Internal structure of Beck Hopelessness Scale: An analysis of method effects using the CT-C(M–1) model

Pablo Ezequiel Flores-Kanter[1], Ronald Toro[2], Jesús M. Alvarado[3]

[1] National Scientific and Technical Research Council (CONICET); Life Sciences Department, Sleep and Memory Laboratory, Buenos Aires Institute of Technology (ITBA), Buenos Aires, Argentina; Centre for Bioethics, Catholic University of Cordoba (UCC), Cordoba, Argentina. https://orcid.org/0000-0002-6712-779X

[2]Psychology Department, Catholic University of Colombia, Bogotá, Colombia. https://orcid.org/0000-0001-6061-3499

[3]Psychobiology & Behavioral Sciences Methods Department, Faculty of Psychology, Complutense University of Madrid, Madrid, Spain. https://orcid.org/0000-0003-4780-0147

Corresponding author:

Pablo Ezequiel Flores-Kanter

Tel: +5493513849201

Email: pablo.floreskanter@conicet.gov.ar; pflores@itba.edu.ar

Word count: 7918

Internal structure of Beck Hopelessness Scale: An analysis of method effects using the CT-C(M–1) model

The construct validity in relation to the dimensionality or factor structure of the Beck Hopelessness Scale (BHS) has long been debated in psychometrics. Irrelevant variance due to item wording (method effects) can distort the factor structure, and recent studies have examined the method factor's role in the factor structure of the BHS. However, the models used to control the method effects have severe limitations, and new models are needed. One such model is the correlated trait-correlated method minus one (CT-C(M-1)), which is a powerful approach that gives the trait factor an unambiguous meaning and prevents the anomalous results associated with fully symmetrical bifactor modeling. The present work compares the fit and factor structure of the CT-C(M-1) model to bifactor models proposed in previous literature and evaluates the convergent validity of the CT-C(M-1) model and its discriminatory capacity by taking suicidal ideation as the criterion variable. This study used a large and heterogeneous open mode online sample of Argentinian people (N = 2,164). The results indicated that the CT-C(M-1) model with positive words as referenced items achieves the most adequate factor structure. The factorial scores derived from this model demonstrate good predictive and discriminating capabilities.

Introduction

The Beck Hopelessness Scale (BHS) is a self-assessment questionnaire designed to measure negative attitudes toward the future (Beck et al., 1974). The BHS has become one of the most popular measurements of the hopelessness construct in international studies (Kliem et al., 2018). One common practice in previous BHS research is, after recoding negatively worded items, to combine the number of endorsed items into a sum-score (Baryshnikov et al., 2020; Brown et al., 2000; Flores-Kanter et al., 2019; Granö et al., 2016; Serafini et al., 2020; Tsujii et al., 2020).

The use of the total BHS score assumes that the psychometric evidence favors a one-dimensional model. However, the factor structure of the scale has been extensively debated, and different internal structure solutions have been suggested for the BHS (Boduszek & Dhingra, 2016; Kocalevent et al., 2016). The three-factor model (Beck et al., 1974) (i.e., feelings about the future, loss of motivation, future expectations) and the one-dimensional model (Aish & Wasserman, 2001) are commonly discussed in the literature. More recently, the influence of method factors (MFs) has been pointed out (i.e., optimism and pessimism in item wording), as discussed by Boduszek and Dhingra (2016), Innamorati et al. (2013), Kliem et al. (2018), and Szabo et al. (2016)). In their research, bifactor models were used to consider specific factors concerning the negative and positive wording of items, as well as a general orthogonal factor of hopelessness.

The inconsistencies between different factor solutions and the lack of clarity of items that make up the factors impede the clear interpretation of BHS scores. This issue is relevant to the applied field because the common use of the raw or total BHS score in predictive studies may be based on a non-optimal factor solution. More importantly, the use of these scores may increase the prevalence of systematic variance that is irrelevant to the measurement of the hopelessness construct. These aspects produce noise in the metrics obtained, which can lead to biased estimates and a loss of predictive capability.

Several reasons may explain the significant divergence in the available evidence on the internal structure of the BHS. Many of these have been clearly identified by Boduszek and Dhingra (2016) and Szabo et al. (2016) and include the heterogeneity in the included samples (clinical vs. general population) and the application of non-conforming estimation methods for the BHS binary response format items that stand out. Beyond these possibilities, another relevant factor to consider is the format in which the items are written or the method effect of the item wording.

Incorporation of negative items in the development of psychometric instruments

In many psychometric instruments, the incorporation of negative items is not due to the generation of a differential factor for a given construct. On the contrary, incorporation of reagents controls response bias, namely acquiescence bias (Savalei & Falk, 2014). However, the incorporation of reversed or negative items in psychometric scales has been widely discussed (see Zhang & Savalei, 2015).

In BHS factor studies reporting more than one factor, the resulting factor structure mainly separated the positively worded items from the negatively worded ones. Positively worded items tended to load on one factor, while negatively worded items tended toward multiple factors. Thus, it is possible to hypothesize that the generation of multi-factor structures in the BHS is a methodological artifact that results from the incorporation of negative items and does not reflect a substantive factor. Identifying the wording of items as dependent factors reflects the tendency of participants to respond in distinct ways to positively and negatively worded items, regardless of the intended content. This tendency yields a systematic variance that is irrelevant to the context of the study (DiStefano & Motl, 2006) and causes biases in the covariance structure of the data (Tomas & Oliver, 1999; Zhang & Savalei, 2015).

Understanding this response bias may help reconcile the mixed reported pattern of the internal structure of the BHS and allow for more conclusive results among applied researchers. The method effect was recently considered in studies of the BHS factor structure (Boduszek & Dhingra, 2016; Innamorati et al., 2013; Kliem et al., 2018; Szabo et al., 2016). In these studies, bifactor models were used to examine the specific factors involved in negative and positive item wordings and a general orthogonal factor of hopelessness. The findings of such studies show an adequate fit for these measurement models. However, these CT-calculometry (CT-CM)-based approaches (i.e., fully symmetrical bifactor modeling) have serious limitations, and their use for method effect control is questionable.

Geiser et al. (2008) have identified two major disadvantages of the CT-CM model: i) it is prone to identification problems, and ii) it often produces solutions with out-of-range parameter estimates. Moreover, the interpretation of the trait and method factors in the CT-CM model can be ambiguous (Heinrich et al., 2018). The correlated trait-correlated method minus one (CT-C (M-1); Eid, 2000) is a recently proposed approach aimed at overcoming these limitations.

The CT-C (M-1) is an approach that has yet to be considered in the study of the BHS's internal structure. The CT-C(M-1) model is a particular variant of the CT-CM, but the number of MFs is one fewer than the number of methods included. The M-1 measure model is a powerful approach that gives the trait factor an unambiguous meaning and prevents the anomalous results associated with fully symmetrical bifactor modeling (Eid, 2020; Heinrich et al., 2018). The advantage of this model over the traditional CT-CM models is that its trait, method, and error components are uncorrelated. This feature allows for the decomposing of variances and covariances of the nonreference measures into variance components, due to the influences of the trait, method, and measurement errors. Thus, M-1 models allow the separation of substantive content from the mechanism or method used to gain the responses (e.g., item wordings; Geiser et al., 2008). In this confirmatory factor analysis (CFA) approach, the first-order CFA model is taken as a starting point. One of the first-order factors is selected as a reference domain that ultimately defines the trait factor (i.e., it reflects the trait assessed by the reference domain and therefore has an unambiguous meaning). The MFs reflect the parts of a domain that the reference domain cannot predict (i.e., residual factors). In the case of the BHS, the items would be determined by a trait factor of pessimism–hopelessness and one MF (i.e., negative or reversed items: the "optimism" or positively worded items in the BHS). Thus, in this case, the M-1 model eliminates one of the suggested BHS MFs (see Figure 1). Negatively worded items (i.e., positive or direct items) are selected as the reference method or trait factor so that the model contains a nonreference (method) factor for the positively worded items.

Given the advantages of applying the M-1 model, we hypothesize that this model will produce better fit indexes than those observed for previous measurement models tested on the BHS (H1).

[Figure 1]

Need for criterion-related validity

The CT-C(M-1) model must present evidence of criterion validity. For this purpose, a semi-partial correlation approach by means of structural equation modelling (SEM) is suggested (Geiser et al., 2008). SEM can be considered a combination of factor analysis models (i.e., measurement models) and regression models (i.e., structural models) (Yuan & Bentler, 2007); the latter can be used to statistically control for variables (i.e., partialling out). In other words, it allows for decomposing the variances and covariances of the measures into variance components due to trait, method, and measurement error influences. As indicated by Geiser et al. (2008), the lack of criterion validity is indicated by a large amount of method-specific variance after statistically controlling for the trait factor. In the present research, we hypothesized that the trait factor that represents pessimism for its positive items strongly explains suicide ideation (SI) (H2). By contrast, the MFs would only have a small or statistically insignificant effect beyond the effect of the trait factor (H3).

We also proposed validating the M-1 model by applying the receiver operating characteristic (ROC) curve procedures with high SI and non-SI groups. This was particularly useful in clarifying the discriminatory capacity of the factor scores obtained. Moreover, it allowed us to appreciate the relative advantage of using the factor score of the M-1 model over the raw total BHS score (i.e., BHS sum-score) in terms of discriminatory capacity. In the ROC analysis, the area under the curve (AUC) measures the performance of a classifier; a higher AUC value indicates a better classification. The AUC is particularly useful when diagnostic tests are being compared (Habibzadeh et al., 2016). We compared the discriminating ability of the M-1 factor score and BHS sum-score. Recent studies have verified that the AUC obtained by using the BHS sum-score for discrimination against SI groups varies within the range of .735 to .798 (Baryshnikov et al., 2020; Granö et al., 2016). We hypothesized that the discriminatory capacity of the M-1 factor score would be higher than observed in the previous literature, resulting in a statistically significant difference in the AUC values (H4). Given the common application of the BHS sum-score, obtaining evidence to confirm H4 allowed us to determine the practical significance of our findings.

Objectives

The main objective of this research was to verify the internal structure of the BHS in a large heterogeneous sample (in terms of gender, age, and clinical status). Unlike previous research, we applied a CT-C(M-1) model as a central objective. We also verified the convergent validity and the relative advantage of using the scores derived from the M-1 model in the prediction of SI.

Materials and methods

Participants

Participants were 2,164 Argentinians obtained through an open mode online sample method (The International Test Commission, 2006). This data collection methodology has proven to be equivalent to traditional forms of collection (i.e., face to face; Weigold et al., 2013), producing equal means, internal consistencies, intercorrelations, response rates, and comfort level when completing questionnaires. For this observational, cross-sectional study, data were collected using an online survey format to gather information through the Google Forms platform and were delivered by Facebook social media. The data were collected in November and December 2018. The socio-demographics of the participants are shown in Table 1.

[Table 1]

All participants were adequately informed of the research objectives, the anonymity of their responses, and their voluntary participation. Likewise, it was clarified that participation would not cause any harm and that they could leave the study whenever they wished. International ethical guidelines for studies with human beings were considered (American Psychological Association, 2017). In this study, no specific incentive was used for participation in the study. In the case of minors under 18 years of age, prior parental consent was additionally requested. The ethics council of the Centre for Bioethics of the Catholic University of Cordoba previously approved the research protocol following APA ethical guidelines.

Measures

Beck Hopelessness Scale (BHS)

For this study, the adapted BHS version for the Argentinian population was used (Mikulic et al., 2009). This scale comprises 20 items with a dichotomous reply format (i.e., true or false) and is used to evaluate the respondent's negative expectations for the future. The instrument shows an adequate level of internal consistency (α = .78). In the present sample, an adequate Cronbach's alpha coefficient for the dimension was observed (α = .88).

Inventory of Suicide Orientation-30 (ISO-30)

The version of the Inventory of Suicide Orientation-30 (ISO-30) validated by Fernandez-Liporace and Casullo (2006) in Argentina was used for this study. The instrument measures the respondent's level of agreement with certain statements using a four-point Likert scale (with responses ranging from 1, "I strongly disagree" to 4, "I strongly agree"). From the inventory, only the questions that were related to the dimension of SI were included (e.g., "In order to stop things from getting worse, I believe suicide is the solution"). Within the ISO-30 scale, the suicidal ideation factor is the one that has shown the highest consistency and evidence of measurement validity in the literature on the internal structure of the scale (Vecco et al., 2021). In the present sample, an adequate Cronbach's alpha coefficient for the dimension was observed (α= .93).

Statistical analyses

First, CFA was applied. We tested the four most commonly discussed models in the BHS literature (see Boduszek & Dhingra, 2016): (1) the original three-factor model by Beck et al. (1974; Model 1), comprising "feelings about the future" (items: 1, 5, 6, 13, 15, 19), "loss of motivation" (items: 2, 3, 9, 11, 12, 16, 17, 20), and "future expectations" (items: 4, 7, 8, 10, 14, 18); (2) the one-dimensional model (all items loading on one overall-general factor; Model 2); (3) the bifactor model as suggested by Szabó et al. (2016; Model 3), with one general factor (all items) and two MFs (pessimistically worded items: 2, 4, 7, 9, 11, 12, 14, 16, 17, 18, 20; and optimistically worded items: 1, 3, 5, 6, 8, 10, 13, 15, 19); and (4) the multitrait-multimethod model suggested by Boduszek and Dhingra (2016; Model 4), with three correlated trait factors (as suggested by Beck et al., 1974) and two correlated MFs. Finally, we added (5) the M-1 model proposed in the present research (Model 5).

Given the categorical nature of the observable variables in the BHS, the models were analyzed using the weighted least square mean and variance (WLSMV) estimator (Flora & Curran, 2004; Rhemtulla et al., 2012). The overall fit of the models was evaluated using the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the Tucker-Lewis index (TLI). RMSEA values below .08 and .05 are considered indicative of reasonable and proper fit, respectively. CFI and TLI values above .90 and .95 reflect an acceptable and good adjustment, respectively (Hu & Bentler, 1999; Marsh et al., 2004). To compare the two given nested models, a T3 second-order correction for the chi-square difference testing method was used (Asparouhov

& Muthén, 2006). This method is appropriate when the WLSMV estimator is used. In addition, the quality of model fit was assessed based on the adequacy and plausibility of the estimated coefficients (Geiser et al., 2008; Heinrich et al., 2018).

The structural model proposed for the prediction of the SI was estimated using the same approach in the measurement model analysis (i.e., CFA). The criterion validity of the reference trait factor is supported if the variance explained by the MFs is smaller than the variance accounted for by the trait factor (Geiser et al., 2008). This was assessed from the standardized regression weights, the estimated correlation coefficients, and the statistical significance obtained for these parameters. Mplus (version 8) was used for CFA and SEM statistical analyses.

Finally, the ROC curves were analyzed, for which the AUC and their respective confidence intervals were calculated (Gönen, 2007). A ROC curve diagram indicates the performance of a binary classification method in terms of its sensitivity (i.e., correctly classified positive observations) and specificity (i.e., correctly classified negative observations). The ROC plot also displays the performance of a binary classification method with discrete ordinal output (Robin et al., 2011). In this research, the ranges of values obtained from the factor scores of the M-1 model were used as a starting point, verifying the performance of a SI group classification. Furthermore, the ROC curve and AUC analysis were based on an extreme case-control design. This type of design is applied to compare groups that lie at both extremes of a specific variable range. (Granat et al., 2017; Zhu et al., 2013). In the present research, the high-SI and no-SI groups were identified based on the scores obtained from the ISO-30 SI subscale (possible range values = 4–16). A total score of 4 (i.e., the participant marked "strongly disagree" on all items) was considered to indicate an absence of SI, while a score of 16 (i.e., the participant marked "strongly agree" on all items) indicated a high SI. Individuals scoring in the middle range were not included in the analyses (see Table 1). In the ROC analysis, the AUC measures the performance of a classifier, and a higher AUC value means a better classification. The AUC can be considered an index of the discriminating ability of a test, and an AUC of 0.5 is equivalent to tossing a coin (i.e., an uninformative test). Statistical analysis was performed using the pROC package (Robin et al., 2011) in R (R Core Team, 2016).

Results

Internal structure of BHS: Comparison of measurement models

As shown in Table 2, Models 3 and 4 fit better than Model 5, and Model 5 fits better than Models 1 and 2.

[Table 2]

While Models 3 and 4 fit better than Model 5, they evidenced anomalous results (i.e., negative variances of specific factors and non-significant and/or incoherent negative loadings in the specific factors; Eid, 2020; Flores-Kanter et al., 2018). In other words, Models 3 and 4 are suboptimal (Heinrich et al., 2018). Based on these results, it can be concluded that the best fit is achieved by Model 5.

Table 3 summarizes the standardized factor loadings corresponding to the M-1 model with positive words as referenced items (Model 5). Except for item 13, all items load strongly in the trait factor (> .50) and in the expected direction (all $p < .001$).

[Table 3]

Criterion-related validity

Testing the structural model

In support of hypothesis (H1), the trait factor presents a larger ($\beta = .754$, $p < .001$) and more significant effect than the low-magnitude effect observed in the case of the MF ($\beta = .201$, $p = .008$). However, the MF has a statistically significant effect on SI. Although the effect is low, this result may suggest the presence of a certain trait and not just a residual method effect (Geiser et al., 2008). Based on the previous CFA findings, we hypothesized that this effect could be mediated mainly by item 13. The analyses—excluding item 13—were repeated to corroborate this. The results indicated that the trait factor has a strong effect on SI (β = .753, p < .001), while the method effect decreases in magnitude, not reaching statistical significance (β = .152, p = .052; see Figure 2).

[Figure 2]

Since relating external variables directly (i.e., direct effects) to general factors and specific factors can lead to biased parameter estimates (Koch et al., 2017), we repeated the analysis with the same SEM model, but we estimated the correlation between the trait factor and MFs with SI (see Gomez & Watson, 2017). In this case, the correlation between the trait factor and the SI remained high, while the correlation between the MF and SI had a marginal statistical significance and showed a very small or null effect. This occurred when item 13 was included (SI with MF r = .088, p = .009; SI with trait factor r = .753, p < .001) and when it was excluded (SI with MF r = .068, p = .047; SI with trait factor r = .753, p < .001).

ROC curve procedure

For comparison testing, we used the comparison based on the AUC and the bootstrap percentile method for paired or correlated ROC curves (Robin et al., 2011). The ranges of values obtained from the factor scores derived from the M-1 model and the BHS sum-score were used as a starting point, verifying the performance of an SI group classification. For comparative purposes, we included the total BHS sum-score due to the widespread use of this procedure in applied research. Figure 3 shows the ROC curve obtained for the trait factor score (solid line) and the BHS sum-score (dashed line). The results showed that the discrimination capacity of the factorial score obtained from the CT-C(M-1) model (AUC = .928) was significantly higher (p < .001) than that observed for the BHS sum-score (AUC = .794).

[Figure 3]

Discussion

The study's main objective was to apply a CT-C(M-1) model to analyze the internal structure of the BHS. We proposed to verify the convergent validity and relative advantage of using the factor scores derived from the M-1 model in the prediction of SI. We first hypothesized that this model would obtain better adjustment rates than those observed for previous models (H1). Furthermore, we hypothesized that the trait factor representing hopelessness for its positive items strongly explains SI (H2). By contrast, we proposed that the MF will only have a small or statistically insignificant effect beyond the effect of the trait factor (H3). Finally, we hypothesized that the discriminatory capacity of the M-1 factor score would be higher than that observed in the previous literature, obtaining a statistically significant difference in the AUC values of the M-1 factor score compared to the BHS sum-score (H4).

Our analysis showed evidence in support of H1. The better fit obtained by the M-1 model is consistent with the initial theoretical approach on which Beck et al. (1974) based the elaboration of the scale. They pointed out that "a person's hopelessness can be objectified by defining it in terms of a system of negative expectancies concerning himself and his future life" (p. 861). In this sense, we consider that the controversies evidenced so far regarding the factor structure or construct validity of the BHS are mainly due to methodological artifacts. In this case, the trait factor could be termed *pessimism measured by negatively worded items* (i.e., positive items; see Geiser et al., 2008).

In the present study, it was verified that item 13 is the only MF indicator that could account for a higher-than-expected explained variance. In much of the earlier research, this item showed problematic performance (Kliem et al., 2018). A more detailed analysis of its content shows that it is the only item that refers to the respondent's current situation and the possibility of positive change in the future. It can be hypothesized that item 13 refers to a cognitive-emotional regulation strategy of an adaptive nature, possibly associated with the refocusing or positive reinterpretation of adverse events (Flores-Kanter & Medrano, 2020). In favor of this possibility, the structural model showed that the MF's effect does not achieve statistical significance when item 13 is subtracted. In this sense, the content validity of item 13 in the context of the BHS is debatable, and it is suggested that in future applications of the scale the results should be verified by considering the inclusion/exclusion of this item.

With respect to H2 and H3, when we applied the structural model based on the M-1 measurement, the presence of a strong effect of the trait factor on the SI was seen, compared to the MF's low and statistically insignificant impact on SI (Figure 2). This result was confirmed by verifying the correlations between the trait factor and MF with the SI. Therefore, this suggests that a trait factor (i.e., pessimism measured by negatively worded items) exerted the greatest amount of variance on a criterion variable of interest, and no substantial variance that can be explained beyond this trait factor was evident (Geiser et al., 2008).

We also obtained evidence in support of H4. The results of the present work show that the discrimination capacity of people with high SI and no SI obtained from the BHS sum-score is similar to that obtained in previous studies (see Baryshnikov et al., 2020 and Granö et al., 2016), while the discrimination capacity of the factorial score of the trait factor obtained from the CT-C(M-1) model is significantly higher. These are very important results since the BHS sum-score is the major statistic interpreted in everyday practice (Kliem et al., 2018). Based on these results, we support can conclude in favor of using the factorial score derived from the CT-C(M-1), as opposed to the widespread practice of calculating the BHS sum-score for discriminatory/predictive purposes.

At this point, it is essential to make some observations regarding the use of bifactor models in psychology and related sciences. Bifactor models are increasingly used for analyses in different research areas in psychology and related sciences as one way to control the method effect (Bonify et al., 2016). In these scenarios, the objective of the analysis is to verify the reliable total score after controlling for some MFs, such as reversed or negative items (Eid et al., 2016). There is strong evidence of the method effect in studies that use other scales, one example being the Rosenberg Self-Esteem Scale. Likewise, in other scales similar to the BHS (see Maydeu-Olivares & Coffman, 2006), the same problem has been observed due to the incorporation of reversed or negative items that show optimism or a positive vision of the future. It is possible to verify the presence of the method effect, and correct modeling has enabled an improvement in the use of the scale and its factor scores (Salerno et al., 2017). In line with these arguments, the results obtained in the present research evidenced the need for and importance of considering factor scores derived from well-fitting CFA models.

In this field, it can be difficult for researchers to choose among several methods to model this type of method-related effect. This issue may explain the frequent misuse of bifactor models (see Bonifay et al., 2016). Therefore, it is important to highlight some evidence-based suggestions to guide good practice (Flores-Kanter et al., 2018). In the case of modeling the method effect on scales similar to the self-esteem or hopelessness-pessimism scales, the CT-C(M-1) models have proven to be advantageous over other modeling possibilities (e.g., random intercepts, symmetries CT-CM; see Geiser et al., 2008; Burns et al., 2020). In this research, the CT-CM and MTMM models showed poor specification errors that prevented their identification or the estimation of the corresponding parameters, presenting an inadequate fit. Finally, it is important to emphasize specific points related to the correct reading of the results obtained in this type of approach. In the case of the M-1 model, the key aspect is the differentiation between (a) a spurious method effect or (b) the presence of a differential feature (Marsh et al., 2010). The present results regarding the BHS allow us to conclude in support of (a).

This study has some limitations that should be noted. First, participants were asked to self-report; thus, the findings should be replicated using other measures (Bonifay et al., 2016; Clark & Watson, 2019). For example, regarding criterion-related validity, other behavioral measures (e.g., suicide attempts) could be incorporated. Second, this study utilized a cross-sectional observational design; however, the implementation of longitudinal and repetitive measures (e.g., ecological momentary assessment; see Kim et al., 2019) would enable future researchers to make more precise inferences regarding the state-trait hopelessness construct as well as mutual relationships with relevant external variables. Finally, future research should extend these findings to other populations (e.g., clinical samples; Boduszek & Dhingra, 2016).

Conclusions

This study provides evidence in support of the CT-C(M-1) model for modeling the BHS internal structure, and the trait factor could, in this case, be termed *pessimism measured by negatively worded items.* The resulting trait factor score showed convergent validity by presenting a substantial effect on SI indicators. Further, the trait factor score evidenced better and optimal discrimination between people with high SI and no SI. In concordance with Wetzel and Roberts (2020) and Flores-Kanter (2017), we hope that the applied field does not undervalue good measurement and evidence-based research practices or ignore the measurement research that has been done.

Funding details: This research was possible due to funding from the research group Cognitive Psychology: Measurement and Modeling of Processes at Complutense University of Madrid.

Declaration of interest statement: The authors have no competing interests to declare.

Data availability statement: The data that support the findings of this study are openly available in “Mendeley Data” at https://data.mendeley.com/datasets/9mjg54kcwv/5, reference number V5.

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness Depression: Theory-Based subtype of depression. *Psychological Review, 96*(2), 358-372. doi:10.1037/0033-295X.96.2.358

Aloba, O., Awe, O., Adelola, A., Olatunji, P., & Aloba, T. (2018). Psychometric Adaptation of the Beck Hopelessness Scale as a Self-Rated Suicide Risk Screening Instrument Among Nigerian University Students. *Journal of the American Psychiatric Nurses Association, 24*(5), 433–443. https://doi.org/10.1177/1078390318762054

American Psychological Association. (2017). *Ethical Principles of Psychologists and Code of Conduct*. Washington, EEUU: Author. Retrieved from http://www.apa.org/ethics

American Psychological Association (APA) (June 18, 2018). *APA calls suicide prevention a public health priority* [Press Release]. Retrieved from APA website https://www.apa.org/news/press/releases/2018/06/suicide-prevention.aspx

Asparouhov, T & Muthen, B. (2006). *Robust Chi Square Difference Testing with Mean and Variance Adjusted Test Statistics*. Mplus Web Notes: N° 10. Retrieved from http://statmodel.com/download/webnotes/webnote10.pdf

Barredo, J., Aiken, E., van't Wout-Frank, M., Greenberg, B. D., Carpenter, L. L., & Philip, N. S. (2019). Neuroimaging correlates of suicidality in decision-making circuits in posttraumatic stress disorder. *Frontiers in Psychiatry, 10*. http://dx.doi.org/10.3389/fpsyt.ȁ9.00044

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology, 42*(6), 861–865. http://dx.doi.org/10.1037/h0037562

Bonifay, W., Lane, S. P., & Reise, S. P. (2016). Three concerns with applying a Bifactor Model as a structure of psychopathology. *Clinical Psychological Science, 5*(1), 184–186. http://dx.doi.org/10.1177/2167702616657069

Boduszek, D., & Dhingra, K. (2016). Construct validity of the Beck Hopelessness Scale (BHS) among university students: A multitrait–multimethod approach. *Psychological Assessment, 28*(10), 1325–1330. http://dx.doi.org/10.1037/pas0000245

Burns, G. L., Geiser, C., Servera, M., Becker, S. P., & Beauchaine, T. P. (2019). Application of the bifactor S – 1 model to multisource ratings of ADHD/ODD symptoms: An appropriate bifactor model for symptom ratings. *Journal of Abnormal Child Psychology, 48*(7), 881–894. https://doi.org/10.1007/s10802-020-00629-4

Chang, E. C., D'Zurilla, T. J., & Maydeu-Olivares, A. (1994). Assessing the dimensionality of optimism and pessimism using a multimeasure approach. *Cognitive Therapy and Research, 18*(2), 143–160. https://doi.org/10.1007/bf02357221

Clark, D. A., Beck, A. T., & Alford, B. A. (1999). *Scientific foundations of cognitive therapy and therapy of depression*. New York: John Wiley and Sons.

Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, 31(12), 1412–1427. doi:10.1037/pas0000626

DiStefano, C., & Motl, R. W. (2006). Further Investigating Method Effects Associated With Negatively Worded Items on Self-Report Surveys. *Structural Equation Modeling: A Multidisciplinary Journal, 13*(3), 440–464. https://doi.org/10.1207/s15328007sem1303_6

Dyce, A. J. (1996). Factor structure of the Beck Hopelessness Scale. *Journal of Clinical Psychology, 25*, 555–556.

Eid, M. (2020). Multi-Faceted Constructs in Abnormal Psychology: Implications of the Bifactor S - 1 Model for Individual Clinical Assessment. *Journal of Abnormal Child Psychology, 48*(7), 895–900. https://doi.org/10.1007/s10802-020-00624-9

Eid, M., Geiser, C., & Koch, T. (2016). Measuring Method Effects. *Current Directions in Psychological Science, 25*(4), 275–280. https://doi.org/10.1177/0963721416649624

Fernández-Liporace, M., & Casullo, M. M. (2006). Validación factor de una escala para evaluar riesgo suicida [Factor validity of a scale to assess suicidal risk]. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica, 21*(1), 9–22.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological methods*, *9*(4), 466-491. https://doi.org/10.1037/1082-989X.9.4.466

Flores-Kanter, P. E. (2020). Commentary: How we know what not to think. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.00306

Flores Kanter, P. E. (2017). El lugar de la psicología en las investigaciones empíricas del suicidio en argentina: Un estudio Bibliométrico [State of psychological empirical research on suicide in argentine: A bibliometric study]. *Interdisciplinaria, 34*(1). https://doi.org/10.16888/interd.2017.34.1.2

Flores-Kanter, P. E., Dominguez-Lara, S., Trógolo, M. A., & Medrano, L. A. (2018). Best practices in the use of bifactor models: Conceptual grounds, fit indices and complementary indicators. *Revista Evaluar, 18*(3), 44-48. https://doi.org/10.35670/1667-4545.v18.n3.22221

Flores-Kanter, P. E., García-Batista, Z. E., Moretti, L. S., & Medrano, L. A. (2019). Towards an Explanatory Model of Suicidal Ideation: The Effects of Cognitive Emotional Regulation Strategies, Affectivity and Hopelessness. *The Spanish Journal of Psychology, 22*. https://doi.org/10.1017/sjp.2019.45

Flores Kanter, P. E. & Medrano, L. A. (2020). Commentary: Putting 'Emotional Intelligences' in Their Place: Introducing the Integrated Model of Affect-Related Individual Differences. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.00574

Flores Kanter, P. E., Medrano, L. A. & Manoiloff, M. L. (2015). Estados de ánimo y juicios de autoconcepto en universitarios: Análisis desde un abordaje basado en Redes Semánticas Naturales [Moods and self-concept judgments in university students: An analysis approach based on natural semantic networks]. *Interamerican Journal of Psychology*, 48(3), 291-296.

Geiser, C., Eid, M., & Nussbeck, F. W. (2008). On the meaning of the latent variables in the CT-C(M-1) model: A comment on Maydeu-Olivares and Coffman (2006). *Psychological Methods, 13*(1), 49–57. https://doi.org/10.1037/1082-989x.13.2.171

Gomez, R., & Watson, S. D. (2017). Confirmatory Factor Analysis of the Combined Social Phobia Scale and Social Interaction Anxiety Scale: Support for a Bifactor Model. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00070

Gönen, M. (2007). *Analyzing Receiver Operating Characteristic Curves with SAS*. New York, EEUU: SAS Publishing.

Granat, A., Gadassi, R., Gilboa-Schechtman, E., & Feldman, R. (2017). Maternal depression and anxiety, social synchrony, and infant regulation of negative and positive emotions. *Emotion, 17*(1), 11–27. https://doi.org/10.1037/emo0000204

Granö, N., Oksanen, J., Kallionpää, S., & Roine, M. (2016). Specificity and sensitivity of the Beck Hopelessness Scale for suicidal ideation among adolescents entering early intervention service. *Nordic Journal of Psychiatry, 71*(1), 72–76. https://doi.org/10.1080/08039488.2016.1227370

Habibzadeh, F., Habibzadeh, P., & Yadollahie, M. (2016). On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochemia Medica,* 297–307. https://doi.org/10.11613/bm.2016.034

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit, & D. So¨rbom (Eds.), *Structural equation modeling: Present and future—A Festschrift in honor of Karl Jo¨reskog* (pp. 195-216). Lincolnwood, IL: Scientific Software International

Hanna, D., White, R., Lyons, K., McParland, M. J., Shannon, C., & Mulholland, C. (2011). The structure of the Beck Hopelessness Scale: A confirmatory factor analysis in UK students. *Personality and Individual Differences, 51*(1), 17–22. https://doi.org/10.1016/j.paid.2011.03.001

Heinrich, M., Zagorscak, P., Eid, M., & Knaevelsrud, C. (2018). Giving G a Meaning: An Application of the Bifactor-(S-1) Approach to Realize a More Symptom-Oriented Modeling of the Beck Depression Inventory–II. *Assessment*, 1-19. https://doi.org/10.1177/1073191118803738

Hill, R. D., Gallagher, D., Thompson, L. W., & Ishida, T. (1988). Hopelessness as a measure of suicidal intent in the depressed elderly. *Psychology and Aging, 3*(3), 230–232. https://doi.org/10.1037/0882-7974.3.3.230

Horwitz, A. G., Berona, J., Czyz, E. K., Yeguez, C. E., & King, C. A. (2016). Positive and Negative Expectations of Hopelessness as Longitudinal Predictors of Depression, Suicidal Ideation, and Suicidal Behavior in High-Risk Adolescents. *Suicide and Life-Threatening Behavior, 47*(2), 168-176. doi:10.1111/sltb.12273

Hu L., & Bentler P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Iliceto, P., & Fino, E. (2015). Beck Hopelessness Scale (BHS). *European Journal of Psychological Assessment, 31*(1), 31–37. http://dx.doi.org/10.1027/1015-5759/a000201

Innamorati, M., Lester, D., Balsamo, M., Erbuto, D., Ricci, F., Amore, M., . . . Pompili, M. (2013). Factor validity of the Beck Hopelessness Scale in Italian medical patients. *Journal of Psychopathology and Behavioral Assessment, 36*, 300-307. http://dx.doi.org/10.1007/s10862-013-9380-3

Joiner, T. E., Brown, J. S., & Wingate, L. R. (2005). The psychology and neurobiology of suicidal behavior. *Annual Review of Psychology, 56*, 287-314. doi:10.1146/annurev.psych.56.091103.070320

Kim, J., Marcusson-Clavertz, D., Yoshiuchi, K., & Smyth, J. M. (2019). Potential benefits of integrating ecological momentary assessment data into mHealth care systems. *BioPsychoSocial Medicine, 13*(1). https://doi.org/10.1186/s13030-019-0160-5

Kliem, S., Lohmann, A., Mößle, T., & Brähler, E. (2018). Psychometric properties and measurement invariance of the Beck hopelessness scale (BHS): results from a German representative population sample. *BMC Psychiatry, 18*(1). http://dx.doi.org/10.1186/s12888-018-1646-6

Kocalevent, R.-D., Finck, C., Pérez-Trujillo, M., Sautier, L., Zill, J., & Hinz, A. (2016). Standardization of the Beck Hopelessness Scale in the general population. *Journal of Mental Health, 26*(6), 516–522. doi:10.1080/09638237.2016.1244717

Koch, T., Kelava, A., & Eid, M. (2017). Analyzing different types of moderated method effects in confirmatory factor models for structurally different methods. *Structural Equation Modeling: A Multidisciplinary Journal, 25*(2), 179–200. https://doi.org/10.1080/10705511.2017.1373595

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320-341. https://doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. Psychological Assessment, 22(3), 471–491. https://doi.org/10.1037/a0019227.supp

Maydeu-Olivares, A., & Coffman, D. L. (2006). Random intercept item factor analysis. *Psychological Methods, 11*(4), 344–362. http://dx.doi.org/10.1037/1082-989x.11.4.344

Medrano, L. A., Flores-Kanter, E., Moretti, L., and Pereno, G. L. (2016). Effects of induction of positive and negative emotional states on academic self-efficacy beliefs in college students. *Psicolgía Educativa, 22*, 135–141. https://doi.org/10.1016/j.pse.2015.03.003

Mikulic I. M., Casullo G. L., Crespi M. C., & Marconi A. (2009). Escala de Desesperanza BHS (A. Beck, 1974): Estudio de las propiedades psicométricas y baremización de la Adaptación

Argentina [Beck Hopelessness Scale BHS (A. BECK, 1974): Psychometric study and standardization of Argentinian adaptation]. *Anuario de Investigaciones, 16*, 365–373.

Mystakidou, K., Parpa, E., Tsilikia, E., Pathiaki, M., Hatzipli, I., & Galanos, A. (2008). The experience of hopelessness in a population of Greek cancer patients receiving palliative care. *International Journal of Social Psychiatry, 54*(3), 262-271. http://dx.doi.org/10.1177/0020764008089857

Naghavi, M. (2019). Global, regional, and national burden of suicide mortality 1990 to 2016: systematic analysis for the Global Burden of Disease Study 2016. *BMJ, l94*. http://dx.doi.org/10.1136/bmj.l94

Nissim, R., Flora, D. B., Cribbie, R. A., Zimmermann, C., Gagliese, L., & Rodin, G. (2009). Factor structure of the Beck hopelessness scale in individuals with advanced cancer. *Psycho-Oncology, 19*(3), 255-263. http://dx.doi.org/10.1002/pon.1540

Pompili, M., & Tatarelli, R. (2007). The Hopelessness Scale: A factor analysis. *Psychological Reports, 100*(2), 375-378. http://dx.doi.org/10.2466/pr0.100.2.375-378

Quiñonez-Tapia, F., Luévano, T. E. M., & Castañeda-Camey, N. (2019). Análisis factor confirmatorio y propiedades psicométricas de la Escala de Desesperanza de Beck en estudiantes en contextos de pobreza en México [Confirmatory factor analysis and psychometric properties of the Beck Hopelessness Scale in students in contexts of poverty in Mexico]. R*evista de Psicopatología y Psicología Clínica, 24*(2), 59-70. https://doi.org/10.5944/rppc.24104

R Core Team (2016). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*, 667-696. http://dx.doi.org/10.1080/00273171.2012.715555

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*, 354-373. https://doi.org/10.1037/a0029315

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12(1). https://doi.org/10.1186/1471-2105-12-77

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98*(3), 223-237. http://dx.doi.org/10.1080/00223891.2015.1089249

Rosenfeld, B., Gibson, C., Kramer, M., & Breitbart, W. (2004). Hopelessness and terminal illness: The construct of hopelessness in patients with advanced AIDS. *Palliative and Supportive Care, 2*(1), 43–53. https://doi.org/10.1017/S1478951504040064

Rueda-Jaimes, G. E., Castro-Rueda, V. A., Rangel-Martínez-Villalba, A. M., Moreno-Quijano, C., Martinez-Salazar, G. A., & Camacho, P. A. (2018). Validation of the Beck Hopelessness Scale in patients with suicide risk. *Revista de Psiquiatría y Salud Mental (English Edition), 11*(2), 86-93. http://dx.doi.org/10.1016/j.rpsmen.2016.09.004

Salerno, L., Ingoglia, S., & Lo Coco, G. (2017). Competing factor structures of the Rosenberg Self-Esteem Scale (RSES) and its measurement invariance across clinical and non-clinical samples. *Personality and Individual Differences, 113*, 13–19. https://doi.org/10.1016/j.paid.2017.02.063

Savalei, V., & Falk, C. F. (2014). Recovering Substantive Factor Loadings in the Presence of Acquiescence Bias: A Comparison of Three Approaches. Multivariate Behavioral Research, 49(5), 407–424. https://doi.org/10.1080/00273171.2014.931800

Sueki, H. (2020). Relationship between Beck Hopelessness Scale and suicidal ideation: A short-term longitudinal study. *Death Studies*, 1-6. doi:10.1080/07481187.2020.1740833

Szabó, M., Mészáros, V., Sallay, J., Ajtay, G., Boross, V., Udvardy-Mészáros, À., Vizin, G., & Perczel-Forintos, D. (2016). The Beck Hopelessness Scale. Specific Factors of Method Effects? *European Journal of Psychological Assessment, 32*, 111-118. https://doi.org/10.1027/1015-5759/a000240

Shahar, G., Bareket, L., Rudd, M. D., & Joiner, T. E. (2006). In severely suicidal young adults, hopelessness, depressive symptoms, and suicidal ideation constitute a single syndrome. *Psychological Medicine, 36*(07), 913. http://dx.doi.org/10.1017/s0033291706007586

Steed, L. (2001). Further validity and reliability evidence for Beck Hopelessness Scale scores in a nonclinical sample. *Educational and Psychological Measurement, 61*(2), 303-316. http://dx.doi.org/10.1177/00131640121971121

Stucky, B. D., Thissen, D., & Orlando Edelen, M. (2012). Using logistic approximations of marginal trace lines to develop short assessments. *Applied Psychological Measurement, 37*(1), 41–57. https://doi.org/10.1177/0146621612462759

Szabó, M., Mészáros, V., Sallay, J., Ajtay, G., Boross, V., Udvardy-Mészáros, À., … Perczel-Forintos, D. (2016). The Beck Hopelessness Scale: Specific Factors of Method Effects? *European Journal of Psychological Assessment, 32*(2), 111–118. http://dx.doi.org/10.1027/1015-5759/a000240

The International Test Commission. (2006). International guidelines on computer-based and internet-delivered testing. *International Journal of Testing, 6*(2), 143–171. http://doi.org/10.1207/s15327574ijt0602_4

Tomas, J. M. & Oliver, A. (1999). Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 84-98. https://doi.org/10.1080/10705519909540120

Tsujii, N., Shirakawa, O., Niwa, A., Yonemoto, N., Kawanishi, C., Yamamoto, K., … Hirayasu, Y. (2020). Hopelessness is associated with repeated suicidal behaviors after discharge in patients admitted to emergency departments for attempted suicide. *Journal of Affective Disorders, 272*, 170–175. http://doi.org/10.1016/j.jad.2020.04.037

Vecco, G. A., Flores-Kanter, P. E., & Luque, L. E. (2021). Análisis psicométrico del inventario de orientación suicida ISO-19, en adolescentes cordobeses escolarizados [Psychometric analysis of the Inventory of Suicide Orientation ISO-19, in scholarized adolescents]. *Evaluar, 21*(1), 40-52. https://doi.org/10.35670/1667-4545.v21.n1.32831

Velting, D. M. (1999). Personality and negative expectations: Trait structure of the Beck Hopelessness Scale. *Personality and Individual Differences, 26*(5), 913–921. https://doi.org/10.1016/S0191-8869(98)00194-9

Weigold, A., Weigold, I. K., & Russell, E. J. (2013). Examination of the equivalence of self-report survey-based paper-and-pencil and internet data collection methods. *Psychological Methods, 18*(1), 53–70. https://doi.org/10.1037/a0031607

Wenzel, A., Brown, G. K., & Beck, A. T. (2009). *Cognitive therapy for suicidal patiens. Scientific and clinical applications*. Washington, DC: American Psychological Association.

Wetherall K., Cleare S., Eschle S., Ferguson E., O'Connor D. B., O'Carroll R. E., & O'Connor R. C. (2018). From ideation to action: Differentiating between those who think about suicide and those who attempt suicide in a national study of young adults. *Journal of Affective Disorders, 241*, 475–483. https://doi.org/10.1016/j.jad.2018.07.074

World Health Organization (2018). *Preventing suicide. A community engagement toolkit*. Geneva, Switzerland: Author.

Young, M. A., Halper, I. S., Clark, D. C., Scheftner, W., & Fawcett, J. (1992). An item-response theory evaluation of the Beck Hopelessness Scale. *Cognitive Therapy and Research, 16*, 579-587. http://dx.doi.org/10.1007/BF01175143

Yuan, K. H., & Bentler, P. M. (2007). Structural Equation Modeling, in C. R. Rao and S. Sinharay, *Handbook of Statistics,* Volume 26: Psychometrics (pp. 297–358). Amsterdam: North-Holland.

Zhang, X., & Savalei, V. (2015). Improving the Factor Structure of Psychological Scales. *Educational and Psychological Measurement, 76*(3), 357–386. https://doi.org/10.1177/0013164415596421

Zhu, Z., Liu, Y., Zhang, C., Yuan, Z., Zhang, Q., Tang, F., … Xue, F. (2013). Identification of Cardiovascular Risk Components in Urban Chinese with Metabolic Syndrome and Application to Coronary Heart Disease Prediction: A Longitudinal Study. *PLoS ONE, 8*(12), e84204. https://doi.org/10.1371/journal.pone.0084204

Table 1. Demographic characteristics of sample

	N	%
Gender		
Female	1,560	72.1
Male	604	27.9
Age group		
Young (≤19)	426	19.7
Adults (20–59)	1,553	71.8
Older (≥60)	185	8.5
Argentina region		
Northwest	269	12.42
Pampean	1,622	74.88
Cuyo	136	6.28
Patagonia	137	6.33
Actual treatment		
No	1,851	85.5
Yes	313	14.5
Presence of SI		
High	82	3.8
Low-Moderate	1,153	39.1
No SI	1,235	57.1

Note. Actual treatment = psychiatric or psychological. The presence of suicide ideation (SI) was identified based on the recategorization of scores obtained from the Inventory of Suicide Orientation-30 (ISO-30; range value = 4–16). No SI = a total score of 4 (i.e., the participant marked "strongly disagree" on all the items); High = a score of 16 (i.e., the participant marked "strongly agree" on all the items); Low-Moderate = correspond to values between High and No SI.

Table 2. Fit indices of the models proposed for BHS

Model Fit	χ^2	df	CFI	TLI	RMSEA	SRMR	Anomalous Results
Model 1	764.60	167	.983	.981	.041	.050	NO
Model 2	913.72	170	.979	.977	.045	.053	NO
Model 3	576.50	150	.988	.985	.036	.044	YES
Model 4	431.14	160	.992	.990	.030	.034	YES
Model 5	711.53	161	.985	.982	.040	.047	NO
Difference Testing	χ^2	df	*p*				
Model 1 vs 5	51.13	6	<.001				
Model 2 vs 5	169.66	9	<.001				
Model 3 vs 5	117.17	11	<.001				
Model 4 vs 5	227.13	15	<.001				

Note. BHS, Beck Hopelessness Scale. Model 1 = the original three-factor model by Beck et al. (1974); Model 2 = the one-dimensional model; Model 3 = the bifactor model, as suggested by Szabó et al. (2016); Model 4 = the multitrait-multimethod model suggested by Boduszek and Dhingra (2015); Model 5 = the correlated trait-correlated method minus one (CT-C(M-1)) model.

Table 3. Standardized factor loadings for CT-C(M-1) model

	Estimate	*SE*	Estimate/SE	*p*
Trait Factor				
BHS1	-.795	.017	46.802	< .001
BHS2	.885	.015	-58.702	< .001
BHS3	-.679	.024	27.790	< .001
BHS4	.545	.026	-21.351	< .001
BHS5	-.554	.025	22.502	< .001
BHS6	-.637	.034	18.992	< .001
BHS7	.888	.012	-76.749	< .001
BHS8	-.697	.02	35.680	< .001
BHS9	.882	.013	-67.318	< .001
BHS10	-.646	.023	27.910	< .001
BHS11	.899	.013	-71.722	< .001
BHS12	.78	.017	-46.966	< .001
BHS13	-.055	.043	1.263	.206
BHS14	.745	.017	-43.273	< .001
BHS15	-.798	.016	49.960	< .001
BHS16	.861	.015	-58.880	< .001
BHS17	.89	.013	-70.791	< .001
BHS18	.725	.017	-42.436	< .001
BHS19	-.719	.02	35.290	< .001
BHS20	.92	.011	-82.167	< .001
Method Factor				
BHS1	.38	.035	10.714	< .001
BHS3	.15	.042	3.544	< .001
BHS5	.141	.044	3.216	< .01
BHS6	.36	.052	6.914	< .001
BHS8	-.012	.042	-.295	.768

BHS10	.144	.041	3.501	< .001
BHS13	.353	.057	6.184	< .001
BHS15	.486	.038	12.668	< .001
BHS19	.218	.038	5.716	< .001

Note. CT-C(M-1), correlated trait-correlated method minus one.

Figure 1. Examples of CT-C(M-1), bifactor, and MTMM models

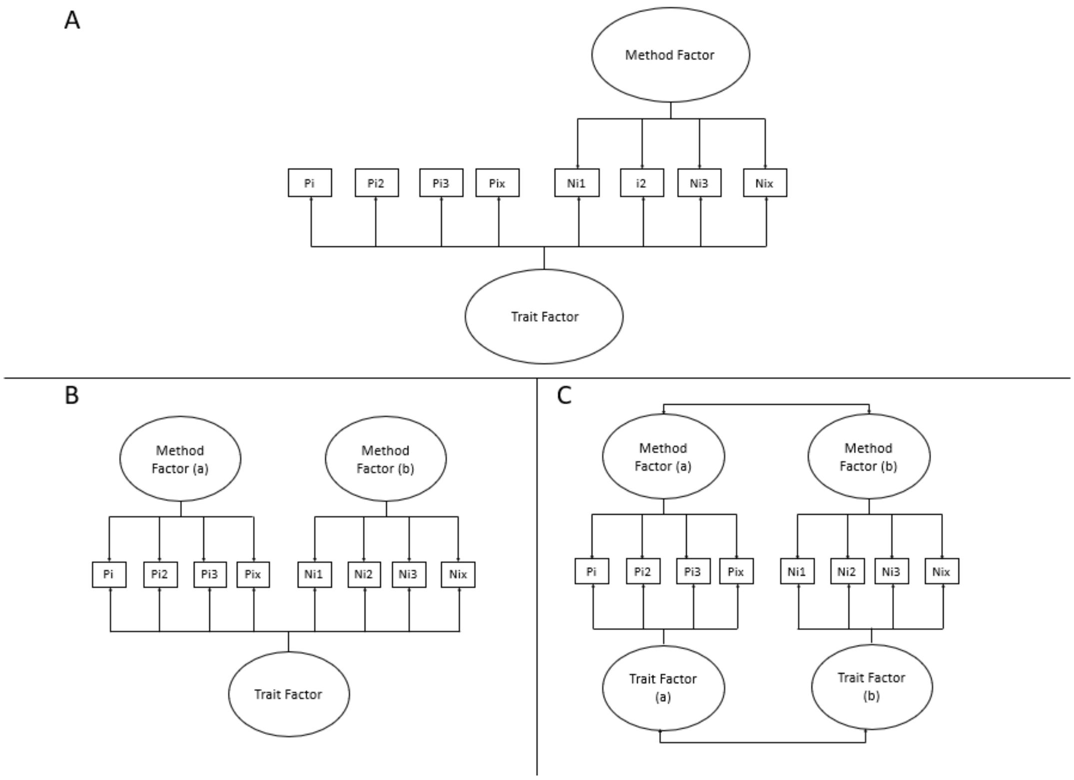

Figure 1. Pi = positive items (negatively worded items in the Beck Hopelessness Scale (BHS)), Ni = negative items (positively worded items in the BHS). A: The correlated traits-correlated methods minus one model. B is a fully symmetrical bifactor model (Szabó, et al., 2016) and C is a multitrait-multimethod model (Boduszek & Dhingra, 2015).

Note. CT-C(M-1), correlated trait-correlated method minus one; MTMM, multitrait-multimethod model.

Figure 2. Structural model: Trait factor and method factor effects on suicide ideation

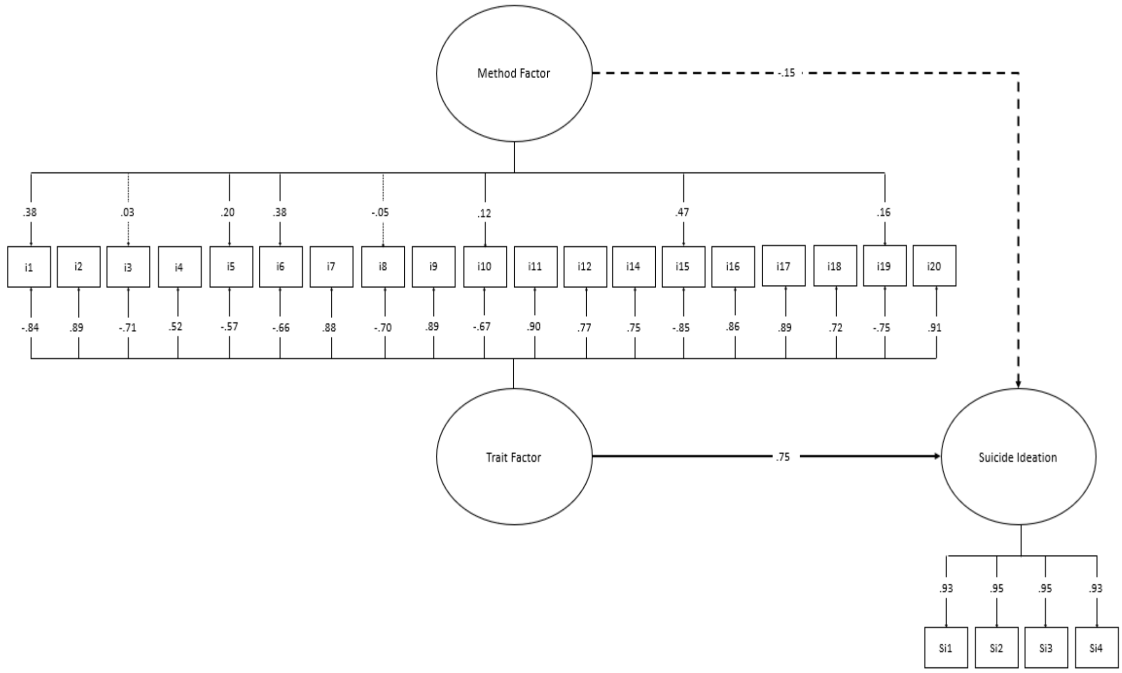

Figure 2. Standardized factor loadings. Dashed lines = statistically insignificant factor loading or regression coefficient ($p > .05$). Solid lines = statistically significant factor loading or regression coefficient ($p < .001$).

Figure 3. ROC curve

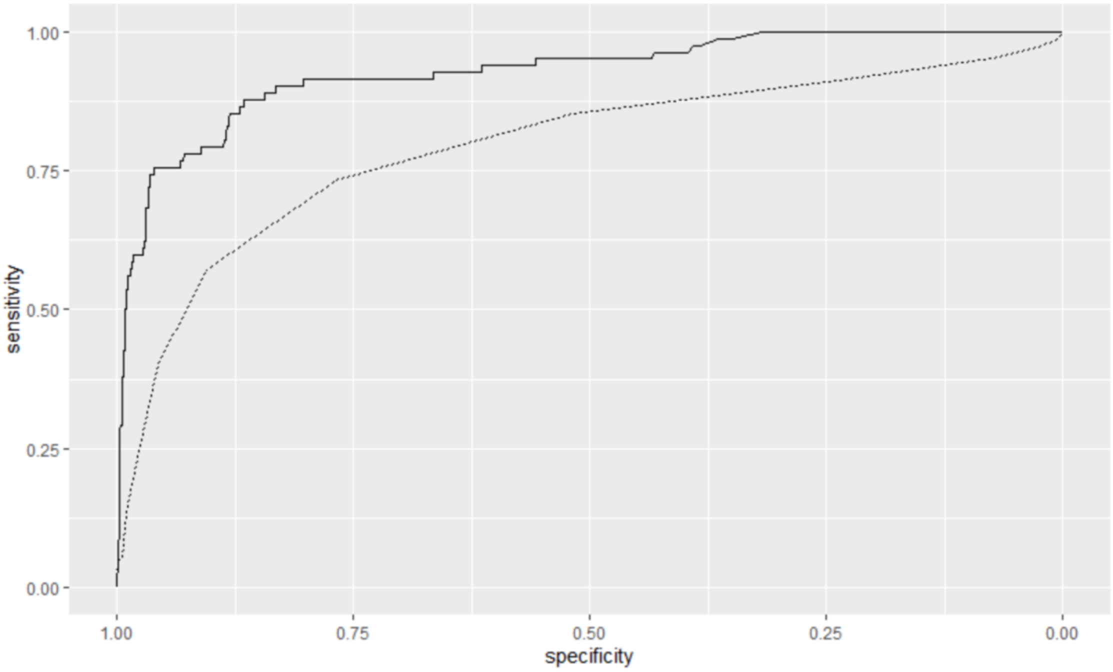

Figure 3. ROC curve obtained for the M-1 trait score (solid line) and BHS sum-score (dashed line).