Career decision self-efficacy Item Bank: A Simulation study

Ana E. Azpilicueta[1,2,4] - Marcos Cupani[1,2] - Fernanda B. Ghío[1,2] - Valeria E. Morán[3] - Sebastián J. Garrido[1,2] - Manuel Bruzzone[2]

1 Facultad de Psicología, Universidad Nacional de Córdoba (UNC), Córdoba, Argentina
2 Instituto de Investigaciones Psicológicas, IIPsi-CONICET-UNC, Córdoba, Argentina
3 Universidad Empresarial Siglo 21, Córdoba, Argentina
4 S/N. Facultad de Psicología, Universidad Nacional de Córdoba. Ciudad Universitaria, 5000 Córdoba, Argentina

Abstract

The transformation of the current world of work, determined by technological advances, affects and challenges the career decision-making process. To face these changes, the career self-management model, derived from social cognitive career theory, offers an optimal framework that allows studying the career adaptive behaviors that favor the preparation for career decision-making. In order to obtain a measure for the core construct, career decision-making self-efficacy, we elaborated an item bank from the selection of reagents from two widely scales used in the field for its estimation. The item analysis was performed from the item response theory. The obtained results evidenced the unidimensionality of the item pool, with difficulty indexes and ability levels covering most of the measured continuum. As regard reliability indexes (persons and items), we observed that the allocation of persons and items might be reproduced in a predictable way. The global fit of items resulted optimum for most of them. We also performed two analyzes on the evidence of the validity of the scores obtained through the Career Decision Making Self-Efficacy-Item Bank. We carried out a simulation study, which allowed to determine the real level of ability of the subjects, providing information on the quality of the item bank constructed to accurately assess career decision

self-efficacy. Thus, we carried out a concurrent validity evidence study to assess test-criterion relationships between career decision self-efficacy and the more significant outcome variables of the career decision-making process: career decision, career indecision, and decisional anxiety. In summary, we developed, calibrated, and validate the first item bank to estimate the central construct of the career decision-making process, which is the first step to develop a Computerized Adaptive Test. Based on this technology, personalized evaluations of specific skills or traits, such as career decision self-efficacy can be obtained and incorporated into computer-assisted career guidance systems.

Key words: career decision-making self-efficacy; item bank; item response theory; simulation study

Career Decision Self-Efficacy Item Bank: A Simulation Study

In the past decades, much research has focused on the study and comprehension of career decision self-efficacy, which refers to beliefs about the people's own capacity to perform adequately activities such as career planning, accurate self-appraisal, acquiring problem-solving abilities, gathering of relevant occupational information, and selecting appropriate goals (Taylor & Betz, 1983).

The social cognitive career theory (SCCT; Lent et al., 1994) has recently proposed the model of career self-management (CSM; Lent & Brown, 2013). The CSM model aims at explaining the processes through which the cognitive variables such as self-efficacy, outcome expectations, goals, personality traits and context variables, influence the development of adaptive career behaviors (e.g., Planning, searching for occupational/educational information, setting of goals, preparation for change, negotiation of transitions, etc.). These adaptive career behaviors promote certain career attainments, such as an adequate level of career decision when completing a career decision-making process (Lent & Brown, 2013).

As a central construct of the CSM model, career decision self-efficacy allows a greater

comprehension of career decision-making, therefore it is really necessary to have instruments that allow its accurate assessment. With reference to the assessment of career decision self-efficacy in high school students, a review of the literature indicated that the Middle School Self-Efficacy Scale (MSSE; Fouad et al., 1997) is one of the most widely used instruments to assess career decision-making self-efficacy (e.g., Ojeda et al., 2012; Ollea & Fouad, 2015; Summers & Falco, 2018) while the Career Exploration and Decision Self-Efficacy Brief Decision Scale (CEDSE; Lent et al., 2016) is a recent measure of career decision self-efficacy developed by the authors of the CSM model. Both scales were developed on the basis of the model proposed by Crites (1978); therefore, they are theoretically compatible, measure the same construct, and use the same response scale.

Although the international literature provides a great variety of instruments to achieve this purpose, no versions of instruments to assess career decision self-efficacy have been found in Spanish, the native language of more than 400 million people worldwide. Thus, cross-cultural studies show that career decision-making, as well as its central variables, are influenced by cultural differences, so the mere translation of the items of a particular instrument can lead to inaccurate results (Brislin, 2000; Lindley, 2006).

Regarding the scales included in this study, it has not been reported that translations have been made into other languages nor adaptations to different cultural and geographic contexts. On the other hand, the aforementioned scales do not have evidence to support their psychometric properties with a rigorous model such as those derived from the item response theory (IRT).

To fill this gap in career assessment literature, in the present study we will adapt and validate the MSSE and CEDSE scales using the methodological approach proposed by the IRT, and the Rasch Analysis in particular. This approach allows to develop and calibrate an item bank (IB), that is, a set of items that measure a latent variable unidimensionally by considering the different dimensions that integrate this variable (Wright & Bell, 1984). The IB also allows to

perform personalized assessments, which is the first step for future developments such as Computerized Adaptive Test (CAT). This standardized tool combines the advancement of technology and innovation in various areas of psychology. Besides, it allows adapting the test to clients, generating a more precise construct estimates, reducing costs and administration time (Barrada et al., 2009; Wainer, 2000). The CAT presents a set of items to the client through a computer. The CAT selects the items according to the client's response. If the client fails, the CAT provides a lower difficulty item; if they respond successfully, the next one will be more difficult. Numerous studies indicate that CATs reduce the number of items and administration time by up to 50%, obtaining measures with a higher level of precision with a smaller number of items (Wainer, 2000).

These advances in the field of psychometrics are possible because the IRT provides a comprehensive and detailed methodology that allows the assessment of psychometric properties of an instrument at item level (Messick, 1994). In particular, Rasch analysis provides more information about the ability/trait of a person because it focuses on the difficulty of the items, rather than the number of items that each participant answers correctly (Linacre, 2002). To analyze the items, the Rasch model first converts the ordinal data from an instrument into interval data, thus fulfilling one of the prerequisites of any measurement (Wright & Linacre, 1994). Then, this psychometric model enables the evaluation of several characteristics, such as the model fit level, the item difficulty and hierarchy, the reliability of persons and item, and the differential item functioning (DIF).

In sum, based on this technology, personalized evaluations of specific skills or traits, such as career decision self-efficacy, can be obtained. As proposed by Ware et al. (2000), the development of IBs generally starts from established questionnaires, which is advantageous since the content, construct validity and the quality of the items have been previously evaluated.

One of the strategies used in the construction of CATs is the simulation study at the different stages of its development: planning, construction of the set of items, quality control, and development of the algorithm (Han, 2018). This procedure makes it possible to get closer to the data more quickly and economically than a real CATs application. Furthermore, it permits to determine the real level of ability of the subjects, obtaining data with a lower level of bias and typical error (Olea et al., 1999). In summary, this procedure provides information on the quality of the IB constructed.

In the current context, the ability to perform appropriate choices in an effective way requires that people count on precise, easy, and adapted to particular contexts measures. Therefore, the main purpose of the present study was to develop and calibrate an IB to estimate career decision self-efficacy and to estimate its quality to precisely assess this construct through a simulation study.

Since evidence of the validity of a given interpretation of test scores for a specified use is a necessary condition for the justifiable use of the test (AERA, APA, & NCME, 2014), we carried out a concurrent validity evidence study to assess test-criterion relationships between career decision self-efficacy and the more significant outcome variables of the career decision-making process: career decision, career indecision, and decisional anxiety.

Method

Participants

For the selection of participants, we used a non-random accidental sampling technique (Kumar & Phrommathed, 2005). Participants were students from the last two years of middle school (5^{th} and 6^{th} year). The sample consisted of 658 adolescents, ethnically homogeneous, 365 (55.5 %) female and 292 (44.4%) male, age range 15-19 years (M= 16,56; SD= .754) from public (32.8%) and private (67.2%) schools in the city of Córdoba, Argentina. The four areas of specialization available in the middle school education system were considered: art (9.7%), social sciences (25.7%), natural sciences (40.3%), and social communication (24.3%). Data were

collected in paper-and-pencil format. Due to the characteristics of the institutions participating in this study, in which students' parents have specialized, professional, and informal jobs, and according to the classification provided by The National Institute of Statistics and Censuses (INDEC for its acronym in Spanish, 2021), the sample was considered representative of the medium-low and medium-high socioeconomic classes from Argentina.

Measures

Career Decision Self-Efficacy

Middle School Self-Efficacy Scale (MSSE; Fouad et al., 1997). This instrument assesses the career decision-making process through 47 items organized into four subscales; one of them is the career decision self-efficacy scale, consisting of 12 items with internal consistency of .97 which is responded using a five-point Likert-type scale varying from 1 (*No confidence at all*) to 5 (*Complete confidence*). It has been found that the scorings of this subscale relate positively with vocational identity, self-esteem, and career decision-making process (Walsh & Betz, 1995), and predict significantly the outcomes expectations and actions related to career exploration (Sawitri et al., 2015).

Career Exploration and Decision Self-Efficacy Brief Decision Scale (CEDSE; Lent et al., 2016). This scale consists of 12 items distributed into two factors that assess career decision self-efficacy (Brief Decisional Self-Efficacy Factor) and coping efficacy. The analyses were performed with the first factor, which assesses career decision self-efficacy. This factor consists of eight items that measure the students' degree of reliability as regard certain activities related to the process of vocational choice; it presents an adequate level of internal consistency ($\alpha = .96$) and is responded using a five-point Likert-type scale, which varies from 1 (*No confidence at all*) to 5 (*Complete confidence*).

Career Decision

Career decision was assessed with the Career Decision Questionnaire (CDQ), belonging to the Career Indecision Profile (CIP-65; Hacker et al., 2013), which allows estimating the level

of self-perceived career decision using four items (Item 2: “How decided about your career direction are you at this time?”) with a Likert-type response scale that ranges from 1 (*Completely undecided*) to 6 (*Very decided).* This measure was selected since the use of short scales is currently recommended to estimate this construct (Lent et al., 2017). The reliability index for this sample was .71.

Career Indecision

Career indecision was estimate with the Career Decision Scale (CDS; Osipow et al., 1987), a 19 items instrument. The first two items indicate the participants’ level of decision in career choice, the next 16 items measure indecision in career choice (Item 3: “Several careers have equal appeal to me. I´m having a difficult time deciding among them”) and the last item is an open question for participants to describe they vocational situation if any of the previous items does not describe it. All the items, excluding the last one, are responded with a scale from 1 (*Not at all like me*) to 4 (*Exactly like me*). The authors reported adequate psychometric properties (α = .81; 18 items). In the validate version, evidence of internal structure was obtained through EFA and predictive validity. The reliability index for this sample was .80 (Azpilicueta et al., 2019).

Decisional Anxiety

We used the Subscale Choice/ Commitment Anxiety (CCA) from the Career Indecision Profile (CIP-65; Hacker et al., 2013) to assess decisional anxiety. The CIP-65 consists of 65 items that allow assessing four aspects that influence the process of career choice: (1) neuroticism/ negative affect (NNA), (2) lack of readiness (LR), (3) interpersonal conflicts (IC), and (4) choice/commitment anxiety (CCA). In the present study, the last subscale was used to determine the level of decisional anxiety through its 10 items (Item 5: “I’m concerned that my interests may change after I decide on a career”). High scores in the CCA scale imply that the person needs to obtain more occupational information and information about him/herself; they also show an incapacity to commit and a high level of anxiety about career decision-

making. Participants must respond to a Likert-type scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). Strong internal consistency reliability estimates have been reported for the CCA (α = .97). Hacker et al. (2013) found evidence to support the factor structure of the CCA, which produced the largest correlation with a measure of career decidedness (r = .71). The validation studies of this scale indicated that the 10 items presented adequate reliability values (.86) and evidence of internal structure validity (EFA) and predictive validity (Azpilicueta et al., 2019).

Procedure

The administration of instruments was collective, in regular class schedules, with authorization and prior consent from the principals of the schools and the teachers of each class, asking for the collaboration of each student and emphasizing the voluntary nature of their participation and the confidentiality of data. The three scales were administered jointly in a 40-minute interval. Although it would have been convenient to carry out the data collection through an electronic format, we opt for a pencil/paper format since most of the schools in our country are not equipped with enough computers to perform the collection electronically. Ethical guidelines for research with humans recommended by the American Psychological Association (APA, 2002) were respected; informed consent forms were used, and measures were taken to guarantee respect for human rights and care of the environment. Besides, there was a strict control to avoid any emergent risk to guarantee the proper use and handling of information. Researchers of this project declared to know and perform the safeguards provided by the Declaration of Helsinki and Law N° 25,326 of Protection of Personal Data.

Data Analysis

Translation of the items

In a first stage, the scales were translated from English into Spanish by three specialists in the English language (direct translation). Then, these versions were compared, and some pertinent idiomatic adjustments were performed trying to keep certain conceptual, semantic,

and functional equivalences, instead of a mere literal translation (Mimura & Griffiths, 2008). This procedure was complemented with a series of cognitive interviews to adolescents with similar characteristics to the target population. Then, it was performed an expert revision of the items to select those that, because of their content and language, assessed more precisely the proposed construct in the target population.

The items referring to a choice of a major in College were not included because these items are neither meant for middle school level students nor correspond with the University Educational System from our country (e.g., “Select one major from a list you are considering”).

Development of the Career Decision-Making Self-Efficacy Item Bank (CDMSE-IB)

The development of an IB is a multi-stage process. Literature suggests performing first a clear definition of the construct to be measured, and then developing or selecting the most representative items for that latent trait. In this study, the development of the CDMSE-IB was based on two widely used instruments to estimate career decision self-efficacy (Choi et al., 2012). The advantage of this approach is that both content validity and item quality have been previously assessed (Bjorner et al., 2007).

Calibration of the CDMSE-IB

The Rasch analysis were carried out with the Winsteps software (Linacre, 2016), which ensures that all the parameters of the persons (θ) and the items (δ) are point locations in a single latent variable, which can be expressed in the same unit of scale (logit) and that make it possible to establish objective comparisons. The CDMSE-IB calibration consisted of the following steps:

Unidimensionality. This feature was evaluated by using Principal Components Analysis of Rasch Residuals (PCAR). The unidimensionality assumption is achieved if: (a) the measurement model explains approximately the 50 % of the variance, and (b) the largest

additional factor (a secondary dimension) has an eigenvalue of less than 3 (a force of three items) and accounts for less than 5% of the unexplained variance (Linacre, 2016). However, the possibility that a single factor fully explains the total variance of the scores is very complex (Muñiz Fernández, 1997). In this sense, Muñiz Fernández (1997) argues that unidimensionality is a gradual concept, that is, the more variance the first factor explains, the more unidimensionality exists. Therefore, a percentage less than 50% does not necessarily indicate the absence of unidimensionality.

Rasch Model Fit. Three analyzes were carried out: (1) the global fit of the data, (2) the fit of the items, and (3) the fit of the people. The global fit of the data allows to verify if, the data matrix fits the predictions of the model. The items fit was analyzed by studying each one independently. The persons fit was analyzed in order to identify the participants who responded unexpectedly and not adequate to the theoretical formulation.

We used the statistical fit indexes Infit (internal fit) and Outfit (external fit). The Infit index is calculated from the unstandardized quadratic means, which allows identifying unexpected behaviors that affect the items that, in the measurement continuum, are close to the level of trait that a person possesses. The Outfit index is the weighted root mean of residuals resulting from persons and items; it allows to evaluate the unexpected behavior of the items that have a difficulty far from the level of latent trait that each person presents (Bond & Fox, 2015). Values provided by the Rasch Model are expressed in logit scale, which is a logistic transformation of the observed scores, with a mean of 0 and standard deviation of 1. When the observed data coincide with those proposed by the model, the values of Infit and Outfit indexes have values close to 1; otherwise, we will obtain values far from 1. That is, a value of Infit of 1 indicates that 100% of the variance of the empirical data are explained by the model, while a value of 1.5 indicates that a 50% of the variance cannot be explained by the model. Following the criteria proposed by Linacre (2002), the region to consider an acceptable adjustment range between .5 and 1.5.

We also calculated the point-biserial correlation discrimination estimates (r_{bp}), which allows to diagnose errors in the coding of items or keys (negative values or cero indicate items or people with response patterns that contradict the variable).

Separation and Reliability. The item separation index indicates the distance between the levels of difficulty or trait, which should be sufficient to identify the meaning of the latent variable (Wright & Stone, 2003). The person's separation index indicates the aptitude of the instrument to discriminate people in the measured variable. A useful set of items must define at least three strata of people (e.g., high, moderate and low levels). An adequate level of separation should be greater than 2, associated to a reliability around 0.80 (Bond & Fox, 2015). To assess the position of items and persons in the continuous, we analyzed the map of items and persons simultaneously.

Differential Item Functioning (DIF*).* We studied the DIF in terms of school year of attendance (5^{th} vs. 6^{th} grade) and the participants' sex. An item is considered to have DIF when the probability of a correct response does not depend solely on the level of the person in the ability/trait intentionally measured by the test (Bond & Fox, 2015). Item severity measures δi were computed for each school year of attendance.

Simulation Study

We carried out a simulation design generated with the R software (R Core Team, 2017) to determine the number of items needed to accurately estimate the career decision self-efficacy level of the examinee. This design consisted of estimating a process of simulating responses to the adaptive test algorithm implemented in the CDMSE-IB for each skill level a thousand times.

The simulation of the adaptive test algorithm responses for a skill level of interest j consists of the following process: (a) selection of the first item, (b) answer 1 is generated randomly, using a Bernoulli distribution with 1 = 1 + - 1.702, (c) selection of a new item

considering answer 1, (d) answer 2 is generated randomly, with the same process used for 1, (e) the two previous steps are repeated, until the adaptive algorithm termination condition.

After estimating a thousand simulation processes for a skill level of interest, we obtained the following: a) a vector with a thousand final estimates of ability using the adaptive test algorithm and, b) a vector with a thousand values that indicate the number of items used for estimating those skills. The skill levels considered were: -2, -1.5, -1, 0, 1, 1.5, and 2. The algorithm completion condition was to reach a standard error of .5. If the condition did not meet 20 items, the algorithm estimated the ability of the responses obtained. The first item was closest to skill level 0.

Concurrent Validity Evidence

Evidence based on relationships with other variables inform about the degree to which these relationships are consistent with the construct underlying the proposed test score interpretations. A concurrent study obtains test scores and criterion information at about the same time. This type of evidence, which avoids temporal changes, is particularly useful for psychodiagnostic tests or in investigating alternative measures of some specified construct for which an accepted measurement procedure already exists. This study was carried out carried out through Pearson´s correlation analysis with the more significant outcome variables of the career decision-making process: career decision, career indecision, and decisional anxiety.

Regarding the correlation indices, it is expected to obtain only moderate correlations between a criterion and any test. Since people and the situations in which the criterion data are obtained are too complex, it is challenging, from test results, to estimate an exact prediction of their performance (Nunnally, 1991).

Results

Calibration of the Item Pool

Unidimensionality

It was examined through the PCAR. The results indicate that the unidimensionality assumption is fulfilled since it was found that the Rasch dimension explained 46.2% of the variance of the data with an eigenvalue of 16.8. The first contrast (the largest secondary dimension) had an eigenvalue of 2.1 and represented 5.7% of the unexplained variance.

Rasch Model Fit

Three analyses were performed: data global fit, item fit, and person fit. Centered in the item analysis (see Table 1), fit has resulted satisfactory for 18 out of the 20 items. Outfit values obtained for item 14 ("Resist attempts of parents or friends to push me into a career I believe is beyond my abilities or not for me") and item 16 ("Choose a career in which most workers are the opposite sex") evidence that these items do not fit; that is, it presents a behavior slightly predictable by the model (Linacre, 2002). The difficulty measure (δi) of the items varied between $.96 \leq \delta i \leq -.79$, with a media of 0.00 (*SD* = .38). Infit item values varied between 2.23 and 1.01 with a media of 0 (*SD* = .38), and Oufit indexes varied between 2.33 and .71, with a media of 1.03 (*SD* = .39). The fit analysis of persons reflects that 85 % of response patterns fit to the model (Infit and Outfit ≤ 1.5). The ability levels varied between $3.38 \leq \theta \leq -2.87$, with a media of .68 (*SD* = .67). None items showed r_{bp} values that were negative or close to zero.

[Table 1]

In the Wright Map of persons and items in Figure 1, it can be jointly seen the contiguous allocation of persons and items. The distribution of the ability levels of the persons in the present study is presented on the left-hand side, whereas the difficulty of items is presented on the right-hand side. It can be observed that most of the items are in a centered position as regard the assessed students, and that they are adequately distributed in the continuum. However, it would be recommendable to add some items to cover the superior area (highest difficulty). On the other hand, some of the items that measure goal selection (item 3 and 6) are similar as regard difficulty (δ between .01 and .02). Likewise, items 12 and 13 (self-appraisal) present similar levels of difficulty (between -.08 and -.09). As regard the

content measured by the items, most of them estimate contents in relation with self-appraisal (six items) and occupational information (six items), requiring a lower level of ability, whereas the items that assess the contents planning (two items) and problem solving (two items) are relatively scarce and implies a higher level of resolution.

[Figure 1]

Separation and Reliability

Item separation (9.16) and item reliability (0.99) indexes were satisfactory, which indicates that the sample used is big enough to confirm the item difficulty hierarchy (construct validity) of the instrument (Linacre, 2016). On the other hand, person separation indexes (2.42) and person reliability indexes (0.85) were considered acceptable.

Differential Item Functioning (DIF)

These analyses were performed according to the year of attendance and the participants' sex. An item is considered to have DIF when the probability of a correct response not only depends on the level attained by the person in the trait measured by the test. To apply DIF, pair-wise analyses were performed with a significance level fixed in $\alpha < .01$, and considering that the DIF contrast must be higher than $\geq .50$ logits (Linacre, 2016). For this analysis, Winsteps uses Welch's *t*-test, obtained when dividing the DIF contrasts by the set standard error of DIF measures (Linacre, 2016). The DIF contrast is the difference between DIF sizes and log-odds estimates.

Results of DIF analysis as regard year of attendance and participants' sex indicate that no item presented a statistically significant difference, which implies that the response given by persons to items do not vary according to the year they are attending (5^{th} or 6^{th} year) or their sex.

Simulation Study

Simulation of 1000 cases was performed to determine the number of items required to accurately estimate career decision self-efficacy, with an error level of .5. The results obtained

in this analysis are presented in Table 2. In the range of -2 to 2 most of the simulations estimated the ability of the subjects with less than 20 items. The averages of items used to estimate the subjects' ability, in the skills range of -1.5 to 1.5, were 19-20 items, that is, almost the same number of items that made up the CDMSE-IB. Thus, the number of items necessary to estimate the value of $\theta = 0.00$ was 17-18 items.

Based on these results, we obtain validity evidence of the quality of the CDMSE-IB to accurately estimate career decision self-efficacy.

[Table 2]

Concurrent Validity Evidence Study

The Pearson's correlations between CDMSE-IB and three career decision-making scales (CDQ, CDS, and CCA) were calculated to explore the concurrent validity evidence of the CDMSE-IB. As documented in Table 2, the Pearson's correlation between CDMSE-IB and CDQ was .50 while the Pearson's correlations with CDS and CCA were -.32 and -.34, respectively. These results indicated that the CDMSE-IB can predict the main outcomes of the career decision-making process.

[Table 3]

Discussion

The career decision self-efficacy construct results crucial in current career choice models (Lent et. al., 2016). Due to the absence of measurement instruments that allow estimating this variable in the Spanish-speaking population, we aimed in the present study at developing an IB to assess it. For item calibration, we used the Rasch model because it guarantees certain conditions, such as the expression in the same units of measure of the parameters of both persons and items, the independence of persons with respect to items, and the scale interval properties.

Overall, the reagents constituting the IB for the career decision self-efficacy assessment presented acceptable psychometric properties. Specifically, a unidimensional structure was evidenced through PCAR analysis, which agrees with previous studies (Peterson & Delmas, 2001). Besides, the model predicted 85% of the participants' response patterns. As regard this, we identified high alignment between item difficulty and the presence of the latent trait in the individuals (Linacre, 2002). Difficulty indexes and participants' ability levels covered most of the measured continuum although, to achieve better coverage of the test and improve the construct validity, it is recommended to include a higher number of reagents to assess higher ability levels. As regard reliability indexes (persons and items), it is observed that the allocation of persons and items could be reproduced in a predictable way (Andrich, 2002). However, DIF according to the participants' sex and year of attendance was not observed.

Concerning the item fit to the model, all the items fit satisfactorily except for item 14 ("Resist attempts of parents or friends to push me into a career I believe is beyond my abilities or not for me") and 16 ("Choose a career in which most workers are the opposite sex"). Regarding item 14, a possible explanation of why adolescents considered this item as challenging could be that the predominant parental style in our culture is determined by low levels of parental education, poorly qualified work, lack of access to jobs and services, isolation, among others. These living conditions and the associated psychological stress, often weaken the ability of parents (Kaiser & Delaney, 1996; Sturge-Apple et al., 2006), making them less accepting and controlling, and more negligent, so they adopt less efficient coping strategies, developing greater emotional conflict with their children (Richaud et al., 2013). Regarding to item 16, many participants asked questions about how to response the item during the administration since they did not consider the content evaluated by that item was relevant. This can be better understood if we consider the perception of greater gender equality in the workforce (Hakim, 2006) and substantial progress in gender-equality laws of the last 50 years in Western countries (Janssen & Backes-Gellner, 2016).

Regarding the quality of the items of the CDMSE-IB to measure career decision self-efficacy, the simulation analysis provided evidence that the instrument accurately estimates this construct. We also obtained evidence regarding the predictive power of the CDMSE-IB regarding the main outcomes of the career decision-making process: career decidedness, career indecision, and decisional anxiety.

In general, the results of this study are satisfactory in terms that they allow submitting the items of this subset to a more rigorous psychometric model. However, certain constraints should be considered. Although satisfactory results have been observed when submitting the items of the classical measurement scales of career decision self-efficacy to a more rigorous psychometric model, some items present a similar or parallel content because these scales have been developed on the basis of the classical test theory, which requires as a necessary condition writing items that are reproductions of others.

The results of this work have both methodological and practical implications. The methodological implications are related to the use of the Rasch analysis as quality control to analyze, assess, and validate the measurement instruments. That is to say, from this approach the construct validity can be analyzed because the items constituting the instrument should be distributed increasingly according to difficulty and be adequately represented by the content of interest. The absence of items corresponding to each ability level is an indicator of the necessity of counting with a higher number of items to achieve a better coverage of the test. Rasch model can also provide a detailed analysis of the response patterns of persons, which allows determining not only the self-efficacy beliefs that each person really has, but also what kind of beliefs the assessed individual should develop or strengthen (Long et al., 2011). On the other hand, because the unidimensionality assumption is fulfilled and the fit model is satisfactory, we provide evidence that the items measure the construct of interest and that the persons have the abilities measured by the instrument.

As regard the practical implications, the construction of an IB allows the development and administration of CAT tests, which are characterized by minimizing the measurement standard error and the possibility of obtaining length measurement without losing precision or reliability (Abad et al., 2010). Thus, IBs allow saving time in the diagnosis process because with one subset of items all the levels of the variable of interest can be estimated, obtaining results comparable in terms of scoring and validity. On the other hand, the possibility of developing parallel tests should be considered because it entails an advantage both in research and in professional interventions; the effect of learning can be eliminated in designs in which the level of attributes is measured before and after the intervention through two versions of the same tests.

On the other hand, the development of a CAT test from the calibration of this IB will allow its incorporation into a CAGC system. According to several authors (Pérez et al., 2005; Sampson, 1999), the justification for the use of these systems is that they are highly effective in performing routine guidance tasks, such as elaboration, administration and evaluation of career profiles; supply of labor, educational and occupational information; and matching personal characteristics with educational/ occupations options. With these new technologies, counselors will be able to carry out more complex interventions (e.g., integrate computerized resources with other guidance resources, help their clients face obstacles in career decision-making, among other advantages). As stated by Nota et al. (2016), these career guidance systems are highly suitable in the present and future context, since they make it possible to capitalize on advances in artificial intelligence and the growing amount of data available on careers. In addition, one of the main advantages of the CACG systems is the potential to reach and involve a much larger and diverse group of people, who could benefit from these services, and who normally do not receive traditional career guidance. Also, different authors (e.g., Rainie, 2010; Reile & Bowlsbey, 2000) suggest that technological innovations offer the possibility of involving younger generations in activities related to career decision-making since

they contemplate different learning styles. In addition, CACG systems can be designed to self-regulate and adapt to the needs of a particular client. That is, students can advance at their own pace through the different evaluation instances (Reile & Bowlsbey, 2000).

In summary, we obtained an IB with adequate psychometric properties to assess career decision self-efficacy. Thus, we spotlight once more the relevance of estimating the psychometric properties of instruments through Rasch analysis to attain higher quality control in the processes of analysis, assessment, and validation of tests (Cupani et al., 2016).

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Abad, F. J., Olea, J., Aguado, D., Ponsoda, V., & Barrada, J. R. (2010). Deterioro de parámetros de los ítems en tests adaptativos informatizados: estudio con eCAT [Deterioration of item parameters in computerized adaptive tests: study with eCAT]. *Psicothema*, 22(2), 340-347. http://goo.gl/fAjIj6

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, 57, 1060–1073. http://www.apa.org/ethics/code2002.html.

American Educational Research Association, American Psychological Association, Joint Committee on Standards for Educational, Psychological, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.

Andrich, D. (2002). Implications and applications of modern test theory in the context of outcomes-based education. *Studies in Educational Evaluation*, 28(2), 103-121. https://doi.org/10.1016/S0191-491X(02)00015-9

Azpilicueta, A. E., Cupani, M., Ghío, B., Morán, V. E., & Garrido, S. J. (2019). Adaptación mediante el modelo de Rasch de tres medidas para estimar la decisión e indecisión de carrera y la ansiedad decisional [Adaptation of three measures to estimate career decision, career indecision and decisional anxiety using the Rasch model]. *Revista Perspectivas en Psicología*, 16(1), 26-37. http://rpsico.mdp.edu.ar/handle/123456789/1108

Barrada, J. R., Olea, J., Ponsoda, V., & Abad, F. J. (2009). Item selection rules in computerized adaptive testing: Accuracy and security. *Methodology*, 5(1), 7-17. https://doi.org/10.1027/1614-2241.5.1.7

Bjorner, J. B., Chang, C. H., Thissen, D., & Reeve, B. B. (2007). Developing tailored instruments: item banking and computerized adaptive assessment. *Quality of Life Research*, 16(1), 95-108. https://doi.org/10.1007/s11136-007-9168-6

Bond, T., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*. Routledge.

Brislin, R. (2000). *Understanding culture's influence on behavior* (2nd ed.). Nelson Thomson Learning.

*Choi, B. Y., Park, H., Yang, E., Lee, S. K., Lee, Y., & Lee, S. M. (2012). Understanding career decision self-efficacy: A meta-analytic approach. *Journal of Career Development*, 39(5), 443-460. https://doi.org/10.1177/0894845311398042

Crites, J. O. (1978). *Career maturity inventory*. CTB/McGraw-Hill.

Cupani, M., Ghio, F. B., Leal, M. F., Giraudo, G. M., Castro Zamparella, T., Piumatti, G., Barrionuevo, L. (2016). Desarrollo de un banco de ítems para medir conocimiento en estudiantes universitarios [Development of an Item Bank to Measure Knowledge in University Students] . *Revista de Psicología*, 25(2), 1-18. http://dx.doi.org/10.5354/0719-0581.2017.44808

Fouad, N. A., Smith, P. L., & Enochs, L. (1997). Reliability and validity evidence for the middle school self-efficacy scale. *Measurement and Evaluation in Counseling and Development*, 30,1, 17. https://doi.org/10.1080/07481756.1997.12068914

Han, K. C. T. (2018). Conducting simulation studies for computerized adaptive testing using SimulCAT: an instructional piece. *Journal of Educational Evaluation for Health Professions*, 15(20), 1-11. https://doi.org/10.3352/jeehp.2018.15.20

Hacker, J., Carr, A., Abrams, M., & Brown, S. D. (2013). Development of the Career Indecision Profile: Factor structure, reliability, and validity. *Journal of Career Assessment,* 21(1), 32–41. https://doi.org/10.1177/1069072712453832.

Hakim, C. (2006). Women, careers, and work-life preferences. *British Journal of Guidance & Counselling*, 34(3), 279-294. https://doi.org/10.1080/03069880600769118

Janssen, S., & Backes-Gellner, U. (2016). Occupational stereotypes and gender-specific job satisfaction. *Industrial Relations: A Journal of Economy and Society*, 55(1), 71-91. https://doi.org/10.1111/irel.12126

Kaiser, A., & Delaney, E. (1996). The effects of poverty on parenting young children. *Peabody Journal of Education*, 71(4), 66-85. https://doi.org/10.1080/01619569609595129

Kumar, S., & Phrommathed, P. (2005). Research methodology. In Kumar, S., & Phrommathed, P. (Eds.), *New Product Development* (43-50). Springer. https://doi.org/10.1007/0-387-23273-7_3

Instituto Nacional de Estadística y Censos. (2021). *Encuesta Permanente de Hogares. Mercado de trabajo. Tasas e indicadores socioeconómicos [Permanent Household Survey. Work market. Socio-economic rates and indicators]*. https://www.indec.gob.ar/uploads/informesdeprensa/mercado_trabajo_eph_4trim20126C4AD8D8.pdf

Lent, R. W., & Brown, S. D. (2013). Social cognitive model of career self-management: Toward a unifying view of adaptive career behavior across the life span. *Journal of Counseling Psychology*, 60(4), 557-568. https://doi.org/10.1037/a0033446.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45(1), 79-122. https://doi.org/10.1006/jvbe.1994.1027.

Lent, R. W., Ezeofor I., Morrison, M. A, Penn, L. T. & Ireland, G. W. (2016). Applying the Social Cognitive Model of Career Self-Management to Career Exploration and Decision-Making. *Journal of Vocational Behavior*, 93, 47-57. https://doi.org/10.1016/j.jvb.2015.12.007

Lent, R. W., Ireland, G. W., Penn, L. T., Morris, T. R., & Sappington, R. (2017). Sources of self-efficacy and outcome expectations for career exploration and decision-making: A test of the social cognitive model of career self-management. *Journal of Vocational Behavior*, 99, 107-117. https://doi.org/10.1016/j.jvb.2017.01.002

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.424.2811&rep=rep1&type=pdf.

Linacre, J. M. (2016). *Winsteps* (Version 3.92.0) [Computer Software]. Winsteps. http://www.winsteps.com/

Lindley, L. D. (2006). The paradox of self-efficacy: Research with diverse populations. *Journal of Career and Evaluation in Counseling and Development*, 26, 137-142. https://doi.org/10.1177/1069072705281371.

Long, C., Wendt, H., & Dunne, T. (2011). Applying Rasch measurement in mathematics education research: Steps towards a triangulated investigation into proficiency in the multiplicative conceptual field. *Educational Research and Evaluation*, 17(5), 387-407. http://dx.doi.org/10.1080/13803611.2011.632661

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23. https://doi.org/10.3102/0013189X023002013

Mimura, C., & Griffiths, P. (2008). A Japanese version of the Perceived Stress Scale: cross-cultural translation and equivalence assessment. *BMC psychiatry*, *8*(1), 85. https://doi.org/10.1186/1471-244X-8-85

Muñiz Fernández, J. (1997). *Introducción a la teoría de respuesta a los ítems [Introduction to item response theory]*. Pirámide.

Nota, L., Santilli, S., & Soresi, S. (2016). A Life-Design-Based Online Career Intervention for Early Adolescents: Description and Initial Analysis. *The Career Development Quarterly*, 64(1), 4-19. https://doi.org/10.1002/cdq.12037

Nunnally, J. (1991). *Teoría psicométrica [Psychometric theory]* (2nd ed.). Paidós.

Ojeda, L., Piña-Watson, B., Castillo, L. G., Castillo, R., Khan, N., & Leigh, J. (2012). Acculturation, enculturation, ethnic identity, and conscientiousness as predictors of Latino boys' and girls' career decision self-efficacy. *Journal of Career Development*, 39(2), 208-228. https://doi.org/10.1177/0894845311405321

Olea, J., Ponsoda, V., & Prieto, G. (1999). *Tests informatizados: fundamentos y aplicaciones [Computerized tests: fundamentals and applications]*. Pirámide.

Olle, C. D., & Fouad, N. A. (2015). Parental support, critical consciousness, and agency in career decision making for urban students. *Journal of Career Assessment*, 23(4), 533-544. https://doi.org/10.1177/1069072714553074

Osipow, S. H., Carney, C., Winer, J., Yanico, B., & Koschier, M. (1987). *The Career Decision Scale* (3rd ed.). Psychological Assessment Resources.

Peterson, S. L., & Delmas, R. C. (2001). Effects of career decision-making self-efficacy and degree utility on student persistence: A path analytic study. *Journal of College Student Retention: Research, Theory & Practice*, 3(3), 285-299. https://doi.org/10.2190/4d9v-dfw1-vdlx-k7gf

Pérez, E., Pássera, J., Olaz, F., & Osuna, M. (2005). *Orientación, información y educación para la elección de carrera [Orientation, information and education for career choice]*. Paidós.

R Core Team (2017). *R: A language and environment for statistical computing* [Computer Software]. R Foundation for Statistical Computing. https://www.R-project.org/.

Rainie, L. (2010). *Internet, broadband, and cell phone statistics*. Pew Research Center.

Reile, D. M., & Bowlsbey, J. (2000). Using the Internet in career planning and assessment. *Journal of Career Assessment*, 8(1), 69-84. https://doi.org/10.1177/106907270000800106

Richaud, M. C., Mestre, M. V., Lemos, V. N., Tur, A., Ghiglione, M. E., & Samper, P. (2013). La influencia de la cultura en los estilos parentales en contextos de vulnerabilidad social [The influence of culture on parenting styles in contexts of social vulnerability]. *Avances en Psicología Latinoamericana*, 31(2), 419-431. https://revistas.urosario.edu.co/index.php/apl/article/view/1951

Sampson, J. P. J. (1999). Integrating Internet-based distance guidance with services provided in career centers. *The Career Development Quarterly*, 47(3), 243-254. https://doi.org/10.1002/j.2161-0045.1999.tb00734.x

Sturge-Apple, M. L., Davies, P. T. & Cummings, M.E. (2006). Impact of hostility and withdrawal in interparental conflict on parental emotional unavailability and children's adjustment difficulties. *Child Development*, 77 (6), 1623–1641. https://doi.org/10.1111/j.1467-8624.2006.00963

Sawitri, D. R., Creed, P. A., & Zimmer-Gembeck, M. J. (2015). Longitudinal relations of parental influences and adolescent career aspirations and actions in a collectivist society. *Journal of Research on Adolescence*, 25(3), 551-563. https://doi.org/10.1111/jora.12145

Summers, J. J., & Falco, L. D. (2018). Evaluating construct validity of the Middle School Self-Efficacy Scale with high school adolescents. Journal of Career Development, 1-18. https://doi.org/10.1177/0894845318773383

Taylor, K. M., & Betz, N. E. (1983). Applications of self-efficacy theory to the understanding and treatment of career indecision. *Journal of Vocational Behavior*, 22, 63-81. https://doi.org/10.1016/0001-8791(83)90006-4

Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., & Mislevy, R. J. (2000). *Computerized adaptive testing: A primer*. Routledge.

Walsh, W. B., & Betz, N. E. (1995). *Tests and assessment* (3rd ed.). Prentice-Hall, Inc.

Ware, J. E., Bjorner, J. B., & Kosinski, M. (2000). Practical implications of item response theory and computerized adaptive testing. *Medical care*, 38(9), 73-82. https://pdfs.semanticscholar.org/2e43/6893c89f845477f9b477bd52d3ed5737ec92.pdf

Wright, B. D., & Bell, S. R. (1984). Item banks: what, why, how. *Journal of Educational Measurement*, 21(4), 331-345. https://doi.org/10.1111/j.1745-3984.1984.tb01038.x

Wright, B., & Linacre, M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions,* 8(3). http:// www.rasch.org/rmt/rmt83b.htm

Wright, B. D., & Stone, M. H. (2003). Five steps to science: Observing, scoring, measuring, analyzing, and applying. *Rasch Measurement Transactions*, 17(1), 912-913. https://goo.gl/ZJagiH

Table 1

Characteristics of the 20 Items Selected from the CEDSE and MSSE Scales

	Rasch Parameters					DIF					
Item Content	δ_i	*SE*	Infit	Outfit	r_{bp}	Rasch-Welch (*t*)	5th grade	6th grade	Rasch-Welch (*t*)	Female	Male
1. Figure out which career options could provide a good fit for your personality (Darme cuenta qué carrera será una buena elección de acuerdo a mi tipo de personalidad)	.10	.04	.76	.79	.53	.94	.06	.14	2.22	.18	.00
2. Identify careers that best use your skills (Identificar carreras en las que pueda hacer un mejor uso de mis habilidades)	-.05	.04	.79	.83	.53	.64	-.08	-.03	2.94	.06	-.19
3.Pick the best-fitting career option for you from a list of your ideal careers (Elegir la mejor opción para mí, de la lista de mis carreras ideales)	.01	.04	.83	.83	.52	-.25	.03	.00	.50	.03	-.01
4. Learn more about careers you might enjoy (Aprender más a cerca de carreras que podrían gustarme)	-.16	.04	1.04	1.16	.34	-3.33	-.01	-.29	.83	-.12	-.19
5. Match your skills, values, and interests to relevant occupations (Comparar si mis habilidades, valores e intereses coinciden con ocupaciones relevantes)	.27	.04	.71	.72	.53	2.06	.18	.34	.24	.27	.25
6. Make a well-informed choice about which career path to pursue (Tomar una decisión fundamentada sobre qué carrera seguir)	.02	.04	1.02	1.03	.52	1.34	-.04	.07	-.08	.01	.02

7. Learn more about jobs that could offer things that are important to you (Aprender más acerca de los trabajos que podrían ofrecerme cosas que son importantes para mí)	-.09	.04	.85	.85	.44	-1.43	-.02	-.14	1.11	-.04	-.13
8. Identify careers that best match your interests (Identificar las carreras que coincidan mejor con mis intereses)	-.28	.04	.75	.74	.57	.70	-.31	-.25	.52	-.10	-.31
9. Find information in the library about five occupations I am interested in (Buscar información en la biblioteca sobre cinco ocupaciones o profesiones que me interesen)	.96	.04	1.26	1.29	.33	.85	.92	.99	-.22	.86	1.12
10. Make a plan of my educational goals for the next three years (Hacer un plan de mis objetivos educativos para los próximos tres años)	.70	.04	1.11	1.11	.45	-.55	.72	.68	-.56	.53	.71
11. Select one occupation from a list of possible occupations I am considering (Seleccionar una ocupación/profesión de la lista de las posibles carreras que estoy considerando)	.08	.04	.76	.76	.60	-1.85	.16	.01	.89	.15	.11
12. Determine what occupation would be best for me (Determinar cuál profesión sería la mejor para mí)	-.08	.04	.85	.84	.57	-.65	-.05	-.10	-1.26	-.34	-.12
13. Decide what I value most in an occupation (Decidir qué es lo que más valoro de una profesión/trabajo)	-.09	.04	.84	.84	.46	-.30	-.08	-.10	-3.14	-.19	-.03
14. Resist attempts of parents or friends to push me into a career I believe is beyond my abilities or not for me (Resistir a los intentos de mis padres o amigos para que ingrese a una carrera que considero que está más allá de mis capacidades o que creo que no es para mí)	-.40	.04	2.27	2.33	.19	3.28	-.55	-.26	.88	.22	-.25

15. Describe the job skills of a career I might like to enter (Describir las habilidades que considero necesarias para desempeñarme en la carrera a la que me gustaría ingresar)	-.04	.04	.77	.78	.50	1.27	-.09	.01	-.07	-.21	-.08
16. Choose a career in which most workers are the opposite sex (Elegir una carrera en la que la mayoría de los profesionales son del sexo opuesto al mío)	.40	.04	1.68	1.71	.23	-.41	.42	.39	-.25	.59	.41
17. Choose a career that will fit my interests (Elegir una carrera que se ajuste a mis intereses)	-.60	.05	.73	.71	.57	-.22	-.59	-.61	-1.12	-.63	-.59
18. Decide what kind of schooling I will need to achieve my career goal (Decidir qué tipo de formación necesitaré para alcanzar mis objetivos académicos/profesionales)	-.11	.04	.71	.74	.53	.02	-.11	-.11	2.52	-.24	-.06
19. Find out the average salary of people in an occupation (Averiguar cuál es el salario promedio de los diferentes profesionales)	.15	.04	1.25	1.31	.26	-2.47	.25	.05	-2.89	.20	.04
20. Talk with a person already employed in a field I am interested in (Hablar con una persona que ya se encuentre trabajando en la profesión que me interesa)	-.79	.05	1.15	1.15	.36	.70	-.82	-.75	-2.21	-.63	-.64

Table 2

Descriptive Statistics of an Adaptive Test with the Parameters of the CDMSE-IB in a Simulation of 1000 Subjects According to Skill Level

Estimated ability level (Error 0.5)	Quantity of items
-2	19.96
-1.5	19.77
-1	19.09
0	17.73
1	19.25
1.5	19.82
2	19.96

Table 3

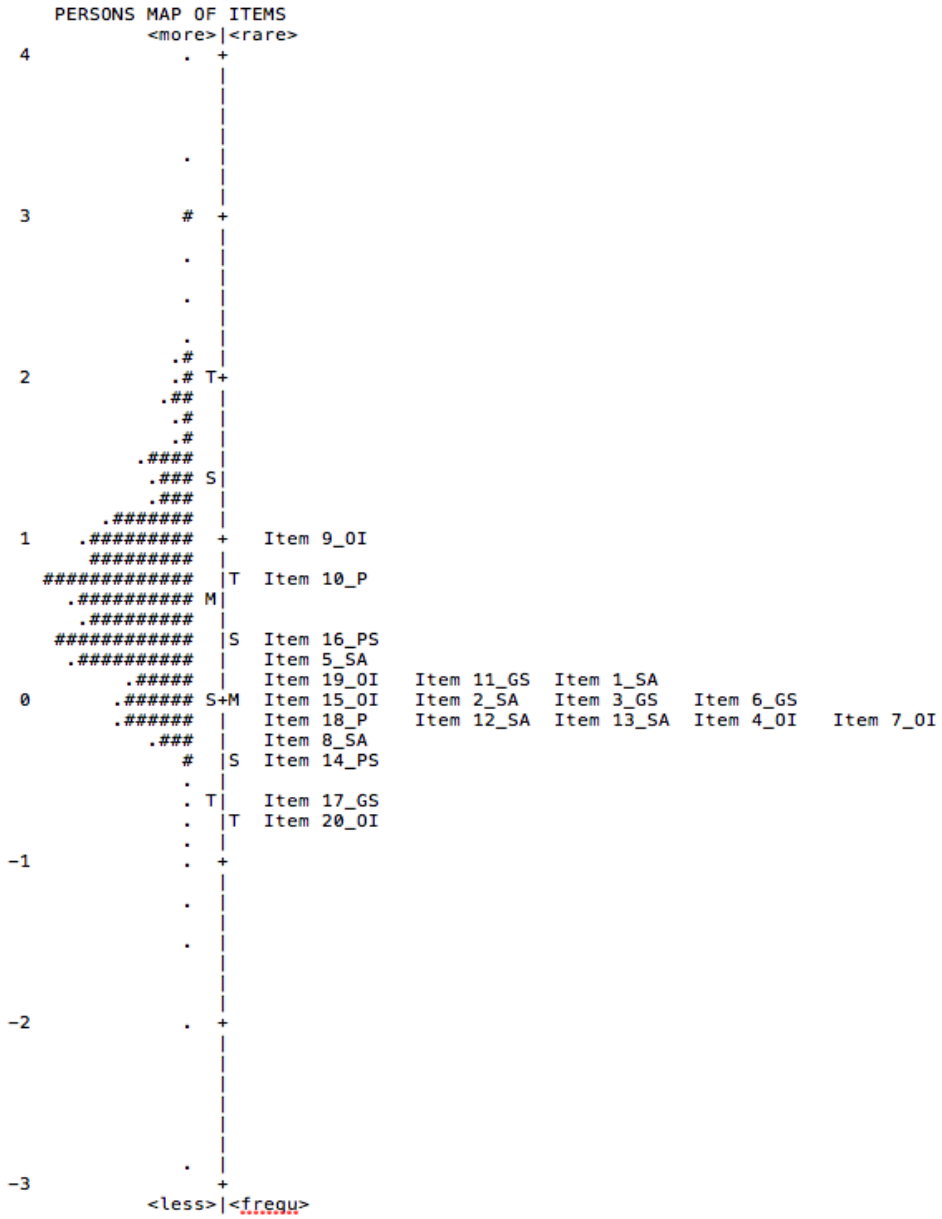
Concurrent Validity of the CDMSE-IB. Correlations, Means, and Standard Deviations

Variables	1	2	3	4	*M*	*SD*
1.CDMSE-IB	1				73.50	11.37
2.CDQ	.50**	1			16.70	4.78
3.CDS	-.32**	-.50**	1		31.30	9.05
4.CCA	-.34**	-.50**	.70**	1	29.08	8.72

Note. **p < .01. CDMSE-IB = Career decision-making self-efficacy item bank; CDQ = Career decision questionnaire; CCA = Career commitment anxiety.

Figure 1

Map of Persons and Items



Note. The left column shows the allocation of persons according to their level of ability. The symbol # represents a group of seven persons, and the point "." represents groups from one to four persons. *M* marks the media of persons and items. *S* is an *SD* distant from the media. *T* is two *SDs* distant from the media. SA = Self-Appraisal, GS = Goal Selection, OI = Occupational Information, P = Planning, PS = Problem Solving.