

The Structure of the Career Decision-Making Difficulties Questionnaire Across 13 Countries

Journal of Career Assessment

2022, Vol. 0(0) 1–20

© The Author(s) 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/10690727221099226

journals.sagepub.com/home/jca

Nimrod Levin¹, Shagini Udayar², Yuliya Lipshits-Brazil¹,
Itamar Gati¹, and Jérôme Rossier² 

Abstract

Assessing the causes of career indecision is among the first steps in career counseling. Gati et al. (1996) proposed a multidimensional taxonomy of career indecision and developed the Career Decision-Making Difficulties Questionnaire (CDDQ), consisting of 10 scales that cohere into three higher-order clusters and a total score. However, studies investigating the CDDQ reported cross-cultural variations in its factor structure. To examine the cross-cultural generalizability of the CDDQ, we compared four alternative factor models using data from 39 diverse samples from 13 countries with nine language versions ($N = 19,562$). Using weighted least squares mean- and variance-adjusted estimation, a robust estimator for nonnormal data, comparison of fit indices supported the original CDDQ structure across countries and languages. These findings support the cross-cultural generalizability of the structure of the CDDQ and the use of 10 scale scores, three cluster scores, and a total score, consistent with the taxonomy underlying the CDDQ.

Keywords

career assessment, career decision-making, career decision-making difficulties, career indecision, confirmatory factor analysis

Introduction

Overcoming career decision-making difficulties is considered a precondition for making satisfying career decisions (Peterson et al., 2002). For this reason, counseling individuals with career decision-making difficulties often begins with assessing the causes of their difficulties,

¹Hebrew University of Jerusalem, Israel

²University of Lausanne, Switzerland

Corresponding Authors:

Jérôme Rossier, University of Lausanne, Switzerland.

Email: jrossier@unil.ch

Nimrod Levin, Hebrew University of Jerusalem, Israel.

Email: nimrod.levin@mail.huji.ac.il

enabling practitioners to tailor interventions to clients' unique needs. The term *career indecision* was adopted to refer to the various difficulties encountered by individuals during the process of choosing a career, suggesting that individuals experiencing career indecision are less likely to identify a preferred alternative or develop sufficient commitment to an educational or vocational choice (Osipow, 1999).

Gati et al. (1996) proposed a taxonomy of career decision-making difficulties and developed a theory-driven, widely used assessment of career indecision: the Career Decision-Making Difficulties Questionnaire (CDDQ). The taxonomy underlying the CDDQ conceptualizes career indecision as multidimensional with the dimensions of career indecision hierarchically arranged at three levels: 32 items yield 10 scale scores; and they can be combined to represent three higher-order clusters as well as a total score. To date, the CDDQ has been translated into 48 languages and implemented in research and practice in more than 60 countries (Levin et al., 2020; Xu & Bhang, 2019). Nevertheless, factor analyses involving different language versions of the CDDQ in many countries did not result in a consistent and theoretically defensible structure. Thus, despite its widespread use, the factor structure of the CDDQ remains a topic for debate. In this research, we addressed this issue by examining and comparing four models of the factor structure of the CDDQ on 39 diverse samples collected in 13 countries using nine CDDQ language versions.

Assessing Career Indecision

The first standardized methods for assessing the *degree* of career indecision were developed in the 1970s. These measures represented unidimensional assessments of the *degree* of career indecision, such as the Career Decision Scale (Osipow et al., 1976) and the My Vocational Situation (Holland et al., 1980). Second-generation measures of career indecision adopted a multidimensional approach for assessing the *nature* of career indecision, namely, its causes; these included the Career Decision Profile (Jones, 1989) and the Career Factors Inventory (Chartrand et al., 1990). However, both first- and second-generation measures were developed independently of theoretical considerations, but rather mostly based on the clinical impressions of practitioners seeking to map systematically the various reasons for being career undecided (Osipow, 1999).

The first theory-grounded multidimensional assessment of career indecision was developed based on Gati et al.'s (1996) taxonomy of career decision-making difficulties. Informed by decision theory, this taxonomy focused on the causes of career indecision, differentiating between three clusters: Lack of Readiness (lack of motivation, general indecisiveness, and dysfunctional beliefs), Lack of Information (about the decision-making process, the self, occupations or careers, and ways of obtaining information or help), and Inconsistent Information (unreliable information, internal conflicts, and external conflicts). This taxonomy postulated that the various causes of career indecision could be combined to form a global indicator of the degree of the overall severity of career indecision. The Career Decision-Making Difficulties Questionnaire (CDDQ) was developed based on this taxonomy, with its structure conceptualized to comprise 10 first-order scales, three second-order clusters, and one third-order total score.

The Empirical Structure of the CDDQ

The first version of the CDDQ included 44 items corresponding to the 10 specific career decision-making difficulties as proposed in Gati et al.'s (1996) taxonomy. In line with this taxonomy, the hypothesized multidimensional, three-order hierarchical structure was supported for the Hebrew

version of the CDDQ using cluster analysis (Gati et al., 1996, 2000). However, analyses of the English and Chinese versions indicated that one of the 10 first-order scales—external conflicts—was positioned within the Lack of Readiness second-order cluster rather than the Inconsistent Information cluster (Gati et al., 1996; Mau, 2001; Osipow & Gati, 1998). Consequently, Mau (2001) suggested adopting a modified structure, positioning external conflicts under the Lack of Readiness cluster, whereas Xu and Tracey (2014) proposed an alternative five-factor structure with lack of motivation, general indecisiveness, dysfunctional beliefs, Lack of Information, and Inconsistent Information (including external conflicts) as five first-order factors of career indecision.

Based on accumulated psychometric data and to shorten the questionnaire, a 34-item version of the CDDQ was developed by either removing items that did not contribute significantly to their scale or combining highly correlated items (Gati & Saka, 2001). Given that most researchers and practitioners in the last two decades have used the shortened 34-item CDDQ, we focused our research on examining the structure of this version. Thirteen validation studies that used confirmatory factor analysis (CFA) to examine the structure of the 34-item CDDQ have been published (see Supplemental Appendix Table 1 in the online supplementary materials for a summary). Of these 13 studies, only three established full structural validation of the original CDDQ taxonomy at the item level. Specifically, validation of the factor structure corresponding to the original taxonomy (32–10–3–1, for 32 items as indicators, 10 first-order scales, three second-order clusters, and a third-order total score, respectively) was reported for the English (Levin et al., 2020) and French (Rossier et al., 2021; Sovet et al., 2017) CDDQ versions. Additional studies—of the Turkish (Bacanli, 2016), French (Atitsogbe et al., 2018), and English and Hebrew (Kleiman et al., 2004; Kleiman & Gati, 2004) versions—supported the original structural model of the CDDQ but did so based on all-item parceling (i.e., validating a 10–3–1 model, for 10 scale scores as indicators, three first-order clusters, and one second-order total score, respectively).

In contrast, six validation studies examining the Chinese, Croatian, Farsi, Greek, and Korean versions of the CDDQ did not confirm the factor model corresponding to Gati et al.'s (1996) taxonomy and proposed alternative factor structures. An exploratory factor analysis of the 34-item Chinese version indicated poor factor loadings of items in the Lack of Readiness second-order cluster. For this reason, an alternative model that excluded the Lack of Readiness cluster and discarded the ten first-order scales (i.e., a 19–2 model) was examined, yielding a better fit than the original model did (Creed & Yin, 2006). Other studies focused on the poor factor loadings of the dysfunctional beliefs first-order scale—a scale included in the Lack of Readiness second-order cluster—providing incremental support for an alternative model that excluded this first-order scale (Croatian version: Babarović & Šverko, 2019; Korean version: Sovet et al., 2015). For the Farsi version of the CDDQ, Vahedi et al. (2012) provided support for a model that excluded the 10 first-order scales (i.e., a 32–3–1 model), and for the Greek version, Vaiopoulou et al. (2019) preferred a model that excluded three first-order scales.

In total, at least five CDDQ factor structures have been promoted to date. Reported differences in the factor structure of the CDDQ emerged across countries and not within countries, thus questioning the cross-cultural equivalence of career indecision as measured by the CDDQ. Nevertheless, previous research supported the measurement invariance of the CDDQ across countries when the same language version was used (Atitsogbe et al., 2018; Levin et al., 2020; Rossier et al., 2021). Thus, rather than being related to cross-cultural variability, the observed structures may reflect that variations are not only related to cross-cultural differences but potentially also to differences across language versions of the CDDQ. We tested this hypothesis in our research by evaluating the factor structure across 13 countries using nine language versions of the CDDQ. The available data allowed us to examine whether different factor structures would

emerge across countries that used different language versions (e.g., Poland vs. Italy) as well as countries in which the same language version was used (e.g., Canada, France, and Switzerland with the French version).

Potential Pitfalls in Previous CFAs of the CDDQ Factor Structure

Upon examining the analytical approaches of previous investigations of the structure of the CDDQ, we identified two recurring methodological decisions that may have undermined the validity of the results and conclusions of previous studies. The first decision involves the issue of parceling, namely the aggregation of items and use of parcels as indicators of latent constructs (Marsh et al., 2013; Sellbom & Tellegen, 2019). Most of the previous structural investigations of the CDDQ that validated its structure employed *all-item parceling* (i.e., all items in each factor are aggregated into one respective parcel). However, Sellbom and Tellegen (2019) argued that “when we evaluate the internal psychometric properties of a scale, the items responded to constitute a key focus” (p. 1429). Indeed, parceling precludes evaluating the parameter properties of individual items and, consequently, impedes factor validation; poor item loadings, which undermine the construct validity of the examined measure, cannot be identified when using parceling. Marsh et al. (2013) found that parceling items with even small cross-loadings can significantly distort results regarding the number of factors. Thus, although remaining a topic of ongoing debate (e.g., Kline, 2016), parceling may be counterproductive when the primary goal is examining the factor structure of a measure. For this reason, in our research, we did not consider CDDQ factor models that involve item parceling.

The second methodological decision implemented in previous studies of the structure of the CDDQ concerns using maximum likelihood (ML) for model estimation without ensuring that indicators conform to a multivariate normal distribution, a core assumption of ML (Beauducel & Herzberg, 2006; Li, 2016). About half of the published studies on the structure of the CDDQ reported using ML, whereas the other half did not indicate which estimator was used. Nonetheless, because ML estimation is the default estimator in most latent variable modeling software, it is likely to have been implemented. Relatedly, Sellbom and Tellegen (2019) noted that ML assumptions are often violated in psychological assessment research, namely, when indicators are Likert-type scale items or response distributions are asymmetrical (see also Kline, 2016), but no study has examined the distribution of the CDDQ items or provided any justification for the choice of ML.

Testing and comparing models with data that violate the assumptions of the implemented estimator can lead to adopting spurious models and rejecting better ones. When the distribution of responses significantly diverges from a normal distribution, indicators should be treated as ordered-categorical (ordinal). In such cases, weighted least squares estimators constitute a better and more robust alternative to ML. Despite several weighted least squares estimators, weighted least squares mean- and variance-adjusted (WLSMV) estimation is considered preferable (Kline, 2016; Sellbom & Tellegen, 2019). Indeed, simulation studies have shown that WLSMV outperforms ML estimation in the precision of estimation of parameters across various conditions (Beauducel & Herzberg, 2006; Li, 2016). Thus, if the use of ML for the examination of the factor structure of the CDDQ is inadequate, then the conclusions of previous studies based on ML can be called into question.

The Present Study

The main goal of the study was to examine the factor structure of the CDDQ and the compatibility of the best-fitting solution with Gati et al.'s (1996) taxonomy of career decision-making

difficulties. Inconsistencies regarding the factor structure of the CDDQ reported in previous studies may have been confounded by the use of various language versions. For this reason, our research examined the factor structure of the CDDQ both within and across countries, with data derived from 39 samples collected in 13 countries from five continents, using nine language versions. In addition, inconsistencies regarding the factor structure of the CDDQ could be the result of the estimators used in the previous analyses of the CDDQ data. For this reason, we first inspected whether the CDDQ items are normally distributed to inform our decision of which estimator to use in the structural analyses—ML or WLSMV, the latter constituting the robust estimator when dealing with nonnormal data.

After determining which estimator to implement, we compared four alternative models: (a) the *original CDDQ model* that corresponds to Gati et al.'s (1996) taxonomy and includes 10 first-order scales, three second-order clusters, and one third-order total score, (b) a *10-factor model* that includes 10 first-order scales and one second-order total score but without the three cluster-level factors, (c) a *3-factor model* that includes three first-order clusters and one second-order total score but without the 10 scale-level factors, and (d) a *unidimensional model* comprising only a single total score. We determined the labels of the three alternative models by the number of first-order factors included in each model. We did not test additional models that involve excluding some of the CDDQ items given substantial evidence supporting the criterion validity of the total score of the CDDQ, which is computed based on the aggregation of all CDDQ items. Finally, we also examined the measurement invariance of the CDDQ across countries in which the same language version was used to examine further the effect of the cultural context on the measurement of career indecision without confounding differences that might be related to translation. The data analyzed in this study represent the various populations to which the CDDQ is typically administered, including middle school, high school, preacademic, college student, mixed (e.g., students and workers), and counseling samples. In this regard, by analyzing diverse data in the study, we aimed to provide conclusions that are more definitive regarding the adequacy of the 10 CDDQ scale scores, three cluster scores, and one total score.

Method

Transparency and Openness

We report the way we determined the sample size and all data exclusions, manipulations, and measures in the study, and we follow journal article reporting standards. The data, analysis codes, and research materials are available by request from the first author. Data preparation and normality analyses were conducted in Python, and structural and reliability analyses were conducted in R. CFAs were conducted using the lavaan package in R (Rosseel, 2012). The design of the study and its analysis were not preregistered.

Procedure

To examine the factor structure of nine language versions of the 34-item CDDQ across 13 countries, we relied on secondary data previously collected for diverse research purposes by the authors (10 samples) and other researchers (29 samples). We did not include 22 additional data sets for the following reasons: 15 samples were excluded because their aggregated country samples were too small for CFA analyses considering the complexity of the compared models ($N > 500$; Kline, 2016; Wolf et al., 2013), and seven samples were excluded because they were used in a previous study testing the factor structure of the English version of the CDDQ and would have

biased the results of the study given their overall large sample size ($N = 32,556$; Levin et al., 2020). We ensured that local institutional review board approvals had been obtained for all included data sets (see Supplemental Appendix 3 in the online supplementary materials for details on the sources and procedures of each sample).

Participants

We obtained the data of 21,852 participants from 39 samples collected in 13 countries. We carried out a series of preliminary validity inspections to ensure that only the data of participants who completed the CDDQ with sufficient attention were included in the main analyses. The data of 2,267 participants were excluded for the following reasons: 757 (3.5%) due to missing data, 203 (0.9%) due to repeatedly selecting only one or two responses, 1,307 (6.0%) because their responses to the two embedded validity items indicated questionable attention to content, and 23 (0.1%) because they were older than 50 to reduce the likelihood of age-related outliers. The number of excluded participants and demographic information are presented separately for each country sample in Table 1. Across samples, the 19,562 participants included participants aged 13 to 50 ($M_{\text{age}} = 19.36$, $SD = 4.19$). Of the 19,415 (99.2%) participants with gender data, 61.6% identified as women. In terms of sample type, participants represented one middle school, 17 high schools, three preacademic programs, 12 college student samples, three mixed samples, and three counseling samples.

Measures

The Career Decision-Making Difficulties Questionnaire. The 34-item version of the CDDQ comprises 32 items representing career decision-making difficulties and two validity items (Gati et al., 1996; Gati & Saka, 2001). The CDDQ measures 10 specific causes of difficulties grouped in three major clusters of Lack of Readiness, Lack of Information, and Inconsistent Information. In addition, a total difficulty score can be computed. The Lack of Readiness cluster includes three scales: (a) lack of motivation (three items; e.g., “I know that I have to choose a career, but I don’t have the motivation to make the decision now”), (b) general indecisiveness (three items; e.g., “It is usually difficult for me to make decisions”), and (c) dysfunctional beliefs (four items; e.g., “I expect that

Table 1. Sample Characteristics.

Country	Language	Initial <i>N</i>	Final <i>N</i>	<i>M</i> Age (<i>SD</i>)	Women (%)
Canada	French	1,130	878	17.4 (0.6)	60.0
Croatia	Croatian	1,369	1,199	16.9 (1.5)	66.6
France	French	1,922	1,800	18.8 (2.4)	71.5
Greece	Greek	2,607	2,390	17.5 (3.4)	64.4
Israel	Arabic	1,535	1,210	17.0 (0.3)	59.9
Israel	Hebrew	3,936	3,646	23.3 (5.1)	64.2
Malaysia	English	977	945	18.9 (2.9)	59.2
Poland	Polish	710	639	18.5 (2.6)	59.9
South Africa	English	782	727	NA	65.6
South Korea	Korean	1,467	1,289	19.2 (2.9)	60.1
Switzerland	French	1,226	1,087	18.4 (4.3)	47.6
Turkey	Turkish	3,207	2,983	15.4 (1.0)	53.8
USA	English	984	769	19.9 (3.1)	65.4

Note. Age, sex, and country statistics are reported for the final *N*s.

entering the career I choose will also solve my personal problems”). The Lack of Information cluster includes Lack of Information about (a) the decision-making process (three items; e.g., “I find it difficult to make a career decision because I do not know what steps I have to take”), (b) the self (four items; e.g., “I find it difficult to make a career decision because I am not yet sure about my career preferences”), (c) occupations or careers (three items; e.g., “I find it difficult to make a career decision because I do not have enough information about the variety of occupations or training programs that exist”), and (d) ways of obtaining information or help (two items; e.g., “I find it difficult to make a career decision because I do not know how to obtain accurate and updated information about the existing occupations and training programs, or about their characteristics”). The Inconsistent Information cluster includes three scales: (a) unreliable information (four items; e.g., “I find it difficult to make a career decision because I have contradictory data about the existence or the characteristics of a particular occupation or training program”), (b) internal conflicts (five items; e.g., “I find it difficult to make a career decision because I’m equally attracted to a number of careers and it is difficult for me to choose among them”), and (c) external conflicts (two items; e.g., “I find it difficult to make a career decision because people who are important to me, such as parents or friends, do not agree with the career options I am considering and/or the career characteristics I desire”). Whereas responses are typically reported on a 9-point Likert-type scale ranging from 1 (*does not describe me*) to 9 (*describes me well*), in three of the samples (Malaysia 03, Turkey 01, and Turkey 03), responses were elicited using a 5-point Likert-type scale. Scale, cluster, and total scores are calculated as the means of the responses; higher scores indicate more severe career decision-making difficulties. For the 34-item English version of the CDDQ, Levin et al. (2020) reported Cronbach alpha internal consistency reliability estimates of .94 for the total score and .66, .94, and .88 for the three clusters, as well as a median internal reliability estimate of .79 for the 10 scale scores. In addition to the original English and Hebrew 34-item versions, seven additional language versions were examined in our study: Arabic, Croatian, French, Greek, Korean, Polish, and Turkish (see Supplemental Appendix 4 in the online supplementary materials for a summary of their psychometric properties).

Data Analysis

We surmised that the inconsistencies of previous results could be related to the issue of item properties and the use of ML estimation. Therefore, in this study, we first examined the degree of item normality, so if severe divergence from normality was detected, the fit of the structural models would be estimated with WLSMV. We relied on three criteria to evaluate the normality of the 32 CDDQ items (excluding the two validity items): (a) kurtosis and skewness values, referring to absolute values greater than 1 as indicative of nonnormality, (b) the Shapiro–Wilk test statistics for univariate normality, and (c) Mardia’s (1970) test statistics for multivariate normality. Choice of an estimator would depend on the results of these analyses: ML, which was applied in previous CDDQ validation studies (see Supplemental Appendix Table 1 in the online supplementary materials), is discouraged for use in analyzing nonnormal data; instead, the WLSMV estimator is preferable in such cases (Li, 2016; Sellbom & Tellegen, 2019).

Structural analyses were conducted using the lavaan package in R (Rosseel, 2012). Four CFA models were tested and subsequently compared to investigate the structure underpinning the CDDQ: (a) the *original CDDQ model* with 10 first-order scales, three second-order clusters, and one third-order total score (32–10–3–1); (b) a *10-factor model* with 10 first-order scales and one second-order total score (32–10–1); (c) a *three-factor model* with three first-order clusters and one second-order total score (32–3–1); and (d) a *unidimensional model* with only a single first-order total score (32–1). In all models, items were only allowed to load on their target factors with cross-loadings constrained to zero. These four models were tested both within and across the 13 country samples.

To assess model fit, we relied on various standard goodness-of-fit indices with their respective thresholds (Hu & Bentler, 1999; Marsh et al., 2005). These indices included the comparative fit index (CFI; $\geq .95$ for good, $\geq .90$ for acceptable), the Tucker-Lewis index (TLI; $\geq .95$ for good, $\geq .90$ for acceptable), the root mean square error of approximation (RMSEA; $\leq .06$ for good, $\leq .08$ for acceptable), and the standardized root mean square residual (SRMR; $\leq .06$ for good, $\leq .10$ for acceptable). Because the χ^2 statistic is sensitive to sample size (Marsh et al., 2005), we did not evaluate model fit using this index. Furthermore, we did not compare models using Akaike's information criterion or Bayesian information criterion because these fit indices are less reliable for weighted least squares estimators. Nevertheless, whereas the purpose of Akaike's information criterion and the Bayesian information criterion is to estimate which models are more likely to replicate in subsequent samples, the design of our study allowed testing of the cross-validation of models directly.

In addition to evaluating models based on fit indices, we also inspected parameter estimates (e.g., factor loadings and correlations) and the theoretical conformity of the models (Marsh et al., 2004; Sellbom & Tellegen, 2019). First, a combination of positive and negative factor loadings denotes that indicators should be reversed. Second, weak factor loadings ($\lambda < .40$; Sellbom & Tellegen, 2019) reflect that indicators do not explain a significant proportion of the variance of their target factors. Third, large factor correlations ($r > .80$; Brown, 2015) indicate that factors may in fact measure the same construct (i.e., poor discriminant validity). Fourth, regarding theoretical conformity, when model fit indices and parameter estimates do not provide enough information to identify the best model, a model with greater theoretical conformity should be preferred.

We also assessed the reliability of the CDDQ items in the retained model. We calculated composite reliability (CR; Raykov, 1998) rather than Cronbach's α because the latter is considered a biased estimator for determining the reliability of factors (Yanyun Yang & Green, 2011). In contrast, CR values are model-based, and their calculation aggregates factor loadings and measurement errors. CR values above .60 are considered acceptable, and CR values above .70 are good (Bagozzi & Yi, 1988).

Finally, we estimated the measurement invariance of the CDDQ across countries in which the same language version was used. A series of multiple group confirmatory factor analyses were performed, sequentially constraining parameters in three models. The initial configural model tests the equivalence of the factor structure across compared groups (i.e., the number of factors and patterns of factor loadings). Then, in the metric model, factor loadings are constrained to be equal across groups. Finally, in the scalar model, indicator intercepts are constrained to be equal across groups. By sequentially comparing and testing the decrease in fit across models, measurement invariance can be evaluated. When a less restrictive model yields better fit, the assumption of measurement invariance for the more restrictive model should be rejected. In our research, we evaluated fit differences using three criteria: changes in CFI, RMSEA, and SRMR values. A change of .010 in CFI supplemented by a change of .015 in RMSEA or .030 in SRMR indicates a lack of metric invariance; a change of .010 in CFI supplemented by a change of .015 in RMSEA or a change of .010 in SRMR indicates a lack of scalar invariance (Chen, 2007).

Results

Normality Analyses

Table 2 presents the findings concerning the univariate normality and multivariate normality of items for the total sample across countries and for each country sample separately. For the total sample, 13 (41%) items had kurtosis values greater than 1 ($|K| = 0.03$ to 1.87, $|K|_{\text{median}} = 0.92$), including all items from the general indecisiveness, lack of information about the process, and lack of information about occupations scales, indicating that participants were relatively polarized in the

Table 2. Statistics and Tests of Univariate and Multivariate Normality Across and Within Country Samples.

Country	Language	Univariate Normality				SW	Multivariate Normality	
		Kurtosis		Skewness			Kurtosis	Skewness
		$ K _{\text{median}}$	Items (%)	$ S _{\text{median}}$	Items (%)			
Total		0.92	41	0.52	12	0%	—	—
Canada	French	0.81	25	0.58	22	0%	93.15*	15,577.99*
Croatia	Croatian	0.87	31	0.41	0	0%	95.98*	14,873.69*
France	French	0.96	44	0.46	9	0%	134.72*	18,067.08*
Greece	Greek	0.95	41	0.60	6	0%	158.27*	21,297.90*
Israel	Arabic	1.06	59	0.52	9	0%	93.06*	16,388.98*
Israel	Hebrew	1.20	72	0.28	9	0%	124.52*	24,942.15*
Malaysia	English	1.08	59	0.27	6	0%	114.69*	15,964.52*
Poland	Polish	0.85	31	0.34	0	0%	61.30*	12,570.45*
South Africa	English	0.76	34	0.65	16	0%	97.35*	16,105.24*
South Korea	Korean	0.70	6	0.29	3	0%	113.89*	15,752.62*
Switzerland	French	0.96	47	0.47	12	0%	80.80*	14,276.48*
Turkey	Turkish	2.00	91	1.19	75	0%	415.47*	51,172.36*
USA	English	0.85	9	0.47	3	0%	98.28*	15,998.21*

Note. % items = percentage of items with values greater than one; SW = percentage of items that were normally distributed based on the Shapiro–Wilk Test at the $p < .001$ level. * $p < .001$.

endorsement of these items. Four (12%) items had skewness values greater than 1 ($|S| = 0.13$ to 1.55, $|S|_{\text{median}} = 0.52$), including two items from the lack of motivation factor and two items from the external conflicts factor, indicating that few participants endorsed these items. A series of Shapiro–Wilk tests was significant at the $p < .001$ for all items for the total sample and all 13 country samples, further indicating consistent univariate nonnormality. Finally, Mardia's (1970) multivariate estimates for the country samples indicated severe multivariate nonnormality. We could not calculate multivariate estimates for the total sample due to its large size. Given the nonnormality of data, we report the results of the structural analyses performed with WLSMV. A summary of results performed with ML is available in [Supplemental Appendix Table 2](#) in the online supplementary materials.

Structural Analyses

As can be seen in the upper part of [Table 3](#), the *original CDDQ model* (32–10–3–1) resulted in good fit in the total sample (CFI = .985, TLI = .983, RMSEA = .041, 95% CI [.040–.041], SRMR = .041) as well as in 11 of the 13 country samples (i.e., CFI and TLI > .95, RMSEA and SRMR < .06; see [Table 3](#)); for the Swiss (French) and Polish samples, the fit was acceptable (CFI = .967/.970, TLI = .964/.967, RMSEA = .055/.048, 95% CI [.052–.057]/[.045–.052], SRMR = .060/.061, respectively). The *10-factor model* (32–10–1) yielded a good but nonetheless slightly lower fit in the total sample (CFI = .981, TLI = .979, RMSEA = .046, 95% CI [.045–.046], SRMR = .046) as well as in six of the 13 country samples, whereas fit indices for the remaining seven country samples were lower and within the acceptable range (i.e., CFI and TLI > .90, RMSEA < .08 and SRMR < .10). The third, *three-factor model* (32–3–1), resulted in good but slightly lower fit indices for the total sample (CFI = .972, TLI = .970, RMSEA = .055, 95% CI [.054–.055], SRMR = .056) and for only the Turkish sample. Finally, the *unidimensional model* (32–1) yielded only acceptable

Table 3. Fit Indices for the Total and Country Samples: CFA with WLSMV Estimation.

Sample	Model	CFI	TLI	RMSEA	95% CI	SRMR
Total	<i>original</i>	.985	.983	.041	[.040-.041]	.041
	<i>1 0-factor</i>	.981	.979	.046	[.045-.046]	.046
	<i>three-factor</i>	.972	.970	.055	[.054-.055]	.056
	<i>unidimensional</i>	.962	.959	.064	[.063-.064]	.065
Canada (French)	<i>original</i>	.988	.987	.038	[.035-.041]	.050
	<i>1 0-factor</i>	.985	.984	.042	[.039-.045]	.054
	<i>three-factor</i>	.974	.972	.057	[.054-.059]	.067
	<i>unidimensional</i>	.965	.962	.065	[.063-.068]	.075
Croatia (Croatian)	<i>original</i>	.974	.971	.054	[.051-.056]	.058
	<i>1 0-factor</i>	.971	.968	.057	[.054-.059]	.061
	<i>three-factor</i>	.958	.954	.068	[.066-.070]	.074
	<i>unidimensional</i>	.951	.948	.073	[.070-.075]	.078
France (French)	<i>original</i>	.973	.971	.052	[.050-.054]	.055
	<i>1 0-factor</i>	.971	.968	.054	[.053-.056]	.058
	<i>three-factor</i>	.964	.961	.060	[.058-.062]	.064
	<i>unidimensional</i>	.957	.954	.065	[.063-.067]	.069
Greece (Greek)	<i>original</i>	.974	.971	.047	[.045-.049]	.050
	<i>1 0-factor</i>	.970	.967	.050	[.049-.052]	.053
	<i>three-factor</i>	.962	.960	.056	[.054-.058]	.059
	<i>unidimensional</i>	.954	.951	.062	[.060-.063]	.065
Israel (Arabic)	<i>original</i>	.980	.979	.045	[.043-.048]	.058
	<i>1 0-factor</i>	.977	.975	.049	[.047-.052]	.061
	<i>three-factor</i>	.973	.971	.052	[.050-.055]	.074
	<i>unidimensional</i>	.965	.962	.060	[.058-.062]	.078
Israel (Hebrew)	<i>original</i>	.958	.954	.054	[.053-.055]	.054
	<i>1 0-factor</i>	.943	.938	.063	[.062-.064]	.062
	<i>three-factor</i>	.933	.927	.068	[.067-.069]	.069
	<i>unidimensional</i>	.897	.890	.084	[.082-.085]	.084
Malaysia (English)	<i>original</i>	.989	.988	.039	[.036-.042]	.049
	<i>1 0-factor</i>	.985	.984	.046	[.043-.049]	.054
	<i>three-factor</i>	.977	.975	.057	[.054-.059]	.063
	<i>unidimensional</i>	.968	.965	.067	[.064-.069]	.072
Poland (polish)	<i>original</i>	.970	.967	.048	[.045-.052]	.061
	<i>1 0-factor</i>	.966	.962	.052	[.048-.055]	.064
	<i>three-factor</i>	.953	.949	.060	[.057-.064]	.071
	<i>unidimensional</i>	.940	.935	.068	[.065-.071]	.078
South Africa (English)	<i>original</i>	.985	.983	.046	[.043-.050]	.058
	<i>1 0-factor</i>	.982	.980	.050	[.047-.053]	.062
	<i>three-factor</i>	.979	.978	.053	[.050-.056]	.065
	<i>unidimensional</i>	.973	.971	.061	[.058-.064]	.072
South Korea (Korean)	<i>original</i>	.971	.968	.049	[.047-.051]	.057
	<i>1 0-factor</i>	.964	.960	.055	[.052-.057]	.062
	<i>three-factor</i>	.955	.952	.060	[.058-.062]	.069
	<i>unidimensional</i>	.938	.933	.071	[.069-.073]	.079
Switzerland (French)	<i>original</i>	.967	.964	.055	[.052-.057]	.060
	<i>1 0-factor</i>	.960	.957	.059	[.057-.062]	.065
	<i>three-factor</i>	.957	.953	.062	[.059-.062]	.068
	<i>unidimensional</i>	.943	.939	.070	[.068-.073]	.076

(continued)

Table 3. (continued)

Sample	Model	CFI	TLI	RMSEA	95% CI	SRMR
Turkey (Turkish)	<i>original</i>	.994	.994	.018	[.016–.020]	.035
	<i>10-factor</i>	.992	.991	.021	[.019–.022]	.039
	<i>three-factor</i>	.985	.984	.028	[.026–.029]	.047
	<i>unidimensional</i>	.981	.979	.032	[.030–.033]	.053
USA (English)	<i>original</i>	.990	.989	.039	[.035–.042]	.052
	<i>10-factor</i>	.989	.988	.041	[.038–.044]	.054
	<i>three-factor</i>	.982	.980	.052	[.049–.055]	.062
	<i>unidimensional</i>	.977	.975	.058	[.055–.061]	.068

Note. CFA = confirmatory factor analysis; WLSMV = weighted least squares mean- and variance- adjusted; CFI = comparative fit index; TLI = Tucker–Lewis Index; RMSEA = root mean square of approximation; SRMR = standardized root mean square residual.

fit for the total sample (CFI = .962, TLI = .959, RMSEA = .064, 95% CI [.063–.064], SRMR = .065) as well as in 12 of the 13 country samples, whereas the fit indices for the Israeli Hebrew sample were below the acceptable thresholds (CFI = .897, TLI = .890, RMSEA = .084, 95% CI [.082–.085], SRMR = .084). Thus, the *original CDDQ model* was slightly but consistently better than the *10-factor model* was, which, in turn, was better than the *three-factor model* and the *unidimensional model* were.

Because the differences between the fit indices of the *original CDDQ model* and the *10-factor model* were minimal, we sought to clarify further which of these two models best fit the data by inspecting their respective standardized factor loadings and factor correlations. As Table 4 shows, in the *original CDDQ model*, factors were mostly well-defined by their respective factor loadings at the level of the 10 first-order scales ($\lambda = .32\text{--}.88$, $M = .69$), the level of the three second-order clusters ($\lambda = .27\text{--}.99$, $M = .78$), and the level of the third-order total score ($\lambda_s = .96, .93, .90$), yielding moderate factor intercorrelations ($r_{\text{median}} = .57$). In comparison, the factor loadings for the *10-factor model*—in which the three cluster-level factors were excluded—were highly similar to those of the *original CDDQ model* at the level of the first-order scales ($\lambda = .31\text{--}.88$, $M = .69$) and the total score ($\lambda = .26\text{--}.93$, $M = .74$), yielding comparable moderate factor intercorrelations ($r_{\text{median}} = .52$). The factor loadings associated with the dysfunctional beliefs scale were low in these two models. Thus, inspections of factor loadings and factor correlations were uninformative for determining which of these two models presented a better fit to the data.

In contrast, the *three-factor model* and the *unidimensional model*—from which the 10 scale-level factors were excluded—resulted in lower factor loadings of indicators on their target first-order factors ($\lambda = .05\text{--}.79$, $M = .58$ and $\lambda = .03\text{--}.76$, $M = .53$, for the *three-factor model* and *unidimensional model*, respectively). As shown in Table 4, these models demonstrated low factor loadings for all four indicators measuring dysfunctional beliefs and for some of the indicators measuring Lack of Readiness. Furthermore, although the *three-factor model* demonstrated high factor loadings at the common total score level ($\lambda_s = .81, .91, .89$), it also resulted in higher factor correlations ($r = .72, .74, .81$) compared with the *original CDDQ model* and *10-factor model* in which the 10 scale-level factors were included. Because the fourth model included only one common total score, factor correlations could not be calculated. These findings undermine the discriminant validity of interpretations of items in the *three-factor model* and *unidimensional model*, which excluded the 10 scale-level factors.

Taken together, the two models that included the 10 scale-level factors of career indecision (i.e., the *original CDDQ model* and the *10-factor model*) demonstrated a better fit to the data than the *three-factor model* and *unidimensional model* that did not include these factors. The *original CDDQ model*

Table 4. Standardized Factor Loadings for the Four Tested Models (Total Sample).

	<i>Original CDDQ</i>			<i>10-Factor</i>		<i>Three-Factor</i>		<i>Unidimensional GF(λ)</i>
	<i>GF(λ)</i>	<i>CF(λ)</i>	<i>SF(λ)</i>	<i>GF(λ)</i>	<i>SF(λ)</i>	<i>GF(λ)</i>	<i>CF(λ)</i>	
LR	.96					.81		
Rm		.59		.55				
Rm1			.86		.86		.59	.49
Rm2			.40		.40		.26	.21
Rm3			.57		.57		.40	.33
Ri		.72		.68				
Ri1			.79		.79		.69	.56
Ri2			.67		.67		.58	.47
Ri3			.60		.60		.52	.43
Rd		.27		.26				
Rd1			.70		.70		.25	.22
Rd2			.32		.31		.05	.03
Rd3			.44		.44		.12	.10
Rd4			.53		.54		.19	.16
Ll	.93					.91		
Lp		.88		.86				
Lp1			.82		.82		.75	.73
Lp2			.83		.83		.76	.73
Lp3			.86		.86		.79	.76
Ls		.92		.90				
Ls1			.78		.78		.74	.72
Ls2			.77		.77		.72	.70
Ls3			.78		.78		.74	.72
Ls4			.79		.79		.74	.73
Lo		.88		.86				
Lo1			.82		.82		.74	.72
Lo2			.82		.82		.74	.71
Lo3			.74		.74		.67	.66
La		.96		.94				
La1			.78		.79		.76	.73
La2			.73		.73		.70	.68
li	.90					.89		
lu		.99		.93				
lu1			.67		.67		.68	.62
lu2			.70		.70		.70	.64
lu3			.58		.70		.68	.65
li		.96		.88				
li1			.60		.58		.57	.52
li2			.65		.60		.59	.54
li3			.63		.65		.64	.58
li4			.63		.63		.62	.56
li5			.64		.63		.62	.56

(continued)

Table 4. (continued)

	<i>Original CDDQ</i>			<i>10-Factor</i>		<i>Three-Factor</i>		<i>Unidimensional GF(λ)</i>
	<i>GF(λ)</i>	<i>CF(λ)</i>	<i>SF(λ)</i>	<i>GF(λ)</i>	<i>SF(λ)</i>	<i>GF(λ)</i>	<i>CF(λ)</i>	
<i>le</i>		.62		.57				
<i>le1</i>			.70		.69		.45	.40
<i>le2</i>			.88		.88		.56	.51

Note. GF = general factor; CF = cluster-level factor; SF = specific factor; LR = Lack of Readiness; Rm = Lack of Motivation; Ri = General Indecisiveness; Rd = Dysfunctional Beliefs; LI = Lack of Information; Lp = Lack of Information About the Process; Ls = Lack of Information About the Self; Lo = Lack of Information About Occupations; La = Lack of Information About Additional Ways of Obtaining Information; II = Inconsistent Information; lu = Unreliable Information; li = Internal Conflicts; le = External Conflicts. λ s < .40 are presented in bold.

demonstrated a slightly better fit than the *10-factor model* in this analysis. A subsequent inspection of the parameter estimates precluded determining which of the two models provided a better fit to the data. Nevertheless, the *original CDDQ model* should be preferred due to its slightly better fit, greater theoretical congruence with the original taxonomy underpinning the CDDQ, and previous findings demonstrating the validity of the three cluster-level factors included in this model.

Reliability Analyses

In the next phase of analyses, we estimated the composite reliability (CR) indices for the *original CDDQ model*, which included 10 first-order scales, three second-order clusters, and one third-order total score. For the total sample, the CR estimates of nine first-order scales reached acceptable levels (CR = .73–.88, CR_{median} = .76); however, the lack of motivation factor was marginally below the acceptable threshold (CR = .58). CRs of the three second-order clusters were good (CRs = .84, .95, and .89 for Lack of Readiness, Lack of Information, and Inconsistent Information, respectively). Finally, an estimation of the CR of the third-order total score resulted in an excellent CR estimate of .97. The CR estimates for the country samples were similar to those calculated for the total sample and are available upon request from the first author.

Measurement Invariance

Our results so far had indicated the equivalence of the CDDQ factor structure across 13 countries and nine language versions. To examine whether participants from different countries who completed the same CDDQ language version (i.e., English or French) responded to items in similar ways, we conducted a series of measurement invariance analyses. First, we tested the measurement invariance of the English version of the CDDQ across three available countries (Malaysia, South Africa, and USA). The configural model demonstrated good fit (CFI = .988, TLI = .987, RMSEA = .041, 95% CI [.039, .043], SRMR = .051), but the metric model resulted in reduced fit (CFI = .972, TLI = .971, RMSEA = .062, 95% CI [.060, .063], SRMR = .067, $\Delta CFI > .010$ and $\Delta RMSEA > .010$), indicating significant differences in the factor loadings among the three subsamples. Subsequent measurement invariance analyses between all country pairs further confirmed that the English version of the CDDQ was measurement variant across these three country samples.

Next, we tested the measurement invariance of the French version of the CDDQ across the three available countries (Canada, France, and Switzerland). The configural model demonstrated good fit (CFI = .976, TLI = .974, RMSEA = .050, 95% CI [.049, .051], SRMR = .054), and was

similar in fit to the metric model (CFI = .968, TLI = .967, RMSEA = .056, 95% CI [.055, .057], SRMR = .061, Δ CFI < .010 and Δ RMSEA < .010). The fit of the scalar model was similar (CFI = .966, TLI = .965, RMSEA = .057, 95% CI [.056, .058], SRMR = .062; Δ CFI < .010 and Δ RMSEA < .010). These findings provided further support for the *original CDDQ model* and demonstrated the measurement equivalence at the level of scalar invariance of the French version of the CDDQ across the samples from the three French countries.

Discussion

The overarching goal of this research was to determine the factor structure of the CDDQ across 13 countries and nine language versions compiled from 39 diverse data sets. Structural analyses and an evaluation of factor loadings and correlations provided support for the model corresponding to the original CDDQ taxonomy. This model, the *original CDDQ model*, includes 10 first-order scales, three second-order clusters, and one third-order total score, corresponding to 10 specific causes of career indecision grouped in three major difficulty clusters and aggregated into a global estimate of indecision. The estimation of composite reliability (CR) indices of the 14 CDDQ scores derived from the *original CDDQ model* confirmed the adequate internal reliability of 13 of its scores; however, the lack of motivation scale yielded a reliability estimate marginally below the acceptable level. The highly similar results across the 13 country samples support the cross-cultural generalizability of the CDDQ as a measure of career indecision.

This research is the most comprehensive examination of the factor structure of the CDDQ performed to date. The good fit of the *original CDDQ model* was demonstrated for the total sample and for 11 of the 13 country samples (acceptable fit was observed for the Polish and Swiss–French samples). In contrast, a *10-factor model*—which included a single total score but excluded the three second-order clusters—demonstrated good but lower fit (relative to the *original CDDQ model*) in the total sample and in six country samples (acceptable fit was observed for the remaining seven language samples). In addition, in five samples—Croatia, Israel (Arabic and Hebrew), South Africa, and South Korea—the fit of the *original CDDQ model* was better than that of the *10-factor model*. In light of these findings and given its lower theoretical compatibility with Gati et al.'s (1996) taxonomy of career decision-making difficulties, the *10-factor model* can be regarded as inferior. Then a *three-factor model*—which did not include 10 first-order scales but rather only three first-order clusters and a second-order total score—yielded good but lower fit in the total sample but in only one of the 13 country samples; an acceptable fit was demonstrated for the remaining 12 country samples. Finally, the *unidimensional model*—including only a single first-order total score—yielded only acceptable fit in the total sample and 11 country samples; it showed below acceptable fit indices in the Hebrew sample and good fit in the Turkish sample. Moreover, compared with the *original CDDQ model* and the *10-factor model*, both the *three-factor model* and the *unidimensional model* resulted in weaker factor loadings and higher factor correlations, thus further supporting their relative inferiority.

The finding that the *original CDDQ model* provides the best fit to data aligns with the conclusions of previous validation studies of its English (Kleiman et al., 2004; Kleiman & Gati, 2004; Levin et al., 2020), French (Atitsogbe et al., 2018; Rossier et al., 2021; Sovet et al., 2017), Hebrew (Kleiman & Gati, 2004), and Turkish (Bacanli, 2016) versions. At the same time, this study is the first to demonstrate the advantage of the *original CDDQ model* over alternative models and to test its internal structure in five additional language versions. Specifically, we also examined and validated the factor structure of the 34-item Arabic and Polish versions, the psychometric properties of which were not previously published. For the Croatian, Greek, and Korean versions, however, our findings are inconsistent with previous work. Babarović and Šverko's (2019) study of the Croatian version yielded a preferred model that excluded the

dysfunctional beliefs scale. Similarly, [Sovet et al.'s \(2015\)](#) study of the Korean version resulted in selecting a nonhierarchical model that included only nine first-order scales (excluding the dysfunctional beliefs scale). For the Greek CDDQ version, a seven-factor model that excluded lack of information about ways of obtaining additional information, unreliable information, and internal conflicts was preferred ([Vaipoulou et al., 2019](#)). Thus, of the nine language versions examined in this study, the *original CDDQ model* was confirmed in previous research for only four language versions.

Four possible explanations for discrepancies in the best-fitting model of a measure tested across countries using different language versions should be considered. The first explanation for structural divergence involves potential cross-cultural differences that contribute to variations in the psychological constructs under study. Cross-cultural discrepancies have been the typical explanation adopted in previous studies reporting differences in the structure of career indecision (see [Xu & Bhang, 2019](#)). However, our findings indicate high structural similarity in career indecision across 13 countries based on data collected using nine language versions of the CDDQ. These findings suggest that at least some of the previously reported structural differences in the CDDQ may be related to methodological issues and not to cross-cultural differences.

Second, structural divergence could also result from imprecise translation processes, producing semantically different items across language versions. In such cases, structural differences would result from one or more translated items measuring something different from what was intended. The observed replicability of the factor structure of the CDDQ across nine language versions in this study precludes a translation bias leading to structural differences. Nevertheless, our findings indicate that the CDDQ is metric variant among the three countries in which the English CDDQ version was used—the USA, South Africa, and Malaysia. These results show that although the same structure of career indecision is applicable in these countries, individuals from these countries respond differently to the CDDQ items. In contrast, the French CDDQ emerged as scalar invariant across Canada, France, and Switzerland. Further investigation of the measurement equivalence of the CDDQ may reveal additional response differences across groups of interest (e.g., gender, age).

Third, validating a factor structure involves analyzing samples with sufficient variance in the measured factors. Thus, potential unaccounted-for differences in sample characteristics constitute another methodology-related explanation for structural variations. [Kline \(2016\)](#) noted that to ensure sufficient variance, larger samples are needed in cases of more complex structural models, nonnormal indicators, and factors characterized by lower reliabilities (see also [Wolf et al., 2013](#)). Relatedly, though a one-size-fits-all heuristic is generally discouraged, [Kline \(2016\)](#) recommended a sample-size-to-parameters ratio of 20:1 or, less ideally, 10:1, with lower ratios diminishing the trustworthiness of the results. Thus, in the case of the *original CDDQ model* with 77 parameters, a sample size of 770 could still be considered small. Furthermore, even with sufficiently large samples, higher homogeneity in terms of factor-relevant variance could undermine attempts at structural validation. [Tien \(2005\)](#), for example, reported structural differences after examining the factor structure of the CDDQ separately among decided and undecided individuals. Similarly, [Levin et al. \(2020\)](#) tested structural differences among six groups of individuals differing in their degree of decidedness. In both studies, the variance in the target factor (e.g., degree of decidedness, career decision status) was reduced, explaining why the *original CDDQ model* was not supported in these studies. Such findings attest to the importance of considering an adequate sample size and the expected heterogeneity of target factors in designing validation studies.

Although an inadequate sample size and reduced variance in target factors may explain some of the previously reported variations in the factor structure of the CDDQ, a fourth explanation relates to the adequacy of the psychometric analyses performed. In this study, the implementation of normality analyses and the subsequent application of WLSMV yielded good fit indices for the tested models. Indeed, whereas the good fit of the *original CDDQ model* was demonstrated with WLSMV—the

accepted method when the assumptions of univariate and multivariate normality are not met (see Table 3)—with ML, this model barely resulted in an acceptable fit and in only one country sample (see Supplemental Appendix Table 2 in the online supplementary materials). Applying ML is justified when indicators are continuous and conform to a multivariate normal distribution, but when this assumption is violated, ML is likely to yield biased parameters (Beauducel & Herzberg, 2006; Li, 2016). The observed violations of the univariate and multivariate nonnormality assumptions in the 13 country samples analyzed in this study suggest that previous failures to validate the *original CDDQ model* may be attributed, at least partially, to improper implementation of a CFA—using ML rather than WLSMV to test the structure of the CDDQ.

Limitations and Recommendations for Future Research

Before discussing the implications of this study for assessing career indecision, its limitations should be acknowledged. An initial limitation is that not all language versions of the CDDQ were examined in this research. Specifically, we failed to obtain sufficiently large samples for other versions, such as Chinese, Dutch, German, Hindi, Italian, Portuguese, Russian, and Spanish. In addition, this research did not examine the factor structure of the 44-item version of the CDDQ, which remains in use in some countries (e.g., Fouad et al., 2009; Willner et al., 2015).

Second, previous research compared CDDQ scores across groups, such as age groups, genders, or countries, but mostly did so without ensuring the invariance of CDDQ scores across the groups under study (for a review, see Levin et al., 2020). Some investigations demonstrated the measurement invariance of the CDDQ across groups from the same country (Levin et al., 2020; Sovet et al., 2015, 2017) or across countries using the same CDDQ language version (Atitsogbe et al., 2018; Levin et al., 2020; Rossier et al., 2021); in this respect, our findings provide further support for the measurement invariance at the scalar level of the French CDDQ across three countries. In contrast, the English CDDQ emerged in our study as measurement variant across Malaysia, South Africa, and the USA, whereas Levin et al. (2020) found that the CDDQ is measurement invariant at the scalar level across seven countries with data collected using its English version (including South Africa and the USA). These discrepancies should discourage researchers from relying on the results of previous demonstrations of measurement invariance; when the focus of a study is comparing groups, conducting measurement invariance analyses is advisable.

A third limitation of this study concerns the scope of the models we considered and evaluated. Given strong previous support for the criterion validity of the CDDQ total score—calculated as the mean of all items—we considered only models that did not exclude any item. For this reason, we did not evaluate the fit of previously proposed models for the factor structure of the CDDQ, such as models excluding the dysfunctional beliefs scale (Babarović & Šverko, 2019; Sovet et al., 2015) or all three Lack of Readiness scales (Creed & Yin, 2006). A previous study on the English version of the CDDQ compared these additional models but provided support for preferring the *original CDDQ model* (Levin et al., 2020). At the same time, our findings indeed reflect the relatively lower psychometric quality of the three Lack of Readiness scales, with weak factor loadings associated with the dysfunctional beliefs scale and a marginally below the acceptable threshold reliability estimate for the lack of motivation scale. Future research could address these issues by designing a revised version with higher quality items to measure lack of motivation and dysfunctional beliefs. In this context, Hechtlinger et al. (2019) developed a measure of dysfunctional beliefs in career decision-making and found that the heterogeneity of dysfunctional beliefs undermines the utility of a total score.

Implications for the Assessment of Career Indecision

The results of this study support using the CDDQ at three levels: a total score, three cluster scores, and 10 scale scores. Each of these three levels has its own advantages and disadvantages for research and practice. The CDDQ total score measures individuals' overall severity or *degree* of career indecision and is especially informative for a global evaluation of the effectiveness of interventions or for the initial prescreening of clients (Fouad et al., 2009; Gati & Saka, 2001). However, deriving information about the *nature* of career indecision, namely, individuals' causes of career indecision, requires using either the three cluster scores or the 10 scale scores. Identifying individuals' specific causes of career indecision facilitates tailoring the intervention to each individual's unique needs.

Specifically, the CDDQ three cluster scores indicate the degree to which individuals' career indecision is caused by difficulties related to Lack of Readiness, Lack of Information, or Inconsistent Information. Cluster scores are especially informative for research seeking to examine the underlying mechanisms of career indecision that are affected by interventions (e.g., Fouad et al., 2009), discern the differential interplay between specific domains of career decision-making difficulties and other constructs (e.g., Udayar et al., 2020), or identify which intervention is most suitable for each client's needs. In addition, whereas the internal reliabilities of two of the three CDDQ cluster scores are high, Lack of Readiness is relatively more heterogeneous. For this reason, a clearer understanding of the specific difficulties related to Lack of Readiness should involve inspecting the three scale scores in this cluster.

The 10 CDDQ scale scores provide information on the specific causes of individuals' career indecision. Although offering the most detailed information, the complexity involved in dealing with 10 scores may explain why research at this level is relatively rare. However, because the 10 CDDQ scale scores provide the best differential assessment of individuals' causes of career indecision, using this level of assessment is encouraged. The 10 scale scores are particularly recommended for evaluating the relative salience of each difficulty to identify individuals' most pressing needs (Gati et al., 2000) or for the classification of individuals into career indecision types (Levin et al., 2022).

Conclusion

The analyses of 13 country samples using nine language versions, supported the alignment of the empirical structure of the CDDQ with Gati et al.'s (1996) taxonomy of career indecision. Based on analyses performed using the adequate WLSMV estimator, our results support the conclusion that career indecision as measured by the CDDQ can be viewed as a general construct comprising three clusters that can be further differentiated into 10 distinguishable components. Previous research investigating the criterion validity of the CDDQ scores supported their utility for research and practice. Our findings complement previous work and provide the needed structural validation for the use of all 14 originally proposed CDDQ scores: a total score, three cluster scores, and 10 scale scores.

Acknowledgment

We are grateful for the generosity of our colleagues who were willing to share their data, thereby making the study possible. The study included the analyses of previously collected secondary data by the authors (10 samples) and other research teams (29 samples). Information on the source of each sample can be found in [Supplemental Appendix 3](#) in the online supplementary materials. This research has been supported in part by the Samuel and Esther Melton Chair of Itamar Gati.

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Jérôme Rossier  <https://orcid.org/0000-0002-9924-3672>

Supplemental Material

Supplemental material for this article is available online.

References

- Atitsogbe, K. A., Moumoula, I. A., Rochat, S., Antonietti, J.-P., & Rossier, J. (2018). Vocational interests and career indecision in Switzerland and Burkina Faso: Cross-cultural similarities and differences. *Journal of Vocational Behavior, 107*, 126–140. <https://doi.org/10.1016/j.jvb.2018.04.002>.
- Babarović, T., & Šverko, I. (2019). The validity of career decision-making difficulties questionnaire in Croatia. *Journal of Career Assessment, 27*(3), 391–407. <https://doi.org/10.1177/1069072717748960>.
- Bacanli, F. (2016). Career decision-making difficulties of Turkish adolescents. *International Journal for Educational and Vocational Guidance, 16*(2), 233–250. <https://doi.org/10.1007/s10775-015-9304-8>.
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science, 16*(1), 74–94. <https://doi.org/10.1007/BF02723327>.
- Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal, 13*(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd edition.). The Guilford Press.
- Chartrand, J. M., Robbins, S. B., Morrill, W. H., & Boggs, K. (1990). Development and validation of the career factors inventory. *Journal of Counseling Psychology, 37*(4), 491–501. <https://doi.org/10.1037/0022-0167.37.4.491>.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. <https://doi.org/10.1080/10705510701301834>.
- Creed, P. A., & Yin, W. O. (2006). Reliability and validity of a Chinese version of the career decision-making difficulties questionnaire. *International Journal for Educational and Vocational Guidance, 6*(1), 47–63. <https://doi.org/10.1007/s10775-006-0003-3>.
- Fouad, N., Cotter, E. W., & Kantamneni, N. (2009). The effectiveness of a career decision-making course. *Journal of Career Assessment, 17*(3), 338–347. <https://doi.org/10.1177/1069072708330678>.
- Gati, I., Krausz, M., & Osipow, S. H. (1996). A taxonomy of difficulties in career decision making. *Journal of Counseling Psychology, 43*(4), 510–526. <https://doi.org/10.1037/0022-0167.43.4.510>.
- Gati, I., Osipow, S. H., Krausz, M., & Saka, N. (2000). Validity of the Career Decision-Making Difficulties Questionnaire: Counselee versus career counselor perceptions. *Journal of Vocational Behavior, 56*(1), 99–113. <https://doi.org/10.1006/jvbe.1999.1710>.
- Gati, I., & Saka, N. (2001). High school students' career-related decision-making difficulties. *Journal of Counseling & Development, 79*(3), 331–340. <https://doi.org/10.1002/j.1556-6676.2001.tb01978.x>.

- Hechtlinger, S., Levin, N., & Gati, I. (2019). Dysfunctional career decision-making beliefs: A multidimensional model and measure. *Journal of Career Assessment, 27*(2), 209–229. <https://doi.org/10.1177/1069072717748677>.
- Holland, J. L., Daiger, D. C., & Power, P. G. (1980). *My vocational situation*. Consulting Psychologists Press.
- Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Jones, L. K. (1989). Measuring a three-dimensional construct of career indecision among college students: A revision of the Vocational Decision Scale: The Career Decision Profile. *Journal of Counseling Psychology, 36*(4), 477–486. <https://doi.org/10.1037/0022-0167.36.4.477>.
- Kleiman, T., & Gati, I. (2004). Challenges of internet-based assessment: Measuring career decision-making difficulties. *Measurement and Evaluation in Counseling and Development, 37*(1), 41–55. <https://doi.org/10.1080/07481756.2004.11909749>.
- Kleiman, T., Gati, I., Peterson, G., Sampson, J., Reardon, R., & Lenz, J. (2004). Dysfunctional thinking and difficulties in career decision making. *Journal of Career Assessment, 12*(3), 312–331. <https://doi.org/10.1177/1069072704266673>.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th edition). The Guilford Press.
- Levin, N., Braunstein-Bercovitz, H., Lipshits-Brazilier, Y., Gati, I., & Rossier, J. (2020). Testing the structure of the Career decision-Making difficulties questionnaire across country, gender, age, and decision status. *Journal of Vocational Behavior, 116*(Part A), Article 103365. <https://doi.org/10.1016/j.jvb.2019.103365>.
- Levin, N., Lipshits-Brazilier, Y., & Gati, I. (2022). The identification and validation of five types of career indecision: A latent profile analysis of career decision-making difficulties. Advance online publication. <https://doi.org/10.1037/cou0000603>.
- Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48*(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*(3), 519–530. <https://doi.org/10.1093/biomet/57.3.519>.
- Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit evaluation. *Contemporary psychometrics* (pp. 275–340). Erlbaum.
- 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: A Multidisciplinary Journal, 11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2.
- Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J., & von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—Camouflaging misspecification with item parcels in CFA models. *Psychological Methods, 18*(3), 257–284. <https://doi.org/10.1037/a0032773>.
- Mau, W.-C. (2001). Assessing career decision-making difficulties: A cross-cultural study. *Journal of Career Assessment, 9*(4), 353–364. <https://doi.org/10.1177/106907270100900403>.
- Osipow, S. H. (1999). Assessing career indecision. *Journal of Vocational Behavior, 55*(1), 147–154. <https://doi.org/10.1006/jvbe.1999.1704>.
- Osipow, S. H., Carney, C. G., & Barak, A. (1976). A scale of educational-vocational undecidedness: A typological approach. *Journal of Vocational Behavior, 9*(2), 233–243. [https://doi.org/10.1016/0001-8791\(76\)90081-6](https://doi.org/10.1016/0001-8791(76)90081-6).

- Osipow, S. H., & Gati, I. (1998). Construct and concurrent validity of the career decision-making difficulties questionnaire. *Journal of Career Assessment*, 6(3), 347–364. <https://doi.org/10.1177/106907279800600305>.
- Peterson, G. W., Sampson, J. P., Lenz, J. G., & Reardon, R. C. (2002). A cognitive information processing approach to career problem solving and decision making. In D. Brown & L. Brooks Brooks (Eds), *Career choice and development* (pp. 423–475). Jossey-Bass.
- Raykov, T. (1998). Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement*, 22(4), 375–385. <https://doi.org/10.1177/014662169802200407>.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2). <https://doi.org/10.18637/jss.v048.i02>.
- Rossier, J., Rochat, S., Sovet, L., & Bernaud, J.-L. (2021). Validation of a French version of the Career Decision-Making Difficulties Questionnaire: Relationships with self-esteem and self-efficacy. *Journal of Career Development: Advance online publication*. <https://doi.org/10.1177/08948453211009975>.
- Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment*, 31(12), 1428–1441. <https://doi.org/10.1037/pas0000623>.
- Sovet, L., DiMillo, J., & Samson, A. (2017). Linguistic identity and career decision-making difficulties among French-speaking Canadian students living in an Anglo-dominant context. *International Journal for Educational and Vocational Guidance*, 17(3), 269–284. <https://doi.org/10.1007/s10775-016-9328-8>.
- Sovet, L., Tak, J., & Jung, S. (2015). Validation of the career decision-making difficulties questionnaire among Korean college students. *Journal of Career Assessment*, 23(4), 661–676. <https://doi.org/10.1177/1069072714553556>.
- Tien, H.-L. S. (2005). The validation of the career decision-making difficulties scale in a Chinese culture. *Journal of Career Assessment*, 13(1), 114–127. <https://doi.org/10.1177/1069072704270327>.
- Udayar, S., Levin, N., Lipshits-Braziler, Y., Rochat, S., Di Fabio, A., Gati, I., Sovet, L., & Rossier, J. (2020). Difficulties in career decision making and self-evaluations: A meta-analysis. *Journal of Career Assessment*, 28(4), 608–635. <https://doi.org/10.1177/1069072720910089>.
- Vahedi, S., Farrokhi, F., Mahdavi, A., & Moradi, S. (2012). Exploratory and confirmatory factor analysis of the career decision-making difficulties questionnaire. *Iranian Journal of Psychiatry*, 7(2), 74–81.
- Vaiopoulou, J., Papavassiliou-Alexiou, I., & Stamovlasis, D. (2019). Career decision-making difficulties and decision statuses among Greek student teachers. *Hellenic Journal of Psychology*, 16(1), 74–94.
- Willner, T., Gati, I., & Guan, Y. (2015). Career decision-making profiles and career decision-making difficulties: A cross-cultural comparison among US, Israeli, and Chinese samples. *Journal of Vocational Behavior*, 88, 143–153. <https://doi.org/10.1016/j.jvb.2015.03.007>.
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 73(6), 913–934. <https://doi.org/10.1177/0013164413495237>.
- Xu, H., & Bhang, C. H. (2019). The structure and measurement of career indecision: A critical review. *The Career Development Quarterly*, 67(1), 2–20. <https://doi.org/10.1002/cdq.12159>.
- Xu, H., & Tracey, T. J. G. (2014). The role of ambiguity tolerance in career decision making. *Journal of Vocational Behavior*, 85(1), 18–26. <https://doi.org/10.1016/j.jvb.2014.04.001>.
- Yanyun Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, 29(4), 377–392. <https://doi.org/10.1177/0734282911406668>.