Weight Concerns Scale Applied to College Students: Comparison Between Pencil-and-Paper and Online Formats

Juliana Chioda Ribeiro Dias, DDS, MSc, PhD,¹ João Maroco, BSc, MSc, PhD,²
and Juliana Alvares Duarte Bonini Campos, DDS, MSc, PhD¹

Abstract

Online data collection is becoming increasingly common and has some advantages compared to traditional paper-and-pencil formats, such as reducing loss of data, increasing participants' privacy, and decreasing the effect of social desirability. However, the validity and reliability of this administration format must be established before results can be considered acceptable. The aim of this study was to evaluate the validity, reliability, and equivalence of paper-and-pencil and online versions of the Weight Concerns Scale (WCS) when applied to Brazilian university students. A crossover design was used, and the Portuguese version of the WCS (in both paper-and-pencil and online formats) was completed by 100 college students. The results indicated adequate fit in both formats. The simultaneous fit of data for both groups was excellent, with strong invariance between models. Adequate convergent validity, internal consistency, and mean score equivalence of the WCS in both formats were observed. Thus, the WCS presented adequate reliability and validity in both administration formats, with equivalence/stability between answers.

Introduction

ALTHOUGH EPIDEMIOLOGICAL STUDIES conventionally use paper-and-pencil questionnaires for data collection, currently, due to technological advancement and information globalization, many researchers have started using the Internet as an alternative to this procedure.¹

According to Calbring and Brunt,² online administration has advantages such as reduced data loss and/or duplication, immediate data entry, and guaranteed privacy of the participant. Nosek and Banaji,³ and Evans and Garcia⁴ report that completing questionnaires online also decreases the effect of social desirability, or the tendency to answer questions in a manner considered more socially acceptable. More comprehensive data collection and representativeness of results are other advantages of this form of data collection.⁵

On the other hand, online completion may have some disadvantages. Yetter and Capaccioli⁶ highlighted that familiarity with technology is not uniform across population groups, especially when taking into account socioeconomic status. Thus, they suggest that online applications may exclude individuals who do not have access to the necessary tools, or who do not know how to use them correctly and safely.

Riva and Teruzzi⁷ indicate that administration of questionnaires via the Internet may be influenced by an inability to control the environment where participants complete the

instrument. Furthermore, they report that in this situation, it is not possible to ensure that all respondents receive the same stimulus for completing the questionnaire. However, this is perceived by Brock and Barry⁸ as an advantage, since the participant is not approached by the researcher to participate in the study and therefore does not experience any kind of influence, thus minimizing the social desirability effect.

Increased access to the Internet and the potential for increased scope of data collection have encouraged the development of online research; several instruments have been adapted for this purpose. However, changing the presentation/form of instruments can affect the quality of information and/or the psychometric properties of the data collected. Some studies have shown that adapting an instrument to an online version does not affect the answers given or the psychometric properties of the data.^{1,5,9,10} Nevertheless, other studies point to significant interference.^{11–14} This lack of consensus can be explained by differences in the theoretical constructs that are evaluated, which suggests the need for specific studies to assess interference when different constructs are measured.

Body image can be regarded as a construct that is highly susceptible to social desirability, since it is strongly influenced by cultural, social, and psychological variables.¹⁵ As it is a latent variable, it is estimated through the evaluation of specific related aspects, and usually these concepts are

¹Pós-graduação em Alimentos e Nutrição, Faculdade de Ciências Farmacêuticas, Universidade Estadual Paulista (UNESP), Araraquara, São Paulo, Brazil.

²Unidade de Investigação em Psicologia e Saúde, UIPE/ISPA, Instituto Universitário, Lisbon, Portugal.

assessed through psychometric scales completed in a paper-and-pencil format.

The Weight Concerns Scale (WCS), proposed by Killen et al.¹⁶ in an English and American context, aims at estimating concern with weight associated with body image in women. The WCS is a one-dimensional instrument, consisting of five items answered using a 7-point Likert scale. Only one study¹⁷ was found that applied the WCS in both paper-and-pencil and online formats and compared the results from the two application methods. The authors confirmed the equivalence of the two formats. Although this study provided important information for researchers, it did not assess construct validity, in terms of factorial and convergent validity, nor did it evaluate the reliability of information obtained from the application of the scale in different modalities. Evaluation of these properties is of paramount importance in understanding how a construct is measured in a sample, and how different administration methods may interfere with its structure.¹ This information is relevant when choosing the best instrument and administration method, and can be verified based on the explained variance, validity, and reliability of the model. These assessments are essential for ensuring quality of the data.

The present study was therefore developed in order to estimate the validity, reliability, and equivalence of paper-and-pencil and online versions of the WCS when applied to university students.

Methods

Study design

A crossover design was used, in which half of the participants (group A) were randomly selected to complete the scale in the paper-and-pencil format and the other half (group B) the online format. After a washout period of 1 week, the groups were reversed, and the participants in group A completed the scale in the online format and group B in the paper-and-pencil format.

Administration using the two methods of completing the scale (paper-and-pencil and online) were matched using identification numbers for the participants. To identify learning effects that can occur for instruments completed twice by the same participants, the method proposed by Guyatt et al.¹⁸ and Streiner and Norman¹⁹ was used. In the present study, this effect was not observed (group A: $\sigma^2_{\text{time } 1} = 1.1$, $\sigma^2_{\text{time } 2} = 1.1$; group B: $\sigma^2_{\text{time } 1} = 1.0$, $\sigma^2_{\text{time } 2} = 1.1$), which indicates that the washout period was sufficient.

Procedures

For the online format, a Google Docs® application was used, and the access link generated was e-mailed to the participants. All scale items could be viewed simultaneously, and participants could change their answers as many times as they wished before submitting them.

The paper-and-pencil format was completed individually at a scheduled time and in a room reserved for this purpose. The researcher did not interfere during the completion of the questionnaire.

Sampling

Estimation of the minimum sample size was performed using power analysis, which indicated 5–10 participants per

item/parameter evaluated in the structural model. Thus, based on the degrees of freedom of the model, a significance level of 5%, and a power of at least 80%, an estimate of the sample size of 50–100 participants was obtained.²⁰ One hundred volunteer female students from the Faculty of Pharmaceutical Sciences, Universidade Estadual Paulista—UNESP, campus of Araraquara, São Paulo, Brazil participated in the study. Nonprobabilistic, convenience sampling was employed. The average age of participants was 23.34 ($SD = 4.88$) years.

Instruments

The Portuguese version of the Weight Concerns Scale (WCS) proposed by the authors of this paper was used.

Statistical analysis

Psychometric sensitivity. The psychometric sensitivity of the WCS items, when completed in the paper-and-pencil format and online, was evaluated through summary measures. Items were considered to present psychometric sensitivity when the response data approached a normal distribution. Absolute values of skewness (Sk) and kurtosis (Ku) below 3 and 7, respectively, were indicative of no severe deviations from the normal distribution²⁰ and, thus, adequate psychometric sensitivity indicators.

Construct validity. The construct validity of the WCS (applied in paper-and-pencil format and online) was estimated using factorial and convergent validity, as described below.

Factorial validity was assessed through confirmatory factor analysis, using local goodness of fit indexes as the factor weights ($\lambda > 0.5$) and the chi-square by the degrees of freedom ratio (χ^2/df), comparative fit index (CFI), normed fit index (NFI), and root mean square error of approximation (RMSEA) as overall fit indices. The fit of the model was considered adequate when $\chi^2/df \leq 3.0$, CFI and NFI ≥ 0.90 , and RMSEA ≤ 0.10 . To compare the WCS models (paper-and-pencil format and online), information theory indices were computed, including the Akaike information criterion (AIC), Bayesian information criterion (BIC), and Browne–Cudeck criterion (BCC). The best model fit was indicated by lower values for the majority of these indices.²⁰

Convergent validity was estimated through average variance extracted (AVE) and composite reliability (CR), as proposed by Fornell and Larcker,²¹ and were considered adequate when AVE ≥ 0.50 and CR ≥ 0.70 .

Factorial invariance. A multi-group analysis was performed to assess the stability of the models for the data collected with each method (paper-and-pencil and online). Invariance of the models obtained was tested using the chi-square difference ($\Delta\chi^2$) for the factor weights (λ) and for specific factors (covariance and residue) of the model.²⁰

Concurrent validity. The Student's t test for repeated measures was used to examine the tendency to report more or fewer symptoms in the two groups (paper-and-pencil and online). For decision making, the level of significance was 5%.

TABLE 1. SUMMARY MEASURES FOR THE ITEM OF THE WEIGHT CONCERNS SCALE COMPLETED BY UNIVERSITY STUDENTS IN PAPER-AND-PENCIL AND ONLINE FORMATS

WCS	Paper-and-pencil					Online				
	Mean	Median	SD	Skewness	Kurtoses	Mean	Median	SD	Skewness	Kurtoses
It1	2.58	2.50	1.01	0.20	−0.65	2.66	3.00	1.12	0.09	−1.02
It2	2.54	2.00	1.21	0.39	−0.78	2.51	2.00	1.26	0.50	−0.86
It3	2.84	1.00	2.42	0.91	−0.91	2.80	1.00	2.36	0.95	−0.77
It4	1.77	2.00	0.75	0.41	−1.11	1.70	2.00	0.76	0.57	−1.05
It5	2.81	3.00	1.25	0.54	−1.05	2.90	3.00	1.34	0.13	−1.08

WCS, Weight Concerns Scale; SD, standard deviation.

Reliability. Internal consistency was estimated using the standardized Cronbach's alpha coefficient.²²

Statistical analyses were conducted using AMOS v20.0 (SPSS, Inc., Chicago, IL) and IBM SPSS Statistics for Windows v20.0 (IBM Corp., Armonk, NY).

Ethical aspects

Invitations to participate in this study were made verbally by the researchers to university students. After presentation of the study's objectives, they were informed that their participation was voluntary. Only students who agreed with free and informed consent participated in the study.

This study was approved by the Ethics Committee in Research of the Faculty of Pharmaceutical Sciences—UNESP (protocol: CEP/FCF/CAR nr.16/2010).

Results

Summary measures of the WCS items for the two completion formats (paper-and-pencil and online) are presented in Table 1. All items showed adequate psychometric sensitivity for both completion formats.

Figure 1 presents the WCS factorial model for the pencil-and-paper and online formats. An appropriate fit of the factor structure to the sample was observed, except for the RMSEA index for the online format. Based on the information theory indexes, there was greater parsimony when completing the WCS on paper. The explained variance was higher in the

online format. Convergent validity and internal consistency were adequate for both administrations (paper-and-pencil: AVE=0.46, CR=0.80, α =0.74; online: AVE=0.47, CR=0.81, α =0.76).

The simultaneous adjustment of the groups was adequate (λ =0.43–0.67, χ^2/df =1.333, CFI=0.978, GFI=0.947, RMSEA=0.041), with strong invariance between models (λ : $\Delta\chi^2$ =1.074, p =0.898; covariance: $\Delta\chi^2$ =0.133, p =0.715; residue: $\Delta\chi^2$ =2.392, p =0.793), confirming the factorial equivalence of the different formats. No significant difference was observed between mean scores of the WCS when completed in different formats (t =0.729, p =0.469).

Discussion

In this study, the psychometric equivalence of the WCS (in paper-and-pencil and online formats) was evaluated and confirmed (Fig. 1). The results point to the possibility of using the Internet for WCS data collection.

Given that the WCS is a simple instrument to estimate concern with weight related to body image, with a short administration time, using the Internet can have benefits. These include privacy and confidentiality when completing the instrument,² minimization of the effect of social desirability,⁴ and increased sample coverage in screening studies.⁵ These advantages can be important when conducting population screening to assess the extent of distortions regarding concerns with weight, and could provide information

FIG. 1. Factor structure of the Portuguese version of the Weight Concerns Scale (WCS) applied to university students in paper-and-pencil (σ^2 =40%; χ^2/df =1.614, CFI=0.979, NFI=0.967, RMSEA=0.079, AIC=28.070, BIC=54.122, BCC=29.361) and online formats (σ^2 =61%; χ^2/df =2.996, CFI=0.935, NFI=0.941, RMSEA=0.142, AIC=34.982, BIC=61.033, BCC=36.272). CFI, comparative fit index; NFI, normed fit index; RMSEA, root mean square error of approximation; AIC, Akaike information criterion; BIC, Bayesian information criterion; BCC, Browne–Cudeck criterion.

for the development of educational, preventive, and/or therapeutic strategies that may assist in the prevention of disorders related to body image.

The methodological rigor of this study is one aspect to be highlighted. A crossover design with the same sample was used, which is rarely reported in the literature. Furthermore, to verify the equivalence of online and paper-and-pencil formats, a multi-group analysis was used and the invariance between models was evaluated. This method has only been used in a few studies.^{1,2,10} It should also be noted that the American Psychological Association¹⁸ states that descriptive statistics, reliability, and construct validity are critical aspects when assessing psychometric scales. These components were sufficiently evaluated in this study. Different scale validation studies have been conducted to present evidence of validity and reliability, in addition to calling to the attention of readers/researchers the importance of these studies to support improving the quality of information when using psychometric scales.^{23–25} Instruments may only be used after the psychometric properties (factor structure, validity, and reliability) have been confirmed. However, if the administration method changes, it is necessary to estimate the stability of the factor structure of the instrument, since the method of administration can influence measurement of the construct.

Although the factor structure of the WCS was more parsimonious for the paper-and-pencil format, the results indicate invariance between the paper-and-pencil and online formats. This supports the idea that the measurement and organization of the construct followed the same structure for both administration types. It should be noted, however, that the variance explained in the online format was significantly higher than that in the paper-and-pencil format (i.e., improved construct measurement when there was no interaction with the researcher, which may point to a decreased social desirability effect). In this context, researchers have the option of using both WCS administration formats, but must consider the advantages and disadvantages of each with respect to their research protocol before making a choice.

Comparing the results presented in this study with the literature is difficult. No studies evaluate all of the psychometric properties of the WCS, nor compare them when the scale is administered in different formats. However, it can be stated that the WCS showed adequate validity and reliability in both administration formats, and equivalence/stability was observed for answers given in both formats.

Acknowledgments

The authors would like to thank the São Paulo Research Foundation (FAPESP) for research assistance granted grant #2010/18279-3, São Paulo Research Foundation (FAPESP) and Raquel Velez Oliveira for the translation and revision of the paper.

Author Disclosure Statement

No competing financial interests exist.

References

1. Campos JADB, Zucoloto ML, Bonafé FSS, et al. Reliability and validity of self-reported burnout in college students: a cross randomized comparison of paper-and-pencil vs. online administration. *Computers in Human Behavior* 2011; 27:1875–1883.
2. Carlbring P, Brunt S, Bohman S, et al. Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior* 2007; 23:1421–1434.
3. Nosek BA, Banaji MR, Greenwald AG. E-research: ethics, security, design, and control in psychological research on the Internet. *Journal of Social Issues* 2002; 58:161–176.
4. Evans DC, Garcia DJ, Garcia DM, et al. In the privacy of their own homes: using the Internet to assess racial bias. *Personality & Social Psychology Bulletin* 2003; 29:273–284.
5. Weigold A, Weigold IK, Russell EJ. Examination of the equivalence of self-report survey-based paper-and-pencil and internet data collection methods. *Psychological Methods* 2013; 18:53–70.
6. Yetter G, Capaccioli K. Differences in responses to Web and paper surveys among school professionals. *Behavior Research Methods* 2010; 42:266–272.
7. Riva G, Teruzzi T, Anolli L. The use of the Internet in psychological research: comparison of online and offline questionnaires. *Cyberpsychology & Behavior* 2003; 6:73–80.
8. Brock RL, Barry RA, Lawrence E, et al. Internet administration of paper-and-pencil questionnaires used in couple research: assessing psychometric equivalence. *Assessment* 2012; 19:226–242.
9. Bates SC, Cox JM. The impact of computer versus paper-pencil survey, and individual versus group administration, on self-reports of sensitive behaviors. *Computers in Human Behavior* 2008; 24:903–916.
10. Naus MJ, Philipp LM, Samsi M. From paper to pixels: a comparison of paper and computer formats in psychological assessment. *Computers in Human Behavior* 2009; 25:1–7.
11. Thorén ES, Andersson G, Lunner T. The use of research questionnaires with hearing impaired adults: online vs. paper-and-pencil administration. *BMC Ear, Nose & Throat Disorders* 2012; 12:12–17.
12. Hirsch O, Hauschild F, Schmidt MH, et al. Comparison of Web-based and paper-based administration of ADHD questionnaires for adults. *Journal of Medical Internet Research* 2013; 15:e47.
13. Hedman E, Ljotsson B, Rück C, et al. Internet administration of self-report measures commonly used in research on social anxiety disorder: a psychometric evaluation. *Computers in Human Behavior* 2010; 26:736–740.
14. Dolnicar S, Laesser C, Matus K. Online versus paper: format effects in tourism surveys. *Journal of Travel Research* 2009; 47:295–316.
15. Gleeson K, Frith H. (De)constructing body image. *Journal of Health Psychology* 2006; 11:79–90.
16. Killen JD, Taylor CB, Hayward C, et al. Pursuit of thinness and onset of eating disorder symptoms in a community sample of adolescent girls: a three-year prospective analysis. *The International Journal of Eating Disorders* 1994; 16:227–238.
17. Luce KH, Winzelberg AJ, Das S, et al. Reliability of self-report: paper versus online administration. *Computers in Human Behavior* 2007; 23:1384–1389.
18. Guyatt GH, Norman GR, Junipera EF, et al. A critical look at transition ratings. *Journal of Clinical Epidemiology* 2002; 55:900–908.
19. Streiner DL, Norma GR. (2008) *Health measurement scales: a practical guide to their development and rise*. New York: Oxford University.

20. Maroco J. (2014) *Análise de equações estruturais*. Lisbon: Report Number.
21. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 1981; 18:39–50.
22. Maroco J, Garcia-Marques T. Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas? *Laboratório de Psicologia* 2006; 4:65–90.
23. Guertler D, Broda A, Bischof A, et al. Factor structure of the Compulsive Internet Use scale. *Cyberpsychology, Behavior, & Social Networking* 2014; 17:46–51.
24. Barke A, Nyenhuis N, Kroner-Herwig B. The German version of the Generalized Pathological Internet Use Scale 2: a validation study. *Cyberpsychology, Behavior, & Social Networking* 2014; 17:474–482.
25. Wartberg L, Petersen KU, Kammerl R, et al. Psychometric validation of a German version of the compulsive Internet use scale. *Cyberpsychology, Behavior, & Social Networking* 2014; 17:99–103.

Address correspondence to:

*Dr. Juliana Alvares Duarte Bonini Campos
Faculdade de Ciências Farmacêuticas de Araraquara
Universidade Estadual Paulista
Rodovia Araraquara-Jaú, km 01
Araraquara
São Paulo 14801-901
Brazil*

E-mail: jucampos@fcfar.unesp.br