Int J Educ Vocat Guidance
DOI 10.1007/s10775-014-9273-3

Validation of a short form of an indecision test: the vocational assessment test

France Picard · Éric Frenette · Frédéric Guay · Julie Labrosse

Received: 19 July 2013 / Accepted: 28 May 2014

Abstract The purpose of this research was to validate the scores of a short form of a new instrument, l'*Épreuve de décision vocationnelle, forme scolaire* (EDV-9S; vocational assessment test), which measures six indecision-related problems (lack of self-knowledge, lack of readiness, lack of method in decision making, lack of information, external barriers, and dysfunctional beliefs). In Study 1 ($n = 778$; $M_{age} = 17.7$ years; female = 55.4 %), we first assess the reliability and factorial validity of the EDV-9S scores. Subsequently, 24 of 48 items were selected using Rasch modeling. In Study 2 ($n_1 = 757$; $M_{age} = 18.0$ years; female = 55.9 %), we assess the reliability, factorial validity, and convergent and divergent validity of the short form scores and their invariance across gender and over time. Findings from both studies suggest that the original as well as the short form scores present

F. Picard (✉)
Centre de recherche et d'intervention sur l'éducation et la vie au travail (CRIEVAT), Département des Fondements et pratiques en éducation, Pavillon des sciences de l'éducation, Université Laval, 2320 rue des Bibliothèques, bureau 574, Quebec, QC G1V 0A6, Canada
e-mail: France.Picard@fse.ulaval.ca

É. Frenette
Département des Fondements et pratiques en éducation, Pavillon des sciences de l'éducation, Université Laval, 2320 rue des Bibliothèques, bureau 454, Quebec, QC G1V 0A6, Canada
e-mail: Eric.Frenette@fse.ulaval.ca

F. Guay
Pavillon des sciences de l'éducation, Université Laval, 2320 rue des Bibliothèques, bureau 942, Quebec, QC G1V 0A6, Canada
e-mail: Frederic.Guay@fse.ulaval.ca

J. Labrosse
Centre d'Étude des COnditions de vie et des BESoins de la population (ÉCOBES), Cégep de Jonquière, 3791, rue de la Fabrique, 6e étage, Jonquière, QC G7X 7W2, Canada
e-mail: julie.labrosse@cjonquiere.qc.ca

Published online: 01 July 2014

adequate psychometric properties and could be used to assess indecision problems in French-speaking students in higher education.

Résumé. Élaboration d'une version abrégée d'un questionnaire d'indécision : l'Épreuve de décision vocationnelle, forme scolaire. La présente recherche a pour but de valider les scores d'une version abrégée de l'*Épreuve de décision vocationnelle, forme scolaire* (EDV-9S). Ce questionnaire comporte six dimensions mesurant les sources de l'indécision vocationnelle : le manque de connaissance de soi, le manque de développement vocationnel, le manque de méthode de décision, le manque d'information, les obstacles externes et les anticipations pessimistes. La première étude ($n = 778$; $M_{\text{âge}} = 17{,}7$ ans; femmes = 55,4 %) a permis de vérifier le niveau de fidélité de chacune des échelles et de tester la validité factorielle de l'EDV-9S. La version abrégée a été mise au point à l'aide du modèle de Rasch (rating scale model); 24 items ont été sélectionnés sur les 48. La deuxième étude ($n_1 = 757$; $M_{\text{âge}} = 18{,}0$ ans; femmes = 55,9 %) a permis d'évaluer le niveau de fidélité de chacune des échelles et de tester la validité factorielle, la validité convergente et divergente, ainsi que l'invariance en fonction du genre et du temps des scores de la version abrégée. Les résultats de ces deux études soulignent que la forme originale et la version abrégée de l'EDV-9S présentent des propriétés psychométriques adéquates. Ce questionnaire pourrait être utilisé afin d'évaluer les sources d'indécision de la population étudiante francophone aux études supérieures.

Zusammenfassung. Validierung einer Kurzform eines Unentschlossenheitstests: Der Berufs Assessment Test. Der Zweck dieser Untersuchung war es, die Werte einer Kurzform von einem neuen Instrument zu validieren, l'*Épreuve de décision vocationnelle, forme scolaire* (EDV-9S; Berufs Assessment Test). Das Instrument misst sechs Unentschlossenheit bezogene Probleme (Mangel an Selbstkenntnis, Mangel an Bereitschaft, Mangel an Verfahren der Entscheidungsfindung, fehlende Informationen, externe Barrieren und dysfunktionale Überzeugungen). In Studie 1 ($n = 778$; $M_{\text{Alter}} = 17{,}7$ Jahre; weiblich = 55,4 %), prüfen wir zunächst die Zuverlässigkeit und die faktorielle Validität der Werte des EDV -9S. Anschließend wurden 24 von 48 Items mittels Rasch -Modellierung gewählt. In Studie 2 ($n_1 = 757$; $M_{\text{Alter}} = 18{,}0$ Jahre; weiblich = 55,9 %), schätzen wir die Zuverlässigkeit, die faktorielle Validität, die konvergente und divergente Validität der Kurzform sowie deren Invarianz über Geschlecht und Zeit. Die Ergebnisse aus beiden Studien deuten darauf hin, dass die ursprüngliche und die Kurzform angemessene psychometrische Eigenschaften zeigen und zur Beurteilung von Unentschlossenheitsproblemen von Französisch-sprechenden Studenten in der Hochschulbildung verwendet werden können.

Resumen. Validación de una forma resumida de un test de indecisión: El *Vocational Assessment Test*. El objetivo de esta investigación fue validar los resultados de una forma resumida de una nueva herramienta, llamada l'*Épreuve de décision vocationnelle, forme scolaire* (EDV-9S; Vocational Assessment Test), que mide seis problemas relacionados con la indecisión (falta de autoconocimiento, falta de determinación, falta de método para la toma de decisiones, falta de información, barreras externas y creencias disfuncionales). En un primer experimento ($n = 778$;

M_{edad} = 17.7 años; mujer = 55.4 %), evaluamos la fiabilidad y la validez factorial de los resultados del EDV-9S. Para ello se seleccionaron mediante modelización de Rasch 24 de 48 preguntas. En un segundo experimento (n = 757; M_{edad} = 18.0 años; mujer = 55.9 %), evaluamos la fiabilidad, la validez factorial, y la validez convergente y divergente de los resultados del test resumido, así como su invariabilidad en cuanto a sexo y a lo largo del tiempo. Se deduce de ambos estudios que los resultados del test original y del resumido presentan propiedades psicométricas adecuadas, y que ambos podrían emplearse perfectamente para evaluar los problemas de indecisión de los estudiantes francófonos de educación secundaria.

Keywords Vocational assessment test · Transcultural validation · Vocational indecision

Introduction

The general purpose of this research was to develop a short form of a new instrument, l'*Épreuve de décision vocationnelle, forme scolaire* (EDV-9S), an indecision test developed by Forner (2009, 2010). Indecision refers to students' inability to make a choice or engage in action required for decision making (Forner, 2007). Although several high-quality instruments are available for assessing decision problems (Kelly & Lee, 2002), there is a lack of indecision measures adapted for college and university students (Germeijs & Verchueren, 2006). French-speaking researchers are faced with an additional challenge, as they must validate the scores obtained on these translated instruments (i.e., transcultural validation) before conducting research.

Academic and career indecision is a common problem in high school and college students (Germeijs & Verchueren, 2006; Kelly & Lee, 2002; Morgan & Ness, 2003). For example, in the education system in the province of Quebec, Canada, approximately 30 % of first-year college students stated that they had no specific ideas about their academic path or profession, which may explain why they eventually switch programs (Conseil supérieur de l'éducation, 2008; Picard, Boutin, & Skakni, 2010). Moreover, indecision has been associated with certain negative consequences, including lower college or university achievement (American College Testing, 2009; Forner, 2000; Rose & Elton, 1971; Taylor, 1982). Conversely, certainty about a major or a career choice has been positively related to persistence intentions (Neuville & Frenay, 2012) and academic success (Graunke & Woosley, 2005; Lent, Brown, & Hackett, 1994; Neuville & Frenay, 2012). Studying vocational indecision and its effects on the pathways of college students is therefore a promising line of research.

In order to gain a better understanding of the role of indecision in students' experience along the college and university pathway, researchers should use indecision scales that (1) are derived from the latest decision-making theories, (2) present good psychometric properties, (3) are adapted to students enrolled in higher education, (4) are available in the students' native language, and (5) are relatively

short, in order to reduce students' fatigue, frustration, or boredom (Robins, Hendin, & Trzesniewski, 2001).

New-generation theories on career decision-making focus on the multidimensionality of career indecision (Forner, 2010; Gati & Asher, 2001; Gati, Gadassi, & Mashiah-Cohen, 2012; Gati, Krausz, & Osipow, 1996; Kelly & Lee, 2002). For example, the decision-making theory proposed by Gati et al. (1996) is based on a hierarchical taxonomy of indecision problems in students prior to or during the decision-making process, in contrast to an ideal career decision maker. Indecision problems are classified into three general categories that include indecision-related factors: lack of readiness (lack of motivation, indecisiveness, and dysfunctional beliefs), lack of information (about the decision-making process, oneself, occupations, and ways of obtaining information), and inconsistent information (unreliable information, internal and external conflicts). Kelly and Lee (2002) provided empirical support for Gati et al.'s (1996) taxonomy. They reported six factors corresponding to indecision problems in university freshmen: trait indecision, choice anxiety, information needs, lack of information, identity diffusion, and disagreement with others.

The indecision problems addressed by the EDV-9S largely correspond to the three general indecision problem categories conceptualized by Gati et al. (1996) and to Kelly and Lee's (2002) components. The 52-item EDV-9S includes scales measuring academic (2 items) and career indecision (2 items) and six subscales measuring indecision-related problems (48 items). Forner (2009, 2010) validated the EDV-9S scores in students ($N = 1152$; $M_{age} = 15.6$ years, $SD = 1.35$; female = 60.5 %) attending high school in France. Six factors emerged from the exploratory factor analysis (EFA) using varimax rotation, supporting the factorial validity of the EDV-9S. Although varimax rotation was used in Forner's validation, we assume that the EDV-9S subscales are correlated, although this has not been tested to date. According to Kelly and Lee's (2002) principal components analysis with varimax and promax rotations, the subscales are correlated on the following three indecision tests: career decision scale (CDS; Osipow, 1987), career factors inventory (CFI; Chartrand & Robbins, 1997), career decision-making difficulties questionnaire (CDDQ; Gati et al., 1996). Internal consistency values of the EDV-9S were considered acceptable to satisfactory, ranging from .66 to .87 (Forner, 2010). Despite adequate reliability and factorial validity, no confirmatory factor analysis (CFA) has been performed to date. Moreover, the EDV-9S is relatively lengthy for research purposes.

Because longer surveys affect the response rate and the amount of missing data (Graham, 2009), researchers often reduce the item pool. The short form should be small enough to allow a rapid assessment of indecision problems, but it should also have sufficient items to present good psychometric properties (Marsh, Hau, Balla, & Grayson, 1998). Whereas using a limited pool of items from a longer scale typically yields poorer psychometric properties than the longer version, it has the advantage of reducing item redundancy, which in turn reduces students' fatigue, frustration, and boredom (Robins et al., 2001). Selecting a pool of items from an original scale has gained in popularity in recent years, particularly for the Self-Directed Search Questionnaire (Holland, Fritzsche, & Powell, 1994; Poitras, Guay, & Ratelle, 2012),

the Inventory of Children's Individual Differences (Deal, Halverson, Martin, Victor, & Baker, 2007), the big five personality factor questionnaire (Gosling, Rentfrow, & Swann, 2003; Muck, Hell, & Gosling, 2007; Rammstedt & John, 2007), and the dyadic adjustment scale (Sabourin, Valois, & Lussier, 2005).

Rasch modeling: an effective method to select a representative pool of items from a scale

Rasch analysis has been used to produce short forms of questionnaires (e.g., Las Hayas, Quintana, Padierna, Bilboa, & Munoz, 2010; Panagiotis, Sideris, Protopapas, & Mouzaki, 2011; Weller et al., 2012). Cole, Rabin, Smith and Kaufman (2004) examined two important criteria for developing a short form using Rasch modeling: item fit statistics and the redundancy level of item endorsement. First, items not meeting the Rasch statistical criteria for model fit are removed from the scale. Second, redundant items are identified when two or more have the same endorsement level.

A multitrait-multimethod approach to convergent/divergent validity

To assess convergent and divergent validity, Campbell and Fiske (1959) proposed the multitrait-multimethod approach (MTMM), in which two or more traits are measured using each of two or more methods. In the present study, the multiple traits are the six subscales of the EDV-9S and the multiple methods are two measurement times, as described in the "Method" section. Although multiple occasions are used less often than multiple assessment sources as the multiple methods in MTMM matrix studies, it is important to keep in mind that Campbell and Fiske (1959) mentioned that multiple occasions *can* be used. Indeed, they use the term "multiple methods" very broadly, in the sense of not only different information sources but also multiple occasions. Moreover, Campbell considered multiple occasions as multiple methods in later publications (Campbell & O'Connell, 1967, 1982). In addition, Marsh and colleagues have demonstrated that multiple occasions can be used as multiple methods (e.g., Marsh & Butler, 1984; Marsh & Hocevar, 1988). However, the MTMM matrix should be interpreted according to the method used. For example, convergent validity based on multiple occasions represents test–retest reliability. Nonetheless, the MTMM approach used here provides a richer framework for assessing test–retest reliability than a simple presentation of test–retest correlations.

Measurement invariance

To compare groups or to obtain valid comparisons over time, a questionnaire should measure identical constructs with the same structure. When this is the case, the questionnaire is called measurement invariant. If measurement invariance can be demonstrated, then participants across all groups interpret the individual question, as well as the underlying latent factor, in the same way (van de Schoot, Lugtig, &

Hox, 2012). CFA using structural equation modeling is used to assess the invariance of the factor structure across gender and measurement times.

Goals of the present research

The main goal of this study was to validate the scores of a short form of a new vocational indecision test, the EDV-9S (Forner 2009, 2010), to be used for research with French-speaking higher education students. To meet this goal, two studies were performed. In Study 1 we (1) assess the psychometric properties (reliability, factorial validity) of the original form of the EDV-9S, (2) propose a short form based on Rasch modeling (rating scale model), and (3) assess the psychometric properties (reliability and factorial validity) of the short form using CFA. In Study 2, we verify the psychometric properties of the short form: reliability, factorial validity, convergent and divergent validity, and its invariance across gender and over time.

Study 1 method

Participants

A total of 778 first-year students attending nineteen French-speaking colleges (post-secondary level) in Quebec took part in the study: 344 males, 431 females, and three with unidentified gender ($M_{age} = 17.68$ years, $SD = 1.78$).

Instrument

Each subscale of the EDV-9S (Forner, 2009, 2010) contains eight items rated on a four-point Likert-type scale (*1 = does not apply*, *4 = definitely applies*). The six subscales are: *lack of self-knowledge*—uncertainty about self-representation that may lead to anxiety at the time of making a decision; *lack of readiness*—inadequate commitment to the orientation process, manifested in a delay in making academic and career choices; *lack of method in decision making*—difficulty in deciding on different but equally attractive goals; *lack of information*—insufficient knowledge on the choice of trades, careers, and academic programs that can slow the decision-making process; *external barriers*—real and tangible events that can hinder the fulfillment of one's plan; and *dysfunctional beliefs*—negative expectations leading to the perception that one's education and career goals are unattainable (Forner, 2010).

Procedure

Colleges in Quebec were invited in 2009 to participate in the study, based on geographical location (urban or rural area), size, and language of instruction. Colleges were asked to recruit two classes of first-year students (freshmen) with similar characteristics in terms of gender and academic pathway. Students had to be

enrolled in a technical program (i.e., 3-year career program), pre-university program (i.e., 2-year pre-university general education program), or a transition-to-college program (*Session d'accueil et d'intégration*—freshman program). Data were collected in August 2009. The questionnaires and instructions were provided to the colleges. During class time, designated administrators explained the goal of the study and asked students to raise their hands if they wanted to participate. Those who agreed filled out a consent form and the EDV-9S. Those who declined were asked to leave the class. No compensation was offered to students for participating.

Statistical analyses

Descriptive analysis

The mean, standard deviation, skewness, and kurtosis for each subscale are provided in Table 1.

Confirmatory factor analysis

The fit of the theoretical model (six interrelated subscales) to the data was assessed with EQS software (Version 6.2; Bentler, 2006) using robust maximum likelihood estimation with categorical variables. The comparative fit index (CFI; Bentler, 1990), non-normed fit index (NNFI, also known as the Tucker & Lewis, 1973, Index), and root-mean-square error of approximation (RMSEA; Steiger, 1990) were used to ascertain model fit. The NNFI and CFI usually vary along a 0 to 1 continuum (although the NNFI could be greater than 1 due to sampling, this is rarely the case in practice), in which values greater than .90 and .95 are typically deemed acceptable and excellent fit to the data, respectively (Hu & Bentler, 1999; Schumacker & Lomax, 1996). Several researchers suggested that a RMSEA less than .05 indicates a close fit and that values up to .08 represent reasonable errors of approximation (Browne & Cudeck, 1993; Marsh, Ellis, Parada, Richards, & Heubeck, 2005).

Rasch analyses

Data were analyzed using the Rasch rating scale model (Andersen, 1977; Andrich, 1978; Wright & Masters, 1982) included in WINSTEPS (Version 3.32; Linacre & Wright, 1999). Goodness of fit statistics and endorsement level were estimated for each item. The item fit statistics included two indexes: infit and outfit (Linacre, 1994; Wright & Masters, 1982). The infit statistic denotes the information-weighted mean square residual difference between observed and expected responses. The outfit statistic, which is more sensitive to outliers, was used as an additional reference. It denotes the usual unweighted mean square residual. Infit and outfit statistics with values around 1 are considered satisfactory, and values within the range 0.5–1.5 are productive for measurement (Linacre, 2002). Once redundant items were identified, two experts analyzed them to identify the best one to retain.

Table 1 Descriptive statistics of the original form of the EDV-9S

Subscales	*M*	*SD*	Skewness	Kurtosis
LSK	2.01	0.84	0.54	−0.77
LR	1.49	0.58	1.49	2.20
LMDM	2.30	0.84	0.17	−0.99
LVI	1.84	0.70	0.84	0.18
EB	1.28	0.44	2.99	11.85
DB	1.94	0.70	0.55	−0.38

LSK lack of self-knowledge, *LR* lack of readiness, *LMDM* lack of method in decision making, *LVI* lack of vocational information, *EB* external barriers, *DB* dysfunctional beliefs

Results

Reliabilities of the six subscales were calculated using Cronbach's alpha coefficient (Cronbach, 1951). Internal consistency values for the six subscales ranged from .81 to .91 and were considered satisfactory (Field, 2005; Kline, 1999; Nunally, 1978). Moreover, item discrimination indices were considered adequate, with item-total correlations above .30 (Crocker & Algina, 1986) for each subscale. Results of the CFA showed that the NNFI (.98), the CFI (.98), and the RMSEA (.045; 95 % CI [.043–.047]) were considered excellent. Correlations among subscales ranged from .30 to .76 ($M = .52$, $SD = .14$). All factor loadings were substantial and significant (z values > 1.96). Average loadings for items in each subscale are high: *lack of self-knowledge* ($M = .82$); *lack of readiness* ($M = .77$); *lack of method in decision making* ($M = .80$); *lack of information* ($M = .75$); *external barriers* ($M = .72$) and *dysfunctional beliefs* ($M = .72$).

In order to select the items to be included in the short form of the EDV-9S, as a preliminary step, calibrations for each item and each student were performed using WINSTEPS separately for the six subscales to respect the unidimensionality assumption (Wright & Linacre, 1989). Reliability criteria and separation indices for students and items are presented in Table 2 for the six subscales. Rasch reliability indices for students were considered satisfactory, except for the *lack of readiness* subscale, which was considered acceptable, and the *external barriers* subscale, which shows low reliability. Rasch reliability indices for items were considered satisfactory for all subscales, allowing item selection based on the calibration results.

For example, the Rasch calibration results for the *lack of readiness* subscale are presented in Table 3. One item (B4) did not meet the Rasch criteria (infit and outfit) for model fit and was therefore removed. Three groups of two items with endorsement redundancy were identified. Based on the experts' judgment, the three items (B2, B3, and B6) were removed.

The same procedure was used for the five other subscales. For the *lack of self-knowledge* subscale, all items met the Rasch criteria for model fit. However, four groups of two items with redundancy were identified. Based on the experts' judgment, four items were removed. For the *lack of method in decision making*, *lack*

Table 2 Real separation index and reliability for items and students' scores on the indecision-related problems subscales

Subscale	Items		Students	
	Real sep.	Reliability	Real sep.	Reliability
LSK	6.54	.98	2.04	.81
LR	5.89	.97	1.32	.64
LMDM	6.33	.98	2.45	.86
LVI	13.67	.99	1.86	.78
EB	5.19	.96	0.84	.41
DB	5.80	.97	2.06	.81

LSK lack of self-knowledge, *LR* lack of readiness, *LMDM* lack of method in decision-making, *LVI* lack of vocational information, *EB* external barriers, *DB* dysfunctional beliefs

Table 3 Rasch calibration results for the *Lack of readiness* subscale

Item number	δ	*SE*	Infit		Outfit		r_t
			MNSQ	ZSTD	MNSQ	ZSTD	
B1	.57	.08	1.07	0.9	0.93	−0.7	.64
B6	.57	.08	1.07	0.9	0.93	−0.7	.64
B3	.09	.07	0.79	−3.3	0.73	−3.6	.73
B8	.09	.07	0.79	−3.3	0.73	−3.6	.73
B5	.01	.07	1.16	2.3	1.29	3.4	.67
B2	−.26	.06	0.81	−3.1	0.80	−3.0	.76
B7	−.26	.06	0.81	−3.1	0.80	−3.0	.76
B4	−.81	.06	1.59	8.3	1.67	8.7	.68

δ Lack of readiness level, *SE* standard error, *MNSQ* mean-square fit statistic, *ZSTD* standardized mean-square statistic, r_t point-biserial correlation between items and total measured *lack of readiness* based on the Rasch calibrated item scores and total scores

of vocational information, *external barriers*, and *dysfunctional beliefs* subscales, one item in each subscale did not meet the Rasch criteria (infit and outfit) for model fit, and was removed. For the four subscales, three groups of two items with endorsement redundancy were identified. Based on the experts' judgment, three items each were removed from the four subscales. The short form of the EDV-9S therefore comprises 24 items, each subscale containing four items: *lack of self-knowledge* ($\alpha = .79$); *lack of readiness* ($\alpha = .74$); *lack of method in decision making* ($\alpha = .88$); *lack of vocational information* ($\alpha = .80$); *external barriers* ($\alpha = .71$), and *dysfunctional beliefs* ($\alpha = .81$).

Results of the CFA for the short form showed excellent NNFI (.98), CFI (.98), and RMSEA (.042, 95 % CI [.037–.046]) values. Correlations among subscales ranged from .32 to .84 ($M = .56$, $SD = .15$). All factor loadings were substantial and significant (z values > 1.96). The mean and standard deviation of each subscale for the short form are presented in Table 4.

Table 4 Mean and standard deviation for the short form of the EDV-9S

Item	Study 1		Study 2, time 1		Study 2, time 2		Study 2, time 1 Female		Study 2, time 1 Male		Study 2, time 2 Female		Study 2, time 2 Male	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
LSK	2.11	0.86	2.21	0.90	2.10	0.87	2.26	0.92	2.16	0.87	2.13	0.89	2.10	0.83
LR	1.43	0.60	1.46	0.60	1.43	0.61	1.36	0.54	1.58	0.65	1.33	0.53	1.57	0.70
LMDM	2.29	0.88	2.37	0.88	2.29	0.86	2.35	0.89	2.41	0.86	2.25	0.88	2.35	0.83
LVI	2.03	0.79	2.07	0.78	1.99	0.77	2.05	0.79	2.11	0.76	1.99	0.78	1.99	0.76
EB	1.29	0.49	1.34	0.56	1.35	0.52	1.33	0.57	1.36	0.55	1.35	0.51	1.37	0.53
DB	2.02	0.81	2.10	0.85	2.14	0.84	2.13	0.87	2.06	0.83	2.18	0.86	2.05	0.80

LSK lack of self-knowledge, *LR* lack of readiness, *LMDM* lack of method in decision making, *LVI* lack of vocational information, *EB* external barriers, *DB* dysfunctional beliefs

Study 2 method

Participants, instrument, and procedure

The procedure and measure (EDV-9S) used in Study 1 were retained in Study 2. However, the design was longitudinal and included two data collection waves. Data were collected in August 2010 (Time 1) and in December 2010 (Time 2). At Time 1, 757 students at eighteen colleges participated: 324 males, 423 females, and 10 with unidentified gender. Participants' mean age at Time 1 was 18.00 years ($SD = 2.18$). At Time 2, 584 students participated: 223 males, 352 females, and nine with unidentified gender. At Time 1, participants' mean age was 17.94 years ($SD = 2.14$). Students who were absent in December were followed up in January 2011 using an online survey, and 17 of these completed the questionnaire.

Statistical analyses

Multitrait-multimethod approach

Campbell and Fiske (1959) proposed four guidelines for assessing convergent and divergent validity using the MTMM matrix. First, convergent validity occurs when the correlation between two different methods assessing the same trait is high and significant. Second, divergent validity occurs when correlations between two traits assessed by a different method are lower than convergent correlations. Third, a method effect exists when correlations between traits of the same method are higher than convergent correlations. Fourth, the pattern of correlations among different traits should be similar for different methods. In sum, convergent and divergent validity is supported when (a) convergent correlations are high, (b) divergent correlations are lower than convergent correlations, (c) the method effect is low, and (d) the pattern of correlations among different traits is similar for different methods.

Confirmatory factor analysis

CFA analyses were run to assess invariance of the factor structure across gender and measurement times. In the CFA solution, separate covariance matrices were computed for males and females and for Time 1 and Time 2. In the case of parallel data for more than one group, the invariance of the solution can be tested by constraining to be equal any one, any set of, or all parameter estimates for the two groups. The minimal condition of factorial invariance is the invariance of the factor loadings. Model comparison was facilitated by positing a nested ordering of models, in which the parameter estimates for a more restrictive model are a proper subset of those in a more general model (Bentler, 1990). Thus, three separate models were run to test: configural invariance (Model 1), measurement invariance (Model 2), and subscale correlation invariance (Model 3). However, uniquenesses were not constrained to equality, because this test is considered excessively stringent (Byrne, 1995).

Results

Internal consistency, assessed by Cronbach's alpha, of the indecision-related problems subscales ranged from .71 to .87 for Time 1 and from .67 to .88 for Time 2. Item discrimination indices for each subscale were considered adequate for Time 1 and Time 2, with item-total correlations above .30 (Crocker & Algina, 1986). The mean and standard deviation for each subscale are presented in Table 4. Results of the CFA for the short form of the EDV-9S showed that the NNFI (.98), CFI (.98), and RMSEA (.042; 95 % CI [.037–.046]) for Time 1 and the NNFI (.99), CFI (.99), and RMSEA (.034; 95 % CI [.028–.040]) for Time 2 were satisfactory, indicating that the theoretical model (six interrelated factors) provided a good fit to the data. Correlations between subscales ranged from .31 to .82 ($M = .52$, $SD = 0.15$) for Time 1 and from .35 to .89 ($M = .56$, $SD = 0.14$) for Time 2. All factor loadings were substantial and significant (z values > 1.96) for Time 1 and Time 2.

Correlational analyses were performed to test the convergent and divergent validity of the EDV-9S scores using the MTMM matrix (see Table 5). This matrix is divided into three components: (a) correlations in bold represent relationships among different traits assessed with the same method (heterotrait-monomethod; HTMM), (b) correlations in italics represent relationships between different traits assessed with different methods (heterotrait-heteromethod; HTHM), and (c) correlations in block represent the convergent validities (monotrait-heteromethod; MTHM). The four guidelines proposed by Campbell and Fiske (1959) were applied in the matrix analysis.

First, convergent validities were relatively high. The six correlations were significant, ranging from .30 to .70 ($M = .53$; $SD = 0.13$), providing good support for this guideline. Second, convergent correlations were higher than HTHM correlations ($M = .28$; $SD = 0.13$), providing good support for the second criterion (divergent validity). Third, convergent correlations were mostly higher than HTMM correlations ($M = .40$; $SD = 0.13$), providing good support for the third guideline (moderate method effect). Finally, the pattern of correlations among different traits was similar across the two methods for Time 1 ($M = .39$, $SD = 0.14$) and Time 2 ($M = .42$, SD = 0.12).

To assess the measurement invariance across gender, we tested for the equivalence of factor loadings and covariances between subscales across gender. In the least restrictive model (Model 1: configural invariance), no parameters were constrained to be equal across gender, and this model provided a good fit to the data (CFI = .99, NNFI = .98, and RMSEA = .040; 95 % CI [.035–.045]). In Model 2, the factor loadings were constrained to be invariant across gender, and the model fit (CFI = .99, NNFI = .99, RMSEA = .039; 95 % CI [.034–.044]) did not differ significantly from that of Model 1 (i.e., the ΔCFI difference is lower than .01). In Model 3, factor loadings and covariances between subscales were constrained to be equal. The fit of this model (CFI = .99, NNFI = .99, RMSEA = .039; 95 % CI [.034–.044]) did not differ significantly from that of Model 2, indicating that covariances between subscales were invariant across gender. In sum, these three

Table 5 Multitrait-multimethod matrix for the short form of the EDV-9S

Measure	LSK[1]	LR[1]	LMDM[1]	LV[1]	EB[1]	DB[1]	LSK[2]	LR[2]	LMDM[2]	LVI[2]	EB[2]	DB[2]
Time 1												
LSK[1]	__											
LR[1]	**.37****	__										
LMDM[1]	**.58****	**.23**	__									
LVI[1]	**.68****	**.32***	**.53****	___								
EB[1]	**.29****	**.38***	**.23****	**.31****	__							
DB[1]	**.51****	**.24***	**.29****	**.47****	**.38****	___						
Time 2												
LSK[2]	.70**	*.35***	*.46***	*.54***	*.15***	*.36***	__					
LR[2]	*.29***	.48**	*.22*	*.25***	*.05*	*.13***	**.45****	__				
LMDM[2]	*.42***	*.19***	.54**	*.38***	*.07*	*.20***	**.56****	**.30**	__			
LVI[2]	*.54***	*.29***	*.39***	.57**	*.16***	*.26***	**.74****	**.40****	**.52****	__		
EB[2]	*.27***	*.16***	*.27***	*.27***	.30**	*.32***	**.40****	**.39****	**.27****	**.40****	__	
DB[2]	*.40***	*.18***	*.19***	*.36***	*.19***	.59**	**.49****	**.25****	**.31****	**.42****	**.41****	__

LSK lack of self-knowledge, *LR* lack of readiness, *LMDM* lack of method in decision making, *LVI* lack of vocational information, *EB* external barriers, *DB* dysfunctional beliefs

* $p < .05$; ** $p < .01$

Correlations in block represent the convergent validities

CFA analyses revealed that the factor structure of the short form of EDV-9S was invariant across gender.

To assess the measurement invariance over time, we tested for the equivalence of factor loadings and covariances between subscales across time. In the least restrictive model (Model 1: configural invariance), no parameters were constrained to be equal across time, and this model provided a relatively good fit to the data (CFI = .99, NNFI = .99, and RMSEA = .038; 95 % CI [.034–.041]). In Model 2, the factor loadings were constrained to be invariant across time, and the model fit (CFI = .99, NNFI = .99, RMSEA = .038, 95 % CI [.034–.041]) did not differ significantly from that of Model 1 (i.e., ΔCFI difference is lower than .01). In Model 3, factor loadings and covariances between subscales were constrained to be equal. The fit of this model (CFI = .99, NNFI = .99, RMSEA = .037, 95 % CI [.034–.041]) did not differ significantly from that of Model 2, indicating that covariances between subscales were invariant across time. In sum, these three CFA analyses revealed that the factor structure of the EDV-9S was invariant over time.

Discussion

The main goal of the present research was to validate the scores of a short form of a new instrument, the Vocational Assessment Test (*Épreuve de décision vocationnelle, forme scolaire*—EDV-9S) adapted for French-speaking students in higher education. To overcome the limitations of previous studies addressing the EDV-9S (Forner, 2009, 2010), we provided more evidence of the psychometric properties of the original version. In Study 1, we showed that the scores of the original version of the EDV-9S presented good psychometric properties (high internal consistency and factorial validity), supporting the six interrelated factors corresponding to indecision-related problems. These results build on previous research by using CFA and testing implicit correlations between subscales. The reliabilities of the six subscale scores were higher than those reported by Forner (2009, 2010). Based on these results, we developed a short version of the scale.

Rasch analyses allowed selecting 24 items from the 48 original items (Study 1). As proposed by Marsh et al. (1998), the short form of the EDV-9S was small enough to allow a rapid assessment of indecision problems. In both studies, the short form scores presented good psychometric properties—high internal consistency and factorial validity. In Study 2, the results of the MTMM analysis revealed good convergent and divergent validity.

Furthermore, the analyses revealed that the factorial structure (six interrelated factors) was invariant across gender and over time. This means that the participants across all groups interpreted the individual item, as well as the latent factor, in the same way (van de Schoot et al., 2012). Another strength of the present research was the sample design. It is very rare that two data waves in a large sample are used to develop and validate a short form of a questionnaire. Hence, the short form of the EDV-9S presented good psychometric properties.

Limitations

Although the EDV-9S is available in English, the scores on the English version were not validated in the present studies. The items selected for the short form, based on the experts' judgment, were based on the French version. The 24 selected items could therefore differ from those that would be selected in English. According to the invariance criterion of the Rasch model, and if the differential item functioning between the French and English versions is low, the same items should be selected. We recommend that the scores on the English version be validated, including an assessment of differential item functioning, by English-speaking researchers.

As mentioned above, researchers may limit the item pool to cope with time constraints. This has the advantages of reducing item redundancy as well as students' fatigue, frustration, and boredom (Robins et al., 2001). Although the 24 items presented good psychometric properties, the short form does not provide exactly the same information to measure indecision-related problems as the original version of the EDV-9S, notably because there are fewer items. In the future, we propose assessing the psychometric properties of the short form in different groups (e.g., university students, first generation in higher education, ethnocultural groups, gays and lesbians).

Attention must be paid to the external barriers subscale scores, which present acceptable reliability for the items but poor reliability for students in the Rasch analyses, low mean in the data waves for both studies, and the lowest convergent validity in the MTMM analysis. Kelly and Lee (2002) conducted a principal component analysis on three instruments—the CDS, CFI, and CDDQ—and provided support for six components. The disagreement with others component is almost similar to the EDV-9S External Barriers subscale, referring to the CDDQ External Conflicts scale items and the CDS item ("I'd like to be a____, but I'd be going against the wishes of someone who is important to me if I did so"). This is a relevant component for career counseling, as it could inhibit academic and career choices (Kelly & Lee, 2002). Subsequent studies could focus on the EDV-9S subscales, following Kelly and Lee's work.

Conclusion

These two studies have filled a methodological gap by validating the scores of a measuring instrument with good psychometric properties that is adapted for research in French-speaking students enrolled in higher education. The EDV-9S proved to be consistent with the new-generation approach to measuring indecision, based on decision theory (Gati & Asher, 2001; Gati, et al., 1996), as it reflected the multidimensionality of decision-related issues. This conclusion was supported by three validations, conducted in French-speaking population in France (Forner, 2009, 2010) and in Canada, of the original and the short form of the EDV-9S scores. The results make a relevant contribution to the research on French-speaking school populations. The present research builds on Forner's (2009, 2010) work by testing correlations between subscales using CFA, convergent and divergent validity

analysis, and invariance analysis for time and gender. The MTMM and CFA results provide empirical support for correlations between subscales. We recommend taking account of these correlations in future studies.

The short form of the EDV-9S could be used to shed light on the effects of vocational indecision on academic success and persistence in higher education. Another promising line of research would be to study the role of indecision on academic pathways (Neuville & Frenay, 2012) in order to understand how vocational indecision is related to educational behaviors such as withdrawal, re-entry, reorientation (changing programs in higher education) and reversal (e.g., enrolling in a college program after university).

Acknowledgments The EDV-9S was modified with special permission of the publisher. This research was supported by Grants from the Social Sciences and Humanities Research Council in Canada.

References

Andersen, E. B. (1977). Sufficient statistics and latent trait models. *Psychometrika, 42*, 69–81. doi:10.1007/BF02293746.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561–573. doi:10.1007/BF02293814.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246. doi:10.1037/0033-2909.107.2.238.

Bentler, P. M. (2006). *EQS structural equations program manual*. Encico, CA: Multivariate Software Inc.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen & R. Stine (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage. doi:10.1177/0049124192021002005.

Byrne, B. M. (1995). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming*. Thousand Oaks, CA: Sage.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105. doi:10.1037/h0046016.

Campbell, D. T., & O'Connell, E. J. (1967). Methods factors in multitrait-multimethod matrices: Multiplicative rather than additive? *Multivariate Behavioral Research, 2*, 409–426.

Campbell, D. T., & O'Connell, E. J. (1982). Methods as diluting trait relationships rather than adding irrelevant systematic variance. *New Directions for Methodology of Social and Behavioral Science, 12*, 93–111. doi:10.1207/s15327906mbr0204_1.

Chartrand, J. M., & Robbins, S. B. (1997). *Career factors inventory: Applications and technical guide*. Palo Alto, CA: Consulting Psychologists.

Cole, J. C., Rabin, A. S., Smith, T. M., & Kaufman, A. S. (2004). Development and validation of a Rasch-derived CES-D short form. *Psychological Assessment, 16*, 360–372. doi:10.1037/1040-3590.16.4.360.

Conseil Supérieur de l'Éducation. (2008). *The student's engagement in his educational plan at the college level: A shared responsibility among all college stakeholders (Abridged version)*. Québec, Canada: Author.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Harcourt Brace-Jovanovich.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334. doi:10.1007/BF02310555.

Deal, J. E., Halverson, C. F., Jr., Martin, R. P., Victor, J., & Baker, S. (2007). The inventory of children's individual differences: Development and validation of a short version. *Journal of Personality Assessment, 89*, 162–166. doi:10.1080/00223890701468550.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Thousand Oaks, CA: Sage. doi:10.1111/j.1365-2648.2007.04270_1.x.

Forner, Y. (2000). Indécision et adaptation à l'université [Indecision and adjustment to the university]. *L'Orientation Scolaire et Professionnelle, 29*, 499–517.

Forner, Y. (2007). L'indécision de carrière des adolescents [Adolescents' career indecision]. *Le Travail Humain, 70*, 213–214.

Forner, Y. (2009). *EDV Épreuve de Décision Vocationnelle [EDV Vocational Decision Test]*. Paris, France: Eurotests Éditions.

Forner, Y. (2010). L'évaluation de l'indécision vocationnelle: L'EDV-9 (forme S). [The validation of the EDV-9S]. *L'Orientation Scolaire et Professionnelle, 39*, iv–xxii.

Gati, I., & Asher, I. (2001). The PIC model for career decision making: Prescreening, in-depth exploration, and choice. In F. T. L. Leong & A. Barak (Eds.), *Contemporary models in vocational psychology* (pp. 7–54). Mahwah, NJ: Erlbaum.

Gati, I., Gadassi, R., & Mashiah-Cohen, R. (2012). Career decision-making profiles vs. styles. Convergent and incremental validity. *Journal of Vocational Behavior, 81*, 2–16. doi:10.1016/j.jvb.2012.03.004.

Gati, I., Krausz, M., & Osipow, S. H. (1996). A taxonomy of difficulties in career decision making. *Journal of Counseling Psychology, 43*, 510–526. doi:10.1037/0022-0167.43.4.510.

Germeijs, V., & Verschueren, K. (2006). High school students' career decision-making process: Development and validation of the Study Choice Task Inventory. *Journal of Career Assessment, 14*, 449–471. doi:10.1177/1069072706286510.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality, 37*, 504–528. doi:10.1016/S0092-6566(03)00046-1.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549–576. doi:10.1146/annurev.psych.58.110405.085530.

Graunke, S. S., & Woosley, S. A. (2005). An exploration of the factors that affect the academic success of college sophomores. *College Student Journal, 39*, 367–377.

Holland, J. L., Fritzsche, B. A., & Powell, A. B. (1994). *Self-directed search: Technical manual*. Odessa, FL: Psychological Assessment Resources.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fix indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(10), 1–55. doi:10.1080/10705519909540118.

Kelly, K. R., & Lee, W.-C. (2002). Mapping the domain of career decision problems. *Journal of Vocational Behavior, 61*, 302–326. doi:10.1006/jvbe.2001.1858.

Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London: Routledge.

Las Hayas, C., Quintana, J. M., Padierna, J. A., Bilboa, A., & Munoz, P. (2010). Use of Rasch methodology to develop a short version of the health related quality of life for eating disorders questionnaire: a prospective study. *Health and Quality of Life Outcomes, 8*(29), 1–12. doi:10.1186/1477-7525-8-29.

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*, 79–122. doi:10.1006/jvbe.1994.1027.

Linacre, J. M. (1994). *FACETS. Rasch model computer program (version 2.7) [Computer software and software manual]*. Chicago IL: MESA.

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions, 16*, 878.

Linacre, J. M., & Wright, B. D. (1999). *WINSTEPS: Rasch-model computer program (Version 2.88) [computer software]*. Chicago, IL: MESA.

Marsh, H. W., & Butler, S. (1984). Evaluating reading diagnostic tests: An application of confirmatory factor analysis to multitrait-multimethod data. *Applied Psychological Measurement, 8*, 307–320. doi:10.1177/014662168400800308.

Marsh, H. W., Ellis, L., Parada, L., Richards, G., & Heubeck, B. G. (2005). A short version of the self description questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment, 17*, 81–102. doi:10.1037/1040-3590.17.1.81.

Marsh, H. W., Hau, K.-T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research, 33*, 181–201.

Marsh, H. W., & Hocevar, D. (1988). A new, more powerful approach to multitrait-multimethod analyses: Application of second-order confirmatory factor analysis. *Journal of Applied Psychology, 73*(1), 107–117. doi:10.1037/0021-9010.73.1.107.

Morgan, T., & Ness, D. (2003). Career decision-making difficulties of first-year students. *The Canadian Journal of Career Development, 2*, 33–39.

Muck, P. M., Hell, B., & Gosling, S. D. (2007). Construct validation of a short five-factor model instrument. *European Journal of Psychological Assessment, 23*, 166–175. doi:10.1027/1015-5759.23.3.166.

Neuville, S., & Frenay, M. (2012). La persévérance des étudiants de 1er baccalauréat à la lumière du modèle *expectancy-value*. [Persistence among first-year college students based on the expectancy-value model]. In M. Romainville & C. Michaut (Eds.), *Réussite, échec et abandon dans l'enseignement supérieur* (pp. 157–175). Bruxelles, Belgium: De Boeck.

Nunally, J. C. (Ed.). (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Osipow, S. H. (1987). *Manual for the career decision scale*. Odessa, FL: Psychological Assessment Resources.

Panagiotis, G. S., Sideridis, G. D., Protopapas, A., & Mouzaki, A. (2011). Psychometric evaluation of a receptive vocabulary test for Greek elementary students. *Assessment for Effective Intervention, 37*(1), 34–49. doi:10.1177/1534509411413254.

Picard, F., Boutin, N., & Skakni, I. (2010). La session d'accueil et d'intégration: Un nouveau parcours d'accès à l'enseignement supérieur québécois [The transition session: A new pathway into Québec's higher education system]. *Éducation et Sociétés, 26*(2), 29–43.

Poitras, C., Guay, F., & Ratelle, C. (2012). Using the self-directed search in research: Selecting a representative pool of items to measure vocational interests. *Journal of Career Development, 39*, 186–207. doi:10.1177/0894845310384593.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less. A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*, 203–212. doi:10.1016/j.jrp.2006.02.001.

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27*, 151–161.

Rose, H. A., & Elton, C. F. (1971). Attrition and the vocationally undecided student. *Journal of Vocational Behavior, 1*, 99–103. doi:10.1016/0001-8791(71)90011-X.

Sabourin, S., Valois, P., & Lussier, Y. (2005). Development and validation of a brief version of the Dyadic adjustment scale with nonparametric item analysis model. *Psychological Assessment, 17*, 15–27. doi:10.1037/1040-3590.17.1.15.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation. approach. *Multivariate Behavioral Research, 25*, 173–180. doi:10.1207/s15327906mbr2502_4.

Taylor, K. M. (1982). An investigation of vocational indecision in college students: Correlates and moderators. *Journal of Vocational Behavior, 21*, 318–329. doi:10.1016/0001-8791(82)90040-9.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1–10. doi:10.1007/BF02291170.

van de Schoot, R., Lugtig, P., & Hox, J. (2012). Developmetrics. A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*, 486–492. doi:10.1080/17405629.2012.686740.

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2012). Development and testing of an abbreviated numeracy scale. A Rasch analysis approach. *Journal of Behavioral Decision Making*. Advance online publication. doi:10.1002/bdm.1751.

Wright, B. D., & Linacre, J. (1989). Observations are always ordinal; measurement, however, must be interval. *Archive of Physical Medicine and Rehabilitation, 71*, 857–860.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. Chicago, IL: MESA.