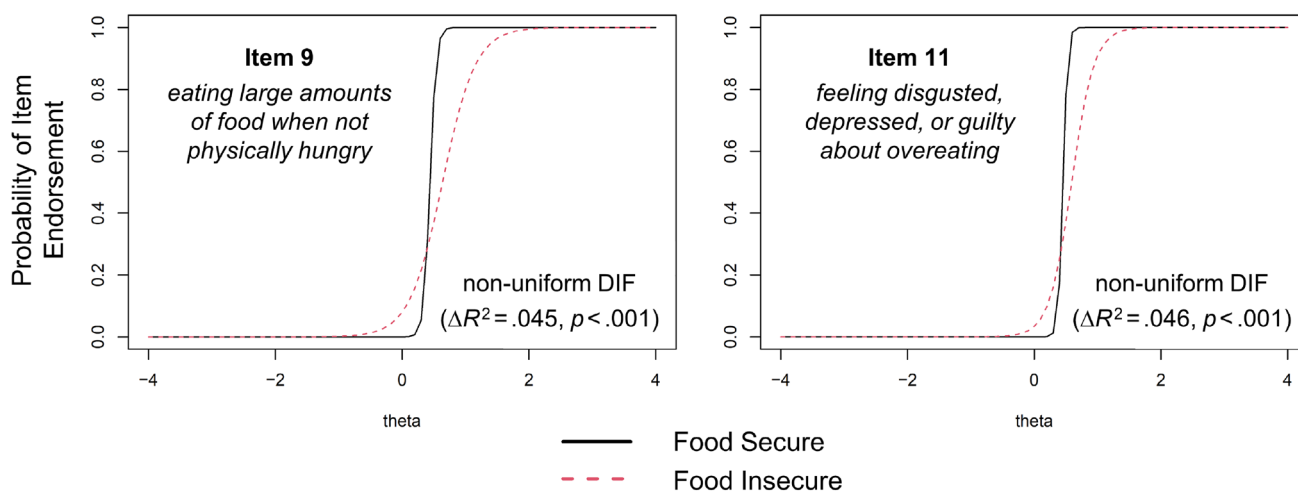


Fig. 1 Item characteristic curves for items from the Eating Disorder Diagnostic Scale for DSM-5 (EDDS-5) demonstrating both statistically and clinically significant differential item functioning for the food insecure group compared to the food secure group

endorsing a particular item score at various levels of the underlying latent variable representing overall ED pathology—were plotted for items demonstrating both statistically and clinically significant DIF.

Results

Short-eating disorder examination-questionnaire

S-EDE-Q global scores ranged from 0 to 6, with $M=2.75$ ($SD=1.74$) in the full sample. DIF model comparison results for the S-EDE-Q are presented in Table S1 (available online). Item 1 (“Have you been consciously trying to restrict the amount of food you eat to influence your shape or weight?”) demonstrated statistically significant, but not clinically significant, non-uniform DIF by food security status. No evidence of DIF by food security status was detected for any other items on the S-EDE-Q.

Eating disorder diagnostic scale for DSM-5

EDDS-5 raw symptom composite scores ranged from 0 to 117, with $M=30.62$ ($SD=27.61$) in the full sample. DIF model comparison results for the EDDS-5 are presented in Table S2 (available online). A total of fourteen items (items 1, 2, 3, 4, 8, 9, 10, 11, 12, 13, 14, 15, 16, and 17) of the EDDS-5 demonstrated statistically significant DIF by food security status, but only two of these items also demonstrated clinically significant DIF. More specifically, results indicated statistically and clinically significant non-uniform DIF for items 9 (“During episodes of overeating with a loss of control, did you eat large amounts of food when you didn’t feel physically hungry?”) and 11 (“During episodes of

overeating with a loss of control, did you feel disgusted with yourself, depressed, or very guilty after overeating?”). Compared to food-secure participants, food-insecure participants were more likely to endorse items 9 and 11 at lower levels of overall ED pathology but less likely to endorse these items at higher levels of overall ED pathology (Fig. 1). Effect sizes were moderate across both items demonstrating DIF, with McFadden pseudo ΔR^2 values ranging from 0.045 to 0.046.

Discussion

The present study examined the possibility of DIF across food-secure and food-insecure women for items on the S-EDE-Q and EDDS-5, two commonly used measures of ED pathology. No evidence of clinically significant DIF was observed within the S-EDE-Q, but two items on the EDDS-5 exhibited statistically and clinically significant DIF with moderate effect sizes. These items (i.e., eating large amounts of food when not physically hungry and feeling disgusted/depressed/guilty about overeating) demonstrated poorer ability to differentiate between different levels of ED pathology among individuals experiencing food insecurity. In other words, both items were endorsed at relatively more consistent rates regardless of levels of ED pathology for individuals with food insecurity relative to those who were food secure.

Eating large amounts of food when not physically hungry was endorsed less at higher levels of ED pathology in the food-insecure group. This finding may suggest that eating in the absence of physical hunger may not be sensitive to disordered overeating in people with food insecurity, perhaps due to the more frequent experience of physiological hunger in this population. Alternatively, this item’s poorer ability to

differentiate across ED pathology severity may suggest that eating large amounts of food may be considered less “pathological” (i.e., less connected to other types of ED pathology) in food-insecure populations relative to food-secure populations. The DIF exhibited by the disgusted/depressed/guilty about overeating item may also support this explanation. Indeed, this item assessing negative affect related to overeating was also endorsed less at higher levels of ED pathology, perhaps reflecting that eating large amounts when food is available may be a normative, and potentially adaptive behavior in the context of food insecurity [37].

Future research is needed to understand how endorsement of ED pathology items may differ with severity of food insecurity, as there may be differences in ability to engage in specific symptoms of ED pathology across the spectrum of food insecurity severity. For example, marginal/low food security is more often characterized by concern about maintaining access to food, or lacking access to a variety of healthy, nutritious foods, but not by lacking food and going hungry. Many households with marginal/low food security have an overabundance of low-cost, high energy–density, palatable food. Conversely, very low food security is characterized by reduced food quantity. These differences may play a role in the ability to engage in particular ED behaviors, such as objective overeating and binge eating, and put marginal/low food-secure individuals in an environment with more triggers for binge eating (e.g., relative over-availability of highly palatable, highly processed foods in their home, and neighborhood food environments; [4–6]) than individuals experiencing very low food insecurity. This idea is also consistent with several findings that the relationship between food insecurity and obesity is curvilinear, with the strongest association and risk for obesity at moderate levels of food insecurity [38]. More research is needed to better understand how different experiences of food insecurity might influence ED pathology.

Moving forward, the moderate effect sizes observed for the two items demonstrating DIF on the EDDS-5 highlight that there should be some concern with using the EDDS-5 symptom composite in food-insecure samples but do not invalidate past results, as the items demonstrating DIF in the present study have not been the focus of past results. Our findings instead emphasize the potential need to adapt various items from commonly used ED measures to ensure ED pathology is being captured appropriately considering the possibility that some eating behaviors potentially serving adaptive purposes within food-insecure populations. Of note, although we expected that items assessing dietary restraint might exhibit DIF, this was not observed in either ED questionnaire. Further research, however, should explore the extent to which dietary restraint motivated by reasons other than weight/shape may be under-reported by these measures.

Strengths and limitations

The current study design has numerous strengths, including the recruitment of a large sample of women who varied in their food security status and who reported a range of severity in ED pathology. To our knowledge, this is the first study designed to test the validity and performance of validated ED measures for people with food insecurity. The S-EDE-Q and EDDS-5 are two of the most commonly used measures of ED pathology. The S-EDE-Q is brief enough to be used as a screening instrument in healthcare settings, and the current findings suggest that this measure assesses ED pathology similarly for women across food security status. However, this study was not without limitations, including the use of a non-clinical population, limiting the generalizability of the findings to clinical samples. Additionally, while items on the EDDS-5 and original EDDS are quite similar and, thus, findings would likely generalize, it is less clear whether the findings from the S-EDE-Q would generalize to the full range of items on the EDE-Q. Further, the use of MTurk to collect a large sample of respondents may limit the generalizability of the data within food-insecure populations; the current study only included participants who read and speak English fluently, and the content of items was not adjusted to make measures more accessible to individuals with lower reading comprehension skills [7]. Given that the full food-insecure population may have relatively lower rates of English fluency and/or literacy than the general population, the current findings may not be generalizable to all food-insecure individuals. However, it is important to note that past studies have demonstrated a substantial proportion of MTurk workers report being underemployed (e.g., [39]), suggesting that exploration of food insecurity within this population would likely be fruitful. Finally, given that the present study was conducted in a U.S. sample, it is possible that findings may not generalize to less developed countries, where food insecurity is often associated with underweight [40].

Conclusion

Findings from the present study highlight a need for continued and more comprehensive exploration of ED pathology in food-insecure populations to evaluate the utility of commonly used ED measures in food-insecure samples. For example, future research is needed to better understand how differences in severity of food insecurity may influence the endorsement of various ED symptoms. Nevertheless, results of this study suggest that the S-EDE-Q may be useful to screen individuals in medical settings. Finally, research designs that allow stakeholders to participate in study design, including qualitative approaches, may be helpful in refining measurement to represent the experiences of food-insecure individuals with ED pathology.

What is already known on this subject?

Food insecurity is associated with elevated rates of ED pathology in adults [e.g., 7–8]. However, measures of ED pathology were largely developed on samples that represent the prototypical ED patient (e.g., White, middle-class, females; e.g., [9–11]), which may lead to under-diagnosis of ED pathology in populations that deviate from this stereotype.

What this study adds?

This study explores whether items on two commonly used ED pathology measures (EDDS-5 and S-EDE-Q) operate similarly in a food-insecure population compared to a food-secure population. Findings indicate no DIF on the S-EDE-Q, but two items on the EDDS-5 demonstrated DIF, showing lower ability to differentiate across levels of ED pathology severity in food-insecure relative to food-secure populations. While DIF exhibited within the EDDS-5 was moderate in effect size and only affected two items, findings highlight a need to consider how the wording of items assessing ED pathology may contribute to differential endorsement by food security status.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40519-021-01289-z>.

Funding All the authors were supported by the National Institute of Mental Health of the National Institutes of Health of the National Institutes of Health under award number T32 MH082761. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Availability of data and material Available upon request.

Code availability Available upon request.

Declarations

Conflict of interest The authors have no conflicts to declare.

Ethical approval All procedures involving human participants were in accordance with the ethical standards of the University of Chicago Institutional Review Board and the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Consent to participate Informed consent was obtained from all participants included in this study.

References

- Coleman-Jensen A, Rabbitt M, Gregory CA, Singh A (2019) Household food security in the United States in 2018. Economic Research Service, U.S. Department of Agriculture, Washington DC, pp 1–47
- Niles MT, Bertmann F, Morgan EH, Wentworth T, Biehl E, Neff R (2020) Food Access and security during coronavirus: a vermont study. College of Agriculture and Life Sciences Faculty Publications; University of Vermont, Burlington, VT, USA
- Wolfson JA, Leung CW (2020) Food insecurity and COVID-19: disparities in early effects for US adults. *Nutrients* 12:1658. <https://doi.org/10.3390/nu12061648>
- Connell CL, Zoellner JM, Yadrick MK, Chekuri SC, Crook MLB, Bogle ML (2012) Energy density, nutrient adequacy, and cost per serving can provide insight into food choices in the lower Mississippi Delta. *J Nutr Educ Behav* 44:148–153. <https://doi.org/10.1016/j.jneb.2011.02.003>
- Drewnowski A, Darmon N (2005) Food choices and diet costs: an economic analysis. *J Nutr* 135:900–904. <https://doi.org/10.1093/jn/135.4.900>
- Leung CW, Epel ES, Ritchie LD, Crawford PB, Laraia BA (2014) Food insecurity is inversely associated with diet quality of lower-income adults. *J Acad Nutr Diet* 114:1943–1953. <https://doi.org/10.1016/j.jand.2014.06.353>
- Becker CB, Middlemass K, Taylor B, Johnson C, Gomez F (2017) Food insecurity and eating disorder pathology. *Int J Eat Disord* 50:1031–1040. <https://doi.org/10.1002/eat.22735>
- Becker CB, Middlemass K, Gomez F, Martinez-Abrego A (2019) Eating disorder pathology among individuals living with food insecurity: a replication study. *Clin Psychol Sci* 7:1144–1158. <https://doi.org/10.1177/2167702619851811>
- Machado PPP, Grilo CM, Crosby RD (2018) Replication of a modified factor structure for the eating disorder examination-questionnaire: extension to clinical eating disorder and non-clinical samples in Portugal. *Eur Eat Disord Rev* 26:75–80. <https://doi.org/10.1002/erv.2569>
- Machado PPP, Grilo CM, Rodrigues TF, Vaz AR, Crosby RD (2020) Eating disorder examination—questionnaire short forms: a comparison. *Int J Eat Disord* 53:937–944. <https://doi.org/10.1002/eat.23275>
- Stice E, Fisher M, Martinez E (2004) Eating disorder diagnostic scale: additional evidence of reliability and validity. *Psychol Assess* 16:60–71. <https://doi.org/10.1037/1040-3590.16.1.60>
- Kelly NR, Cotter EW, Mazzeo SE (2012) Eating disorder examination questionnaire (EDE-Q): norms for Black women. *Eat Behav* 13:429–432. <https://doi.org/10.1016/j.eatbeh.2012.09.001>
- McEntee ML, Serier KN, Smith JM, Smith JE (2021) The sum is greater than its parts: intersectionality and measurement validity of the eating disorder examination questionnaire (EDE-Q) in Latinx undergraduates in the United States. *Sex Roles* 84:102–111. <https://doi.org/10.1007/s11199-020-01149-7>
- Schaefer LM, Anderson LM, Simone M, O'Connor SM, Zickgraf H, Anderson DA, Rodgers RF, Thompson JK (2019) Gender-based differential item functioning in measures of eating pathology. *Int J Eat Disord* 52:1047–1051. <https://doi.org/10.1002/eat.23126>
- Serier KN, Smith JE, Yeater EA (2018) Confirmatory factor analysis and measurement invariance of the eating disorder examination questionnaire (EDE-Q) in a non-clinical sample of non-hispanic White and hispanic women. *Eat Behav* 31:53–59. <https://doi.org/10.1016/j.eatbeh.2018.08.004>
- Middlemass KM, Cruz J, Gamboa A, Johnson C, Taylor B, Gomez F, Becker CB (2020) Food insecurity and dietary restraint in a diverse urban population. *Eat Disord*. <https://doi.org/10.1080/10640266.2020.1723343>
- Sonneville KR, Lipson SK (2018) Disparities in eating disorder diagnosis and treatment according to weight status, race/ethnicity, socioeconomic background, and sex among college students. *Int J Eat Disord* 51:518–526. <https://doi.org/10.1002/eat.22846>
- Curran L, Schmidt U, Waller G (2007) Variables that influence diagnosis and treatment of the eating disorders within primary

- care settings: a vignette study. *Int J Eat Disord* 40:257–262. <https://doi.org/10.1002/eat.20355>
19. Gordon KH, Perez M, Joiner TE Jr (2002) The impact of racial stereotypes on eating disorder recognition. *Int J Eat Disord* 32:219–224. <https://doi.org/10.1002/eat.10070>
 20. Gordon KH, Brattole MM, Wingate LR, Joiner TE Jr (2006) The impact of client race on clinician detection of eating disorders. *Behav Ther* 37:319–325. <https://doi.org/10.1016/j.beth.2005.12.002>
 21. Striegel-Moore RH, Dohm FA, Kraemer HC, Taylor CB, Daniels S, Crawford PB et al (2003) Eating disorders in White and Black women. *Am J Psychiatry* 160:1326–1331. <https://doi.org/10.1176/appi.ajp.160.7.132>
 22. Veillette LA, Serrano JM, Brochu PM (2018) What's weight got to do with it? Mental health trainees' perceptions of a client with anorexia nervosa symptoms. *Front Psychol* 9:2574. <https://doi.org/10.3389/fpsyg.2018.02574>
 23. Osterlind SJ, Everson HT (2011) Description of DIF. Differential item functioning. SAGE Publications, Thousand Oaks, CA, pp 8–12. <https://doi.org/10.4135/9781412993913>
 24. Buhrmester MD, Kwang T, Gosling SD (2011) Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspect Psychol Sci* 6:3–5. <https://doi.org/10.1177/1745691610393980>
 25. American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, 5th edn. Author, Washington, DC
 26. Chmielewski M, Kucker SC (2020) An MTurk crisis? Shifts in data quality and the impact on study results. *Soc Psychol Personal Sci* 11:464–473. <https://doi.org/10.1177/1948550619875149>
 27. Hoaglin DC, Iglewicz B, Tukey JW (1986) Performance of some resistant rules for outlier labeling. *J Am Stat Assoc* 81:991–999. <https://doi.org/10.2307/2289073>
 28. Bickel G, Nord M, Hamilton W (2006) US household food security survey module: three-stage design with screeners. US Department of Agriculture, Economic Research Service, Washington, DC
 29. Grilo CM, Reas DL, Hopwood CJ, Crosby RD (2015) Factor structure and construct validity of the eating disorder examination questionnaire in college students: further support for a modified brief version. *Int J Eat Disord* 48:284–289. <https://doi.org/10.1002/eat.22358>
 30. Fairburn CG, Beglin SJ (1994) Assessment of eating disorders: interview or self-report questionnaire? *Int J Eat Disord* 16:363–370. [https://doi.org/10.1002/1098-108X\(199412\)16:4%3c363::AID-EAT2260160405%3e3.0.CO;2-%23](https://doi.org/10.1002/1098-108X(199412)16:4%3c363::AID-EAT2260160405%3e3.0.CO;2-%23)
 31. Choi SW, Gibbons LE, Crane PK (2011) lordif: an R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and monte Carlo simulations. *J Stat Softw* 39:1–30. <https://doi.org/10.18637/jss.v039.i08>
 32. Crane PK, Gibbons LE, Jolley L, Van Belle G (2006) Differential item functioning analysis with ordinal logistic regression techniques: DIFdetect and difwithpar. *Med Care* 44:115–123. <https://doi.org/10.1097/01.mlr.0000245183.28384.ed>
 33. Zumbo BD (1999) A handbook on the theory and methods of differential item functioning (DIF). National Defense Headquarters, Ottawa, pp 1–57
 34. Jodoin MG, Gierl MJ (2001) Evaluating type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Appl Meas Educ* 14:329–349. https://doi.org/10.1207/S15324818AME1404_2
 35. Menard S (2000) Coefficients of determination for multiple logistic regression analysis. *Am Stat* 54:17–24. <https://doi.org/10.1080/00031305.2000.10474502>
 36. Gadermann AM, Chen MY, Emerson SD, Zumbo BD (2018) Examining validity evidence of self-report measures using differential item functioning: an illustration of three methods. *Methodol* 14:164–175. <https://doi.org/10.1027/1614-2241/a000156>
 37. Connell CL, Lofton KL, Yadrick K, Rehner TA (2005) Children's experiences of food insecurity can assist in understanding its effect on their well-being. *J Nutr* 135:1683–1690. <https://doi.org/10.1093/jn/135.7.1683>
 38. Franklin B, Jones A, Love D et al (2012) Exploring mediators of food insecurity and obesity: a review of recent literature. *J Community Health* 37:253–264. <https://doi.org/10.1007/s10900-011-9420-4>
 39. Behrend TS, Sharek DJ, Meade AW, Wiebe EN (2011) The viability of crowdsourcing for survey research. *Behav Res Methods* 43:800–813. <https://doi.org/10.3758/s13428-011-0081-0>
 40. Moradi S, Mirzababaei A, Dadfarma A, Rezaei S, Mohammadi H, Jannat B, Mirzaei K (2019) Food insecurity and adult weight abnormality risk: a systematic review and meta-analysis. *Eur J Nutr* 58:45–61. <https://doi.org/10.1007/s00394-018-1819-6>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.