Translation and validation of the Ten-Item-Personality Inventory into Spanish and Catalan

Vanessa Renau[a], Ursula Oberst[a], Samuel D. Gosling[b], Jordi Rusiñol[a], Andrés Chamarro[c]

[a]Universitat Ramon Llull, [b]University of Texas, Austin (USA), [c]Universitat Autònoma, Barcelona

Translation and validation of the Ten-Item Personality Inventory into Spanish and Catalan

Summary. In response to an increasing need for ever-shorter personality instruments, Gosling, Rentfrow, and Swann (2003) developed the Ten-Item-Personality Inventory (TIPI), which measures the dimensions of the Five Factor Model (FFM) using 10 items (two for each dimension) and can be administered in about one minute. In two studies and using a multi-judge (self and observer) and multi-instrument design, we develop Spanish (Castilian) and Catalan versions of the TIPI and evaluate them in terms of internal consistency, test-retest reliability, convergent, discriminant, and content validity, as well as self-observer correlations. Test-retest correlations were strong, and convergence with the NEO-PI-R factors was significant. There were also strong correlations between observer ratings and the participants' self-ratings. Despite some inconsistencies with respect to the Agreeableness scale, the Catalan translation and both translations into Spanish of the original TIPI demonstrated sufficient psychometric properties to warrant use as a Five Factor personality measure when the use of longer instruments is not convenient or possible. Furthermore, as the first translation of a brief standard Big Five Instrument into Catalan, this work should facilitate future research on personality in the Catalan-speaking population.

Keywords: Five-factor model; Ten Item Personality Inventory (TIPI); measurement; validation; psychometric properties

Traducció i validació del Ten-Item Personality Inventory a l'espanyol i al català

Resum. Com a resposta a la necessitat creixent de disposar d'instruments curts de personalitat, Gosling, Rentrow i Swann (2003) van desenvolupar el Ten Item Personality Inventory (TIPI), que mesura les dimensions del model dels cinc grans (Five Factor Model, FFM) utilitzant deu ítems (dos per a cada dimensió) i que es pot administrar aproximadament en un minut. En dos estudis, en els quals hem utilitzat un disseny de multi-jutges (self i observadors) i de multi-instruments, hem desenvolupat una versió espanyola i catalana del TIPI i les hem avaluades en termes de consistència interna, fiabilitat test-retest, validesa convergent, discriminant i de contingut, així com quant a les correlacions amb observadors. Les correlacions test-retest van ser fortes i la convergència amb els factors del NEO-PI-R va ser significativa. Tanmateix, també hi van haver correlacions fortes entre les puntuacions dels observadors i les dels participants mateixos. Tot i que hi ha inconsistències respecte de l'escala d'afabilitat, la versió catalana i les dues versions espanyoles del TIPI original han demostrat tenir propietats psicomètriques suficients per garantir-ne l'ús com a mesura de la personalitat, quan emprar altres instruments més llargs no és convenient o possible. A més, com a primera traducció al català d'una mesura breu dels cinc grans pot facilitar investigacions futures en el camp de la personalitat en població catalanoparlant.

Paraules clau: model de cinc factors; Ten Item Personality Inventory (TIPI); mesura; validació; propietats psicomètriques

Correspondence:
Ursula Oberst, Ph.D.

Facultat de Psicologia, Ciències de l'Educació i de l'Esport,
Universitat Ramon Llull
c/Císter, 34, 08022 Barcelona (Spain)
Phone: (0034) 932533000
Email: ursulao@blanquerna.url.edu

Introduction

Filling out questionnaires and other personality measures can be a time-consuming enterprise. Researchers must balance their wish to assess their participants' personalities adequately and accurately against the need to alleviate the burden on participants and to allocate time to other components of the study (Yarkoni, 2010). There are numerous contexts where extremely short personality measures are beneficial, ranging from Internet-based studies that rely on participant goodwill and large-scale panel studies to other contexts where boredom and fatigue in participants may increase the possibility of careless or random responding (Burisch, 1997; Gosling, Rentfrow, & Swann, 2003).

In response to such needs, several very short instruments have emerged over the past few years to assess personality. Many of them have been based on the Five-Factor Model (FFM), the most widely accepted model of broad personality traits. These instruments have demonstrated sufficient levels of reliability and good convergence with other measures of the FFM, such as the NEO-PI-R (Costa & McCrae, 1992), which many researchers consider to be the gold-standard instrument. The small family of instruments validated against the NEO-PI-R includes three very brief FFM measures: The 20-item Mini-IPIP (Cooper, Smillie, & Corr, 2010) based on the International Personality Item Pool IPIP (Goldberg, 1999), the Ten-Item Big Five Inventory (Rammstedt, 2007; Rammstedt & John, 2007) derived from the longer 44-item Big Five Inventory (John & Srivastava, 1999), and the Ten-Item-Personality Inventory (TIPI; Gosling et al., 2003). The briefest measures of the FFM contain only five items (e.g., Wood & Hampson, 2005).

Credé, Harms, Niehorster, and Gaye-Valentine (2012) recently evaluated eight short measures of the FFM. Their results showed that the highly abbreviated scales can be associated with significant decrements in validity and increased vulnerability to Type 1 and 2 errors, especially in the case of single-item measures. However, the results also indicated that two-item measures represent a substantial improvement over single-item measures with respect to criterion validity. Therefore, we focus on a two-item measure (for 10 items in total) because it represents a good tradeoff between brevity and psychometric performance.

The TIPI, which contains ten items (with each item consisting of two adjectives with similar, but not identical meanings) is by far the most widely used very brief measure (with more than 1100 citations according to Google Scholar), and has proved to be useful in a wide variety of fields in the social sciences (e.g., Back et al., 2010; Batey, Furnham & Safiullina, 2010; Bunevicius, Katkute & Bunevicius, 2008; Caprara, Francescato, Mebane, Sorace, & Vecchione, 2010) and to have adequate psychometric properties (Gosling et al., 2003; Muck, Hell & Gosling, 2007) in psychometric studies that used exploratory as well as confirmatory factor analyses (Ehrhart, Holcombe-Ehrhart, Roesch, Chung-Herrera, Nadler, & Bradshaw, 2009). The TIPI has been translated into numerous other languages, but to our knowledge, only two German translations (Herzberg & Brähler, 2006; Muck et al., 2007), a Dutch translation (Hofmans, Kuppens & Allik, 2008), a Japanese translation (Oshio, Abe, & Cutrone, 2012), and a Spanish translation (Romero, Villar, Gómez-Fraguela, & López-Romero, 2012) have been validated. The purpose of the present paper is not to evaluate the FFM or personality structure more generally, but to describe a translation and validation of Spanish and Catalan versions of the TIPI. The existing Spanish translation (Romero et al., 2012) was published while the present research was being prepared. The current work builds on that instrument in two ways. First, for two of the items, Romero et al. use verbal phrases (more akin to the item structure in the Big Five Inventory; John & Srivastava, 1999) instead of adjectives, whereas the current instrument retains the original adjectival item structure for all the items. More importantly, the present work also includes a translation into Catalan, the first such translation of a brief standard Big Five instrument into that language.

Study 1

Introduction

The goal of Study 1 was to translate the TIPI into Spanish and Catalan, and to assess the psychometric properties of the TIPI-SPA and TIPI-CAT in terms of internal analysis, test-retest reliability, and structural as well as convergent validity. The goal in translating an instrument is to achieve the best possible equivalence with the original meaning of the original items. Doing so is a major challenge because even thorough by-the-book translations may yield different shades of meaning due to the different cultural contexts, especially when the languages in question are etymologically distant. With respect to Spanish, the meaning of many words is different in different countries, or some expressions are more frequently used in one country than in another. For example, relatively large differences can be found when comparing words and phrases used in Spain (Europe) and in Latin American countries; as a result several different Spanish versions of the TIPI have been proposed (see http://homepage.psy.utexas.edu/homepage/faculty/gosling/scales_we.htm).

Here we use "Spanish" to refer to the Spanish spoken on the Iberian Peninsula (also called Castilian or European Spanish), which is slightly different from the Spanish spoken in Latin America; some expressions are even quite different in meaning. Catalan is the local and co-official language spoken in the autonomous region of Catalonia in Spain. Both Catalan and Spanish are members of the Romanic languages, but there are some important differences between them in terms of wording and grammar. There are approximately six million native Catalan speakers and

Catalan is the language used in the public school system in Catalonia, so it is important to develop a Catalan version of the TIPI to allow personality research to be undertaken in that population.

For our studies, three people, two native bilingual (Spanish and Catalan) and one multilingual speakers independently translated the 20 adjectives into Spanish and Catalan, which were then backtranslated into English by a bilingual English teacher. In cases where the backtranslation turned out to be different from the originals, a consensus among the three translators was chosen. Some adjectives turned out to be extremely difficult to translate, because any existing literal translation had a highly negative connotation when applied to oneself; this was the case for some of the "negative" poles of the dimensions, especially for "quarrelsome", "careless", "conventional", and "uncreative". Strongly evaluative items should be avoided because they tend to result in social desirability and other rater biases in self and informant reports (Furr, 2011). To minimize such effects, where possible, the translators selected the less evaluative meaning of the adjectives (see appendix for the complete questionnaire). As in the original TIPI, the translated versions use a 7-point-Likert scale (from disagree strongly to agree strongly).

Method

Recruitment and participants

A total of 118 second-year students of Psychology (Group I) completed two questionnaires during a practical class of Personality Psychology (self-rating of NEO-PI-R and TIPI); the ratings were done in one of the first practical exercises, before the students had any formal knowledge about psychological testing. Each student had to indicate his or her preferred language (Catalan or Spanish) and then received the materials in that language.

Following the procedures devised by Muck et al. (2007), the participants were asked to recruit a volunteer who also completed self-report versions of the two instruments; the volunteer had to be of similar age, different sex, and not a psychology student or psychologist (these "recruits" will be referred to as Group II). This procedure increased the sample size and the diversity of the sample (i.e., they were not all psychology students). The Psychology students received course credit for their task. Using a coding system we associated each student to his or her partner while preserving anonymity; each participant could retrieve his or her individual results by reconstructing this code and thus obtain feedback on the results.

Another group of students of Sport Science and Teacher studies (Group III) provided self-ratings on the TIPI (Catalan or Spanish) only. The final sample consisted of 309 participants (118 in Group I, 108 in Group II, and 83 in Group III). The age and sex distributions are shown Table 1.

Instruments and procedure

Two instruments were used, the NEO-PI-R (in its standard Spanish or Catalan adaptations), and the TIPI in Spanish (TIPI-SPA) or in Catalan (TIPI-CAT). The students of Group I were handed out a complete dossier in their preferred language with detailed instructions about the procedure, two NEO-PI-R questionnaires and two TIPI questionnaires (one for their own self-rating and one for their recruit's self-rating). According to the preferred language, the students (Group I) received either the Spanish or the Catalan version of the NEO-PI-R (Costa & McCrae, 1992) and the Spanish or Catalan version of the TIPI (TIPI-SPA or TIPI-CAT). The students were asked to reply to the NEO-PI-R first. After completing both instruments, the experimenter read aloud the instructions detailed in the dossiers about administering the instruments to the volunteers to be recruited (Group II); the Group II participants received and replied to all materials in the same language as that chosen by their partner in Group I. The students had to turn in the completed dossiers (their recruit's self rating) within two weeks. The students of Sports and Teacher studies (Group III) rated themselves on the TIPI in their preferred language in a classroom context. One month after the first assessment, the Psychology students (Group I) were asked to answer the TIPI again, in the same language as in the first assessment. 80 of them did so (49 in Catalan and 31 in Spanish). Responding to the questionnaires and obtaining ratings by other individuals was part of the mandatory activities for receiving course credit; however, turning in the results was voluntary and anonymous.

Results

Descriptive statistics, internal, and structural analysis

Table 1 shows the descriptive statistics of the scales used in our study. A three-factor analysis of variance (group, language, and gender) for age, the NEO-PI-R and the TIPI scores showed no significant differences for age among the three groups. There were significant differences for gender with respect to Neuroticism[1] ($F = 8.82$, $p < .01$, $df = 1$, $\eta^2 = .04$), Agreeableness ($F = 8.41$, $p < .01$, $df = 1$, $\eta^2 = .04$), and Conscientiousness ($F = 7.75$, $p < .01$, $df = 1$, $\eta^2 = .03$) in the NEO-PI-R, with females scoring higher than males. There were no differences for group or language.

With respect to the TIPI, there were also significant gender differences for Emotional Stability ($F = 13.32$, $p < .01$, $df = 1$, $\eta^2 = .04$) and Conscientiousness ($F = 7.33$, $p < .01$, $df = 1$, $\eta^2 = .02$), but not for Agreeableness, as in the NEO-PI-R. In the case of the TIPI, there were also significant differences for participant type

---

1. The TIPI keys the dimension of Neuroticism (N) in the other direction, i.e. as Emotional Stability (ES). In order to preserve the original nomenclature of both scales, we write Neuroticism when referring to the NEO-PI-R and Emotional Stability when referring to the TIPI.

Table 1. Descriptive Statistics of the NEO-PI-R and the TIPI-CAT and TIPI-SPA scales for study 1

			Students 1				Partners (I)				Other students (III)			
			CAT		SPA		CAT		SPA		CAT		SPA	
			M n = 15	F n = 58	M n = 10	F n = 35	M n = 51	F n = 16	M n = 30	F n = 11	M n = 19	F n = 18	M n = 22	F n=24
AGE		Mean	24.67	20.40	21.30	19.69	21.49	20.88	21.47	21.64	21.74	22.94	21.91	21.92
		SD	2.44	2.22	1.83	4.73	3.07	2.09	2.30	2.73	2.42	3.80	3.82	2.48
NEO-PI-R	N	Mean	76.13	94.79	85.30	97.14	85.92	85.62	78.53	86.91	–	–	–	–
		SD	13.03	21.92	18.87	21.05	18.48	26.10	17.75	18.41	–	–	–	–
	E	Mean	119.13	118.50	123.10	120.49	120.90	129.44	120.83	122.18	–	–	–	–
		SD	13.83	17.19	21.63	19.48	20.32	18.70	14.12	14.70	–	–	–	–
	O	Mean	128.60	122.64	119.80	119.46	137.16	128.56	109.70	120.55	–	–	–	–
		SD	9.42	15.13	18.06	16.57	157.63	20.34	15.57	19.64	–	–	–	–
	A	Mean	110.27	118.57	107.50	115.86	104.37	116.56	110.17	118.00	–	–	–	–
		SD	18.20	19.62	14.03	24.66	20.55	14.65	15.93	16.00	–	–	–	–
	C	Mean	100.93	117.52	104.10	119.03	112.78	116.56	109.97	112.36	–	–	–	–
		SD	18.91	18.44	17.90	24.55	21.37	22.47	19.92	23.78	–	–	–	–
TIPI	ES	Mean	5.10	4.20	4.55	4.13	4.40	4.00	4.80	3.81	4.47	3.92	4.54	3.94
		SD	.89	1.49	1.40	1.35	1.55	1.08	1.09	1.63	1.46	1.21	1.25	1.07
	E	Mean	4.87	4.86	5.20	4.97	5.10	5.40	5.18	5.54	5.21	5.17	4.79	5.23
		SD	1.19	1.35	1.49	1.40	1.28	1.35	1.30	1.03	1.16	.92	1.25	1.52
	O	Mean	5.57	5.40	6.15	5.43	5.42	5.13	4.97	5.32	5.21	5.64	5.11	5.46
		SD	.80	1.05	.78	1.21	1.10	.89	1.09	1.03	1.08	1.17	.84	1.28
	A	Mean	4.33	4.60	4.20	4.66	3.84	3.97	4.42	3.82	4.24	4.78	4.31	4.27
		SD	.70	.74	1.13	.64	1.08	.55	.79	1.23	.67	.93	.65	.69
	C	Mean	4.43	5.21	4.25	5.27	4.72	4.97	4.85	4.77	4.55	5.33	5.52	5.35
		SD	.88	1.23	1.36	1.21	1.22	1.17	1.34	1.40	1.43	1.04	.94	1.16

Note. N=neuroticism, E=extraversion, O=openness, A=agreeableness, C=conscientiousness, ES= emotional stability

(group), with respect to Openness ($F = 3.08$, $p < .05$, $df = 2$, $\eta^2=.02$) and Agreeableness ($F = 6.47$, $p < .01$, $df = 2$, $\eta^2=.04$); a post-hoc analysis (Scheffé) showed that Group I (the Psychology students) and Group III (students of Sports and Teacher studies) presented significantly higher scores in Agreeableness than did Group II (participants recruited by the Psychology students). As in the case of the NEO-PI-R, there were no significant differences for language in the TIPI.

The main focus of this study was the validation of the TIPI in Catalan and in Spanish, so in Table 2 we present our results for the TIPI-SPA and the TIPI-CAT in comparison to the original data (TIPI), the results of the German study (TIPI-G), and those from the Spanish study by Romero et al. (2012).

As can be seen in the table, the values were generally similar; the most notable exception was for the alpha for A, which was low, especially for the Spanish version. However, it should be noted that using alphas in a two-item instrument that is designed to capture a very broad domain is of questionable meaningfulness. In contrast to multiple-item scales that can afford to have high content overlap, a scale with only two items such as the TIPI inevitably creates lower internal consistency

Table 2. Means, standard deviations and Cronbach's s of the TIPI-CAT and TIPI-SPA (study 1), the TIPI-CAT (retest) and TIPI-SPA (retest), as well as the TIPI-SPA-v2 and the observer ratings (study 2) in comparison with the original data (TIPI, Gosling et al., 2003) and the German adaptation (TIPI-G, Muck et al., 2007)

		TIPI	TIPI-G	TIPI (Romero et al., 2012)	TIPI-CAT N=178	TIPI-CAT-R N=49	TIPI-SPA N=134	TIPI-SPA-R N=31	TIPI SPA-v2 N=191	TIPI-OBS N=94
ES	Mean	4.85	5.10	4.01	4.30	3.86	4.35	4.56	4.42	4.63
	SD	1.45	1.20	1.54	1.42	1.31	1.30	1.26	1.25	1.05
		.73	.67	.59	.67	.66	.51	.70	.45	.84
E	Mean	4.87	4.87	4.83	5.03	4.88	5.10	5.26	4.81	5.14
	SD	1.48	1.21	1.4	1.25	1.24	1.35	1.26	1.33	1.05
		.68	.57	.54	.62	.71	.71	.61	.61	.62
O	Mean	5.49	5.49	4.92	5.39	5.06	5.33	5.22	5.07	5.19
	SD	1.06	.97	.94	1.09	1.05	1.12	1.10	1.25	.95
		.45	.54	.48	.53	.43	.51	.44	.55	.707
A	Mean	5.20	5.20	5.73	4.29	4.77	4.38	4.97	5.22	.518
	SD	1.12	.95	.92	.90	.88	.80	.91	.98	.95
		.40	.42	.38	.28	.26	.08	.32	.21	.751
C	Mean	5.85	5.85	5.15	4.90	5.05	5.13	5.10	5.15	5.32
	SD	1.13	.93	1.35	1.25	.88	1.25	1.21	1.30	1.15
		.50	.66	.54	.59	.60	.48	.63	.53	.82

Note. The alphas of observers are extracted from the means of the scores between the different observers

estimates in order to preserve content validity. High alphas could have been achieved by simply choosing lexically similar descriptors for the two items defining each factor. But in cases where broad constructs are being measured, doing so, would improve internal consistency at the expense of construct validity. For these reasons, Gosling et al. (2003) recommend the use of reliability measures such as test-retest reliability and inter-observer-reliability in lieu of internal consistency estimates like Cronbach's alpha. For the same reason, the meaning of exploratory factor analyses (as Romero et al. [2012] do for their adaptation) is hard to interpret because various rating effects (e.g., acquiescence) could interfere with content-based factors. Similarly, Muck et al. (2007) and Hofmans et al. (2008), following Kline (2005), have also discussed the vulnerability to estimation problems of factor models that have only two indicators per factor. However, previous German (Muck et al.) and Dutch (Hofmans et al.) translations calculated a confirmatory factor analysis as an additional piece of evidence for the construct validity of their TIPI adaptations, so we followed their procedure. In our study, the goodness of fit indices for this model were acceptable (see Schreiber, Stage, King, Nora, & Barlow [2006]): 2 = 59.16 (*df* = 25; *p* < .05); 2/*df* = 2.36; comparative fit index (CFI) = .912; root mean square error of approximation (RMSEA) = .066 [.044-.088].

Test-retest reliability

The means, standard deviations, and alphas of the repeated TIPI-CAT and TIPI-SPA are shown in columns 7 and 9 of Table 2. Test-retest correlations for the subscales of each translated instrument are presented in Table 3 (study 1). All test-retest correlations are strong (at least .60) and are significant on a *p* = .001 level.

Convergent correlations

The convergent and discriminant correlations between the NEO-PI-R factors and facets and the corresponding TIPI-CAT factors as well as the individual TIPI-CAT items are shown in table 4; for the TIPI-SPA, results are shown in section (a) of Table 5.

With very few exceptions, all convergent correlations were significant to at least *p*<.01, for both the Catalan and the Spanish version. However, in the Spanish version, the Openness and especially the Agreeableness scales fared less well; item T7 (*comprensiva, afectuosa*) presented no convergent correlation with the NEO-PI-R Agreeableness scale and converged with only two of its subscales. We also computed the intercorrelations among the 10 items in both versions (Table 6 for Catalan, upper section of Table 7 for Spanish).

As can be seen, correlations between the items of the same dimensions were significant and higher than correlations between items of different dimensions, with the exception of the Agreeableness items; in both versions, there was a higher correlation between *comprensiva, amable* (Agreeableness) and *fiable, auto-disciplinada* (Conscientiousness) than between the two Agreeableness items.

Table 3. Test-retest Pearson correlations for Studies 1 and 2

	Study 1		Study 2
	TIPI-SPA	TIPI-CAT	TIPI-v2
ES	.76**	.82**	.56**
E	.81**	.85**	.55**
O	.72**	.70**	.59**
A	.61**	.69**	.64**
C	.77**	.81**	.78**

Note. ** = *p* <. 01

Discussion

The results of the ANOVA showed no differences with respect to group, nor, as expected, with respect to language. The gender differences found in the NEO-PI are as might be expected based on previous research; generally, women tend to score lower in Emotional Stability (higher in Neuroticism) and higher in Agreeableness and Conscientiousness (see Srivastava, John, Gosling, & Potter, 2003, among others).

With the TIPI, things were slightly different. We found the expected differences in gender concerning ES and C, but not for A. There were no differences found for language group, but the participants of Groups I and III, (i.e., the Psychology and the Sports and Teacher Studies students) had significantly higher scores in O and in A than did those in Group II (the recruited volunteers). Unfortunately, there is no further information available on the educational and social context of the volunteers (the Psychology students had only been instructed to use peers-same age, different sex, never studied Psychology), so we cannot determine whether these differences found are due to experimental artifacts (classroom setting for Group I and III versus natural setting for Group II), to an artifact of the instrument, or to real existing personality differences between groups that the NEO-PI-R had failed to detect; it could be the case that Psychology students are more agreeable and open than an average non-selected sample, and it is also possible that in the presence of a teacher or graduate student, the student groups might have wanted to present themselves as more open-minded and more agreeable.

In general, our results show that both the TIPI-SPA and the TIPI-CAT have acceptable psychometric properties. The mean scores as well as internal consistency are similar to the original version and/or the German adaptation, except for the factor Agreeableness, which was considerably lower. Both versions yielded high test-retest reliability. The results of the convergent correlations between the NEO-PI-R factors and facets and the corresponding TIPI dimensions and items were highly satisfactory, because we would not expect a two-item instrument to correlate with all facets of the NEO-PI-R dimensions. Yet, most of the dimensions and individual items did, with the notable exception of Openness and Agreeableness. Actually, Agreeableness item T2 (*critica, combativa*) had been the most difficult to translate; perhaps, being critical or "combative" also has a highly negative social desirability, but all other possible translations seemed to

Vanessa Renau, Ursula Oberst, Samuel D. Gosling, Jordi Rusiñol, Andrés Chamarro

Table 4. Convergent correlations between NEO-PI-R factors and facets and TIPI single items and dimensions (TIPI-CAT)

Study 1 Catalan	T4 (ES)	T9 (ES)	TIPI ES	T1 (E)	T6 (E)	TIPI E	T5 (O)	T10 (O)	TIPI O	T2 (A)	T7 (A)	TIPI A	T3 (C)	T8 (C)	TIPI C
NEO-PI-R															
N	–.42**	–.54**	–.55**	–.17*	–.17*	.02	–.20*	–.04	–.05	.040	.03	.18**	–.13	–.12	.20*
N1	–.40**	–.34**	–.45**	–.15	–.18*	–.19*	–.18*	–.14	–.19*	.07	.22*	.18*	.04	.05	.05
N2	–.39**	–.45**	–.48**	–.10	.02	–.04	–.24**	.10	–.07	–.10	–.35**	–.29**	–.23**	–.14	–.20*
N3	–.32**	–.53**	–.48**	–.17	–.15	–.18*	–.22**	–.08	–.18*	.11	.10	.14	–.08	–.10	–.11
N4	–.20*	–.25**	–.26**	–.23**	–.32**	–.32**	–.15	–.02	–.10	–.04	–.03	–.05	–.04	–.04	–.05
N5	–.21*	–.25**	–.26**	.07	.01	.04	.20*	.17*	.23**	–.04	.06	.01	–.18*	–.22**	–.24**
N6	–.32**	–.52**	–.47**	–.20*	–.13	–.19*	–.23**	–.19*	–.26**	.16	.08	.17	–.16	–.12	–.15
E	–.03	.08	–.20*	.58**	.49**	.61**	.25**	.20*	–.00	–.23*	.11	.02	.03	.07	–.06
E1	.067	.12	.10	.49**	.47**	.55**	.23**	.15	.23**	–.02	.36**	.20*	.16	.11	.14
E2	–.06	–.09	–.09	.45**	.41**	.50**	.06	.09	.09	.02	.17*	.12	.03	.03	.04
E3	–.00	.16	.08	.38**	.46**	.50**	.10	.16	.16	–.34**	–.20*	–.36**	.01	.03	.03
E4	–.08	–.00	–.25	.41**	.36**	.44**	.20*	.16	.22**	–.16	.02	–.10	–.08	.05	–.00
E5	–.07	–.00	–.04	.21*	.18*	.22**	.16	.15	.19*	–.28**	–.18*	–.31**	–.18*	–.11	–.16
E6	.04	.11	.08	.46**	.29**	.42**	.27**	.19*	.27**	.01	.24**	.15	.08	.04	.07
O	–.05	–.05	–.13	.01	–.01	.27**	.17*	.09	.16	.00	.08	–.08	.03	–.06	–.11
O1	.11	.02	.08	.15	.06	.11	.19*	.21*	.25**	–.10	.13	.01	–.06	–.07	–.07
O2	.00	–.25**	–.13	.04	.12	.10	.05	.09	.09	.09	.38**	.30**	.05	.09	.08
O3	.04	–.15	–.05	.25**	.15	.22**	.33**	.28**	.37**	–.08	.43**	.20*	.07	–.06	–.02
O4	.19*	–.05	.08	.18*	.02	.10	.35**	.25**	.36**	.03	.14	.11	–.13	–.10	–.13
O5	.28**	.15	.26**	.09	–.01	.03	.27**	.18*	.27**	–.21*	.04	–.13	–.00	.01	.00
O6	.18*	.12	.17*	.22*	.11	.18*	.33**	.13	.27**	–.09	.23**	.07	.05	.02	.04
A	.15	.17*	.05	.05	–.01	–.10	.09	–.21*	.05	.17*	.50**	.42**	.16	.17*	–.03
A1	.11	.21*	.18*	.28**	.21*	.27**	.09	–.15	–.04	.09	.30**	.25**	.17	.20*	.22*
A2	–.09	.04	–.04	.02	–.04	–.02	–.05	–.19*	–.15	19*	.30**	.31**	.06	.06	.07
A3	.15	.21	.21*	.28**	.17*	.25**	.18*	.07	.15	.05	.52**	.36**	.30**	.16	.25**
A4	.27**	.20*	.27**	–.05	–.09	–.08	–.06	–.28**	–.22*	.27**	.36**	.42**	.22*	.08	.15
A5	.00	.00	.00	–.11	–.22*	–.19*	.03	–.13	–.06	.10	.35**	.30**	.16	.08	.13
A6	–.01	.09	.04	.33	.26**	.34*	.26**	–.05	.12	.05	.51**	.35**	.12	.07	.11
C	.07	.28**	–.15	.02	–.11	.06	–.04	–.14	–.03	–.17	.10	.20*	.44**	.61**	.63**
C1	.14	.21*	.20*	.07	.02	.04	.06	–.02	.02	–.12	.15	.01	.33**	.44**	.46**
C2	–.06	.11	.02	.08	–.07	–.00	–.14	–.16	–.18*	–.08	.03	–.03	.36**	.65**	.62**
C3	.05	.27**	.17*	.02	–.09	–.05	.06	–.21*	–.10	.11	.10	–.01	.39**	.41**	.46**
C4	–.07	.09	.01	.13	.03	.08	–.00	.04	.02	–.08	–.01	–.07	.28**	.42**	.43**
C5	.10	.30**	.22**	.12	–.00	.06	.03	–.05	–.01	–.08	.06	–.02	.38**	.54**	.55**
C6	.10	.30**	.18*	–.21*	–.29**	–.29**	–.05	–.08	–.08	–.02	.14	.07	.37**	.30**	.38**

Note. TIPI ES= Emotional Stability; TIPI E= Extraversion; TIPI O= Openness; TIPI A= Agreeableness; TIPI C= Conscientiousness. NEO-PI-R facets: N1: Anxiety, N2: Angry Hostility, N3: Depression, N4: self-Consciousness, N5: Impulsiveness, N6: Vulnerability, E1: Warmth, E2: Gregariousness, E3: Assertiveness, E4: Activity, E5: Excitement-Seeking, E6: Positive Emotions, O1: Openness to Fantasy, O2: Openness de Aesthetics, O3: Openness to Feelings, O4: Openness to Actions, O5: Openness to Ideas, O6: Openness to Values, A1: Trust, A2: Straightforwardness, A3: Altruism, A4: Compliance, A5: Modesty, A6: Tender-Mindedness, C1: Competence, C2: Order, C3: Dutifulness, C4: Achievement, C5: Self-Discipline, C6: Deliberation. * = $p < .05$. ** = $p < .01$.

have even more negative connotations than the ones that were finally accepted. Apparently, the Dutch and the German versions as well as the adaptation by Romero et al. had similar problems of ambiguity with the Agreeableness scale. In general, we consider the results of our translation and validation of the TIPI into Spanish and Catalan as promising. However, the Agreeableness scores must be interpreted with caution. In response to our concerns about Agreeableness, we decided to try to improve the scales, which was the purpose of Study 2.

Study 2

Introduction

The results of study 1 were promising but we were not satisfied with the results with respect to Agreeableness. As noted above, we are cautious about relying heavily on the Cronbach's alpha reliability estimates; nonetheless, it is worth noting that the alphas for Agree-

ableness were considerably lower than those for the others. More importantly, the convergent correlations between one item of the Agreeableness scale (Item 2: *crítica, combativa*) were lower than for all the other items. So we decided to revise the items of the TIPI-SPA and run a second study to evaluate the new scale.

The choice of the new items was based on two major considerations. First, we interviewed a subgroup of the Spanish speaking students (Group I) about how they had interpreted the items. They informed us that being "critical" was something estimated as highly positive (especially in an academic context), while *combativo* was considered something more negative and undesirable. "Critical" was seen as falling closer to aspects of Extraversion, while *combativo* was associated with being less kind. As a result of the divergent meanings the participants were unsure about how to score themselves on this item. As for the other Agreeableness item (Item 7), the students informed us that the expression *afectuoso* (affectionate) was very much

Table 5. Convergent correlations between NEO-PI-R factors and facets and TIPI single items and dimensions (TIPI-SPA study 1 and study 1)

Study 1	T4 (ES)	T9 (ES)	TIPI ES	T1 (E)	T6 (E)	TIPI E	T5 (O)	T10 (O)	TIPI O	T2 (A)	T7 (A)	TIPI A	T3 (C)	T8 (C)	TIPI C
NEO–PI–R															
N	–.40**	–.30**	–.40**	–.15	–.13	.01	.07	.03	–.04	.11	.02	.13	.04	.07	.20*
N1	–.34**	–.12	–.29**	–.03	–.25*	–.18	.04	–.21*	–.14	.12	.16	.18	.27*	.10	.21
N2	–.48**	–.31**	–.49**	–.04	.02	–.00	–.04	.01	–.01	–.05	–.08	–.09	–.05	.14	.06
N3	–.21*	–.28**	–.30**	–.14	–.17	–.17	.11	–.03	.02	.30**	.09	.26**	.04	.09	.08
N4	–.11	–.04	–.09	–.35**	–.33**	–.38*	–.06	–.10	–.10	.19	–.00	.14	–.03	–.07	–.06
N5	–.26*	–.04**	–.37**	.05	.20	.16	.01	.25*	.20	–.15	–.07	–.15	–.16	–.09	–.14
N6	–.22*	–.40**	–.38**	–.14	–.20	–.19	–.09	–.04	–.07	.17	–.01	.12	–.01	.08	.05
E	–.09	.12	–.16	.48**	.28**	.41**	.27*	.24*	.12	–.15	.05	–.05	–.05	–.12	–.06
E1	–.01	.20	.09	.40**	.25*	.35**	.36**	.25*	.35**	.03	.16	.12	.07	–.07	–.01
E2	–.11	–.03	–.09	.27*	.12	.21	.18	.21	.24*	.00	.12	.08	.05	–.01	.02
E3	–.09	.06	–.02	.37**	.34**	.39**	.00	.12	.10	–.28**	–.02	–.21*	.04	.00	.02
E4	–.23**	.14	–.07	.31**	.28**	.33**	.21*	.18	.23*	–.24**	–.05	–.21	.02	–.02	–.00
E5	–.11	–.01	–.08	.19	.13	.17	.19	.33**	.33**	–.21*	–.03	–.17	–.32**	–.11	–.24*
E6	.07	.30**	.21*	.39**	.16	.29**	.25*	.15	.22*	–.10	–.04	–.09	–.05	–.30**	–.23*
O	.05	–.13	.05	.23*	.03	.30**	.30**	.30**	.35**	–.20	.15	–.16	–.21*	–.34**	–.11
O1	.08	–.14	–.02	.03	.04	.04	.17	.31**	.31**	–.10	.01	–.06	–.23*	–.24*	–.28**
O2	.13	–.07	.04	.13	.01	.07	.21	.15	.21	–.07	.17	.06	–.05	–.30**	–.23*
O3	–.18	–.05	–.15	.39**	.23*	.33**	.30**	.28**	.35**	–.10	.10	–.01	–.18	–.26*	–.27*
O4	–.09	–.10	–.12	.03	–.03	–.01	.44**	.42**	.51**	–.16	.07	–.07	.15	–.02	–.09
O5	.07	.03	.06	.28**	.01	.14	.15	.04	.10	–.20	.18	–.02	–.06	–.21*	–.17
O6	.07	.07	.09	.04	–.05	–.05	.08	.13	.13	–.14	–.03	–.12	–.18	–.20	–.23*
A	.13	.08	.09	.03	–.10	–.07	.10	–.08	–.05	.23*	.15	.05	.08	.06	–.03
A1	.32**	.19	.32*	.07	.07	.08	.03	.06	.06	.16	.02	.24*	.02	.13	–.08
A2	.05	–.07	–.01	–.10	.04	–.07	.07	–.70	–.02	.11	.05	.11	.10	.11	.12
A3	.02	.22*	.14	.34**	.07	.20	.29**	.04	.16	.05	.37**	.26*	.21	–.04	.07
A4	.28	.17	.28**	–.04	–.20	–.14	.13	–.17	–.06	.40**	.11	.35**	.02	.08	.06
A5	.06	–.06	.01	–.09	–.14	–.13	.23*	–.11	.02	.22*	.09	.21*	.06	.12	.11
A6	.10	.14	.14	.17	.01	.08	.19	–.06	.04	.08	.33**	.26*	.13	.01	.07
C	–.02	.28**	.07	.15	–.09	–.11	–.02	–.26*	–.34**	–.02	.16	–.03	.57**	.51**	.63**
C1	.10	.24*	.20	.28*	.01	.13	.03	–.10	–.06	–.17	.13	–.04	.38**	.20	.32**
C2	–.05	.19	.07	.06	–.03	.01	.05	–.23*	–.14	.10	.09	.12	.48**	.70**	.71**
C3	–.10	.22*	.07	.11	–.09	–.01	–.07	–.24*	–.21	.08	.12	.13	.45**	.37**	.49**
C4	–.21	.12	–.07	.10	–.02	.04	–.01	–.12	–.09	–.12	.03	–.06	.38**	.45**	.50**
C5	.03	.56*	.16	.16	–.02	.06	–.04	–.23*	–.19	–.07	.12	.02	.50**	.45**	.59**
C6	.14	.32**	.27*	.03	–.26*	–.16	–.04	–.30**	–.24*	.08	.25*	.21	.41	.17	.33**

associated with expressing feelings of affection and less with being kind and good-natured.

Second, we drew on previous research examining the FFM in Spanish-speaking contexts. Benet-Martínez and John (2000) examined self-ratings on a list of 299 indigenous Spanish (Castilian) personality descriptive adjectives along with defining phrases and compared them with self-ratings on the Spanish Big Five Inventory (BFI; Benet-Martínez and John, 1998). They identified a set 60 of indigenous Spanish personality descriptors containing local, culturally relevant terms exhibiting a five-factor structure. Benet-Martínez and John suggest that these quasi-indigenous personality descriptors can be used as an alternative to imported (translated) personality measures, when the goal is to measure the Big Five domain with indigenous terms, and when one does not need to measure all the facets of the Big Five. Among the highest factor loadings for Agreeableness were *colérica, que se enfada facilmente* (choleric, somebody who gets easily angered -.45) and *amable* (kind; .43).

Based on these considerations, for the revised version of the TIPI-SPA (TIPI-SPA-v2), we changed both Item 2 and Item 7. The descriptors for Item 2 was changed from "*crítica, combativa*" into "*colérica, discutidora*" and Item 7 from "*comprensiva, afectuosa*" into "*comprensiva, amable*".

The main goal of the second study was to test whether the revised items would improve the validity of the scales, especially the Agreeableness scale. We also decided to use the rating of observers as additional measure of validity. We focused on one language only (Spanish) because in the first study, the results for the Catalan version had not shown these problems.

Method

Study 2 took place 9 months after the first study. Similar to Study 1, 120 psychology students (Group IV) were asked to respond to the questionnaires (NEO-PI-R and TIPI-SPA-v2); furthermore, each student had to administer the questionnaires to a recruited volunteer of similar age but opposite sex and who should not be student of Psychology (Group V). Additionally, each student had to find another 5 relatives or friends ("observers") who were willing to rate the student on the TIPI-SPA-v2.

Vanessa Renau, Ursula Oberst, Samuel D. Gosling, Jordi Rusiñol, Andrés Chamarro

Table 5. Convergent correlations between NEO-PI-R factors and facets and TIPI single items and dimensions (TIPI-SPA study 1 and study 2).

Study 2		T4 (ES)	T9 (ES)	TIPI ES	T1 (E)	T6 (E)	TIPI E	T5 (O)	T10 (O)	TIPI O	T2 (A)	T7 (A)	TIPI A	T3 (C)	T8 (C)	TIPI C
NEO–PI–R																
	N	−.36**	−.41**	−.47*	−.28**	−.18*	−.26**	−.22**	−.18*	−.23**	−.23**	−.03	−.21**	−.19*	−.08	−.14
	N1	−.22**	−.34**	−.34**	−.02	−.13	.010	−.10	−.09	.12	.05	.09	.01	.03	−.12	.00
	N2	−.41**	−.15**	−.36**	−.24**	−.07	−.17*	−.24**	−.24**	−.29**	−.46**	−.18*	−.47**	−.14*	−.02	−.12
	N3	−.19*	−.31**	−.31**	−.29**	−.17*	−.26**	−.16*	−.14	−.19*	.13	.06	.11	.18*	−.08	−.10
	N4	−.15*	−.18*	−.20*	−.35**	−.26**	−.37**	−.30**	−.15*	−.26**	−.14*	.04	−.14	−.10	−.03	.03
	N5	−.32**	−.32**	−.40**	.10	.13	.14	.11	.052	.09	−.12	−.05	.13	−.165*	.03	.18*
	N6	−.22*	−.40**	−.38**	−.29**	−.17*	−.26**	−.18*	−.12	−.17*	−.07	.02	.06	−.20**	−.15*	.16*
	E	−.03	−.05	−.05	.48**	.60**	.45*	.43**	.20**	.36**	.13	.02	.12	.07	−.01	.02
	E1	.08	−.01	.05	4.37**	.27**	.40**	.30**	.10	.22**	.21**	.32**	.31**	.20**	.03	.10
	E2	−04	−.09	.08	.17*	.18*	.21*	.22**	.05	.15*	.16*	.05	.16*	.07	.10	.10
	E3	.04	.12	.10	.32**	.33**	.38**	.21**	.13	.20**	.00	−.03	−.01	.12	.03	.08
	E4	−.16*	−.11	−.17*	.34**	.21*	.30**	.30**	.20**	.29**	−.07	−.05	−.08	.05	−.03	.00
	E5	−.11	−.06	−.11	.06	−.02	.02	.15*	.14	.17*	−.06	−.20**	−.13	−.24**	−.24**	−.28
	E6	−.02	−.09	−.06	.45**	.20*	.36**	.32**	.07	.22**	.20**	−.22**	.25**	.16	−.01	.06
	O	.01	−.05	−.02	.25**	.08	.18*	.35**	.45**	.50*	.15*	.26**	.23**	.15*	−.06	.02
	O1	−.11	−.11	−.14	.14	.06	.11	.21*	.44**	.40**	.10	.10	.12	−.06	−.13	−.13
	O2	−.01	−.08	−.05	.16*	−.04	.05	.13	.26**	.24**	.12	.28**	.21**	.18*	.05	.11
	O3	−.08	−.14	−.13	.19**	.10	.16*	.20*	.13	.19*	.04	.25**	.13	.20**	−.01	.08
	O4	.14	.08	.13	.22**	.13	.19**	.46**	.43**	.53**	.20	.17*	.24**	.17*	−.00	.07
	O5	.06	.06	.07	.05	.02	.03	.10	.27**	.24**	−.03	.07	−.00	.00	−.10	−.07
	O6	.08	.02	.06	.26**	.12	.21**	.33**	.27*	.32**	.22**	.18*	.26**	.16*	−.06	.03
	A	.22**	−.03	.12	.11	.07	.10	.13	.06	.11	.25**	.35**	.35*	.22**	.10	.17*
	A1	.21**	.09	.19**	.11	.16*	.16*	.12	.06	.10	.25**	.17*	.28**	.22**	.14	.19**
	A2	.17*	−.01	.10	01	−.15*	−.10	−.04	−.04	−.05	.15*	.18*	.20*	.18*	.16*	.19**
	A3	.09	−.05	.03	.29**	.15*	.24**	.24**	.11	.21**	.26**	.41**	.38**	.25**	.01	.11
	A4	.28**	−.05	.15*	.03	−.04	−.02	−.02	.08	.04	.35**	.26**	.40**	.06	.07	.08
	A5	.11	.06	.12*	−.01	−.01	−.01	.12	.00	.07	.04	.25**	.08	.15*	.07	.12
	A6	−.06	−.06	−.08	.16*	.10	.15*	.07	.03	.06	.10	.28**	.19*	.14	.10	.13
	C	.14*	.19**	.20**	.06	−.14	−.06	−.10	−.13	−.14	.06	.22**	.03	.59**	.60**	.70*
	C1	.13	.16*	.18*	.08	−.09	−.02	−.06	−.12	−.11	−.09	.18*	−.01	.40**	.38**	.45**
	C2	.08	.04	.08	.04	−.13	.07	−.09	−.17*	−.16*	−.08	.19**	−.00	3.8**	.69**	.67**
	C3	.10	.10	.12	.02	−.14*	.09	−.04	−.06	−.06	−.00	.21**	.07	.48**	.37**	.48**
	C4	−.6	.13	.04	.15*	−.02	.06	−.02	−.07	−.06	−.12	.12	−.06	.47**	.36**	.47**
	C5	.11	.18*	.17	.15*	.01	.07	.03	−.05	−.02	−.01	.17*	.05	.56**	.53**	.64**
	C6	.30**	.27**	.35**	−.10	−.22**	−.20**	−.20**	−.10	−.18*	.05	.10	.09	.41**	.39**	.46**

Note. TIPI ES= Emotional Stability; TIPI E= Extraversion; TIPI O= Openness; TIPI A= Agreeableness; TIPI C= Conscientiousness. NEO-PI-R facets: N1: Anxiety, N2: Angry Hostility, N3: Depression, N4: self-Consciousness, N5: Impulsiveness, N6: Vulnerability, E1: Warmth, E2: Gregariousness, E3: Assertiveness, E4: Activity, E5: Excitement-Seeking, E6: Positive Emotions, O1: Openness to Fantasy, O2: Openness de Aesthetics, O3: Openness to Feelings, O4: Openness to Actions, O5: Openness to Ideas, O6: Openness to Values, A1: Trust, A2: Straightforwardness, A3: Altruism, A4: Compliance, A5: Modesty, A6: Tender-Mindedness, C1: Competence, C2: Order, C3: Dutifulness, C4: Achievement, C5: Self-Discipline, C6: Deliberation. * = p < .05. ** = p < . 01

Table 6. Intercorrelations among the 10 TIPI items for TIPI-CAT

					TIPI item					
TIPI item	1	2	3	4	5	6	7	8	9	10
Extraversion										
1.Extravertida, entusiasta	—									
2.Reservada, callada	−.47**	—								
Agreeableness										
3.Comprensiva, afectuosa	.12	−.03	—							
4.Crítica, combativa	.15	.03	−.16*	—						
Conscientiousness										
5.Fiable, auto-discilplinada	.16*	.13	.36**	.13	—					
6.Desorganizada, descuidada	−.06	−.07	−.12	−.15*	−.47**	—				
Emotional Stability										
7.Serena, emocionalment estable	.08	.04	.08	.04	.26**	−16*	—			
8.Ansioso, fácilment alterable	.09	−.06	−.03	.12	−.03	.08	−.51**	—		
Openness to New Experiences										
9.Oberta a noves experiències, polifacètic	.27**	−.07	−20**	.01	−.05	.07	.13	−.08	—	
10.Tradicional, poc imaginativo	−.16*	.02	−.06	−.01	.07	−.18*	.01	.09	−.36**	—

Note. ** = p < .01; * = p < .05

Table 7. Intercorrelations among the 10 TIPI items for TIPI-SPA (Study 1 and study 2)

	TIPI item									
TIPI item (Study 2)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Extraversion										
1. Extravertido, entusiasta	—									
2. Reservado, callado	–.46**	—								
Agreeableness										
3.Comprensiva, amable	.15*	.06	—							
4.Colérica, discutidora	–.11	.05	–.16*	—						
Conscientiousness										
5.Fiable, auto-discilplinada	.07	.09	.33**	–.08	—					
6.Desorganizada, descuidada	–.05	–.08	–.10	.04	–.42**	—				
Emotional Stability										
7.Serena, emocionalmente estable	–.01	.01	.02	–.19**	.25**	–.20**	—			
8.Ansioso, fácilmente alterable	–.11	–.04	–.15*	.48**	–.13	.14	–.29**	—		
Openness to New Experiences										
9.Abierto a nuevas experiencias, polifacético	.37**	–.24**	.03	–.24**	.03	.12	.03	–.10	—	
10.Tradicional, poco imaginativo	–.29**	.24**	.02	.15*	.02	–.12	–.02	.17*	–.38**	—

	TIPI item									
TIPI item (Study 1)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Extraversion										
1. Extravertido, entusiasta	—									
2. Reservado, callado	–.60**	—								
Agreeableness										
3.Comprensiva, afectuosa	.19*	.01	—							
4.Crítica, combativa	.06	–.09	–.04	—						
Conscientiousness										
5.Fiable, auto-discilplinada	.20*	.01	.22**	–.04	—					
6.Desorganizada, descuidada	.09	–.03	.00	.18*	–.35**	—				
Emotional Stability										
7.Serena, emocionalmente estable	.05	.09	.15	–.02	.08	–.08	—			
8.Ansioso, fácilmente alterable	.08	–.07	–.08		.08	.03	–.35**	—		
Openness to New Experiences										
9.Abierto a nuevas experiencias, polifacético	.35**	–.27**	.12	.11	–.03	.06	.03	.06	—	
10.Tradicional, poco imaginativo	–.20*	.30**	.07	–.10	.09	–.09	.19*	–.04	–.36**	—

Note. ** =p < .01; * = p < .05

Responding to the questionnaires and obtaining ratings by other individuals was part of the mandatory activities for receiving course credit; however, turning in the dossiers for this research study was voluntary and anonymous. After one month, 105 dossiers containing the scores for Group IV and Group V were returned. Three dossiers had to be eliminated from the final data set because they presented incomplete data. The final sample consisted of 102 psychology students (Group IV; 17 males and 85 females) and 89 recruits (Group V; 73 males and 16 females), in total, 191 participants. All 102 individuals of Group IV obtained the ratings of five observers each.

Results

Descriptive statistics and internal analysis

Means, SDs, and Cronbach's alpha of the dimensions of the revised TIPI-SPA-v2, shown separately for group and gender, are presented in Table 8.

A two-factor analysis of variance (group and gender) for age, the NEO-PI-R scores, and the TIPI scores showed no significant differences for age; there was a gender difference for the NEO-neuroticism scale (F = 3.978, p = .048, df = 1, η^2=.02) and a difference between groups for Openness (F = 4.384, p = .038, df = 1, η^2=.02) and Agreeableness (F = 7.641, p = .006, df = 1, η^2=.04), with the Psychology students scoring higher than their partners. For the TIPI-SPA-v2 scores, there was only one significant difference between groups: the Psychology students scored higher on Conscientiousness than their recruits (F = 4.682; p = .032, df = 1, η^2=.02). There were no significant differences for the interactions of gender and groups.

The following analyses were run combining Groups IV and V. For the purpose of comparison, we calculated the descriptive statistics (means and standard deviations) as well as Cronbach's alpha of the TIPI-SPA-v2, though we maintain our reservations expressed in Study 1 with respect to the meaningfulness of using alphas in this context. The results are presented in the same table as Study 1 (column 10 of Table 2).

Convergent correlations

The convergent and discriminant correlations between the NEO-PI-R factor and facet scores on the one hand and the dimensions and individual items

Vanessa Renau, Ursula Oberst, Samuel D. Gosling, Jordi Rusiñol, Andrés Chamarro

Table 8. Descriptive statistics for study 2

Study 2			Students		Partners	
			M	**F**	**M**	**F**
			n=17	**n=85**	**n=73**	**n=16**
AGE						
		Mean	20.47	20.35	21.58	21.38
		SD	1.12	2.35	3.38	4.60
NEO-PI-R						
	N	Mean	78.53	93.06	91.07	91.69
		SD	16.85	20.84	19.16	20.03
	E	Mean	113.29	120.69	116.45	115.88
		SD	22.04	15.86	20.18	9.56
	O	Mean	115.00	118.55	108.37	111.00
		SD	17.40	16.86	19.36	13.50
	A	Mean	117.65	117.39	104.99	110.88
		SD	13.79	19.40	18.08	14.31
	C	Mean	114.59	120.22	111.41	113.88
		SD	23.45	20.84	22.37	21.80
TIPI						
	ES	Mean	5.03	4.41	4.36	4.13
		SD	1.20	1.30	1.22	1.01
	E	Mean	4.76	4.75	4.89	4.84
		SD	1.36	1.39	1.34	.94
	O	Mean	4.88	5.16	5.05	4.81
		SD	1.30	1.23	1.27	1.25
	A	Mean	5.06	5.44	5.05	5.00
		SD	1.14	.95	.96	.89
	C	Mean	5.29	5.48	4.79	4.94
		SD	1.20	1.24	1.26	1.45

Note. N=neuroticism, E=extraversion, O=openness, A=agreeableness, C=conscientiousness, ES= emotional stability

of the TIPI-SPA-v2 on the other hand are shown in section (b) of Table 5. For all dimensions of the TIPI-SPA-v2, convergent correlations with their NEO-PI-R counterparts were higher than any discriminant correlations found, with one exception: Item 2 (*colérico, discutidor,* the negative pole of Agreeableness that had been changed) and therefore, overall Agreeableness had the strongest correlation with N2 ("angry hostility") of the NEO-PI-R. (r = -.47, p < .05).

We also computed the intercorrelations among the 10 items. As shown in the lower section of Table 7, correlations between the items of the same dimensions were higher than correlations between items of different dimensions, with the exception of, again, the Agreeableness scale: like in the first study, item 7 (here: *comprensiva, afectuosa)* presented a higher correlation with one of the Conscientiousness items (*fiable, autodisciplinado)*.

Inter-rater reliability

The ratings of five observers for each participant of Group IV were taken as additional indicators of reliability for the TIPI-SPA-v2. The inter-rater reliability (alpha) among the five observers was calculated for the five TIPI-dimensions. The values ranged between .7 and .9 (column 2 of table 9).

Self-other agreement

The means and standard deviations of the five observer ratings were calculated and are presented in the last column of Table 2. Paired t-tests showed no significant differences between the self-ratings of Group IV and the average of the observer ratings, except for Extraversion (t=2.94, p=.00, df=93); the observers rated their acquaintances to be significantly more extraverted than the participants rated themselves. The means of the observers' scores were then correlated with the means of respondents' scores from Group IV. Results are shown in Table 9 (column 3 to 7). Again, the strongest correlations between self and observer ratings are the convergent correlations for the dimensions. However, there was also a strong discriminant correlation between the observer ratings of Emotional Stability and the self-rating of Agreeableness (.50) and vice versa (observer ratings of Agreeableness and self-rating of Emotional Stability [.40]).

Discussion

In Study 2, our main focus was in evaluating a new instrument –the TIPI-SPA-v2– revised on the basis of the findings from Study 1. As in Study 1, we examined the convergent and discriminant correlations of the individual TIPI items as well as the convergent and discriminant correlations of the TIPI dimensions and items with the NEO-PI-R factors and facets. All convergent correlations for the dimensions far exceeded the discriminant correlations, although there were secondary loadings of some individual items on a discriminant factor. The pattern of correlations was superior to that obtained in Study 1, especially with respect to the Openness and the Agreeableness scale (i.e., with more significant convergent correlations and/or higher values), a result that suggests the changes made with respect to the Agreeableness items were somewhat successful. In both studies, test-retest reliability was strong and significant for all dimensions on p = <.01.

Table 9. Inter-rater reliability and correlations between self- and observer ratings

	α observer	TIPI-SPA v2 ES	TIPI-SPA v2 E	TIPI-SPA v2 O	TIPI-SPA v2 A	TIPI-SPA v2 C
TIPI obs ES	.78	.55**	–.17	.05	.50**	.19
TIPI obs E	.81	.05	.48**	.212	.05	.06
TIPI obs O	.75	.17	.20	.45**	.22	.05
TIPI obs A	.77	.40**	–.11	.06	.60**	.09
TIPI obs C	.88	.17	–.23*	–.11	.12	.67**

Note. ** = p < .01.

To broaden the evaluative bases of the study, Study 2 implemented some key methodological changes. Ratings of five acquaintances (observer ratings) for the participants of Group IV were included. With the exception of Extraversion, there was no difference in means of the self and observer ratings. Strong correlations were obtained between the observer ratings and the participants' self-ratings; here again, convergent correlations exceeded the discriminant correlations.

General discussion

For all three measures (TIPI-SPA, TIPI-CAT, and TIPI-SPA-v2), our results were similar to those obtained in the original study of the English-language TIPI (Gosling et al., 2003) and to the translated German instrument (Muck et al., 2007), suggesting that the Spanish and Catalan versions can stand as reasonable alternatives when the use of a longer instrument is not convenient. We consider the second version of the Spanish TIPI psychometrically superior to the first, and the choice of the items linguistically better, so we will limit further discussion to the findings from Study 2. The TIPI-SPA-v2 reached adequate levels in each of the criteria against which it was validated: Convergent and discriminant validity, inter-correlation of the items, test-retest reliability, and convergence between self- and observer ratings. As in the original study of the English-language instrument (Gosling et al., 2003) and the recent Spanish translation (Romero et al., 2012), best results were achieved for Emotional Stability, Extraversion, and Conscientiousness, while Openness and Agreeableness worked less well. The Agreeableness scale deserves special mention. On the basis of a cultural and linguistic analysis the TIPI used in Study 1 was revised for use in Study 2. This revised scale still presented some difficulties. Apart from the expected convergent correlations, we found substantial discriminant correlations for the Agreeableness Item 2 (*colérica, discutidora*) with respect to three criteria: A discriminant correlation of this item with N2, which is not surprising, because N is "angry hostility", and being quarrelsome is related to anger; a higher discriminant than convergent correlations for the two Agreeableness items, and a correlation between the self-ratings of Agreeableness and observer ratings of Emotional Stability and vice versa. These findings indicate a secondary association of at least one item of Agreeableness (the negative pole) with aspects of (low) Emotional Stability. In other words, being quarrelsome is associated with Neuroticism. This association might be due to the instrument itself or to its different adaptations into other languages. Apparently, the Agreeableness scale fared less well in all other adaptations evaluated. Romero et al. (2012) discuss this problem in the context of other short scales, and argue that this dimension is generally difficult to assess with few adjectives or sentences, possibly as a result of social desirability effects. It is also possible that the effects are driven by real existing differences between the way

personality is expressed in the U.S. (where the original TIPI was developed) and Spain. Our study offers some detailed data (e.g., convergent and discriminant correlations of the individual TIPI items) not available in the other Spanish translation so it is not possible to determine whether these problems generalize to other versions of the instrument. However, Benet-Martínez and John (1998) note that Latin cultures (Spanish and Hispanic) value *simpatia*, a concept the authors describe as "(…) the need for interpersonal behaviors that promote smooth and harmonious relationships, such as expressing positive emotions and avoiding interpersonal conflict" (Benet-Martínez & John, 1998, p.729). In their study using the Spanish version of the BFI, Benet-Martínez and John found that two items of the Agreeableness scales of the Spanish BFI had secondary loadings on Extraversion. Benet-Martínez and John used the term "migration" to refer to this phenomenon in which certain items moved toward Extraversion, explaining it with the strong communal values Spaniards attach to Extraversion (in comparison to Anglo-Saxon North Americans) and a more negative view this culture has of Introversion. However, our finding of a secondary association of the negative pole of Agreeableness with the negative pole of Emotional Stability is not inconsistent with those of Benet-Martínez and John who found in their Spanish sample that "lack of anger-proneness" was an important facet of Agreeableness. Together these findings suggest that for Spaniards being quarrelsome has a connotation of low Emotional Stability. Additional research is needed to further examine this possibility.

Conclusions

Despite the abovementioned inconsistencies with respect to the Agreeableness scale, the Catalan translation and the TIPI-SPA-v2 have shown sufficient psychometric properties to warrant use as a Big Five personality measure when the use of instruments with more items is inconvenient. Further studies are needed to examine the predictive and criterion-related validity of our adaptations. However, it is yet to be analyzed why in the Catalan version the Agreeableness items fared well, whereas in the first Spanish version they did not, despite the words being identical. Cultural differences may account for this.

Based on our findings and until further studies may provide more data on this difference, we recommend: a) the use of the TIPI-SPA-v2 (in Spain; its appropriateness for other Spanish speaking countries in Latin America is still to be determined), and b) the use of the TIPI-CAT for Catalan speaking individuals. Both instruments are provided in the Appendix along with scoring instructions.

References

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Fa-

cebook profiles reflect actual personality, not self-idealization. *Psychological Science, 3*, 372-374. http://dx.doi.org/10.1177/0956797609360756

Batey, M., Furnham, A., & Safiullina, X. (2010). Intelligence, general knowledge and personality as predictors of creativity. *Learning and Individual Differences, 5*, 532-535. http://dx.doi.org/10.1016/j.lindif.2010.04.008

Benet-Martínez, V., & John, O. P. (1998). Los Cinco Grades across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology, 75*, 729-750. http://dx.doi.org/10.1037/0022-3514.75.3.729

Benet-Martínez, V., & John, O. (2000). Toward the development of Quasi-Indigeneous Personality Constructs. *American Behavioral Scientist, 44(1)*, 141-157. http://dx.doi.org/10.1177/0002764200044001011

Bunevicius, A., Katkute, A., & Bunevicius, R. (2008). Symptoms of anxiety and depression in medical students: Relationship with big-five personality dimensions and vulnerability to stress. *International Journal of Social Psychiatry, 6*, 494-501. http://dx.doi.org/10.1177/0020764008090843

Burisch, M. (1997). *Test length and validity revisited. European Journal of Personality, 11*, 303-315. http://dx.doi.org/10.1002/(SICI)1099-0984(199711)11:4<303::AID-PER292>3.0.CO;2-#

Caprara, G., Francescato, D., Mebane, M., Sorace, R., & Vecchione, M. (2010). Personality foundations of ideological divide: A comparison of women members of parliament and women voters in Italy. *Political Psychology, 5*, 739-762. http://dx.doi.org/10.1111/j.1467-9221.2010.00780.x

Cooper, A., Smillie, L. D., & Corr, P. J. (2010). A confirmatory factor analysis of the Mini-IPIP Five-Factor Model personality scale. *Personality and Individual Differences, 48,* 688-691. http://dx.doi.org/10.1016/j.paid.2010.01.004

Costa, P. T., Jr., & McCrae, R. R. (1992). NEO PI-R professional manual. Odessa, FL: Psychological Assessment Resources, Inc.Costa, P. T., Jr., & McCrae, R. R. (1992).

Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology, 4*, 874-888. http://dx.doi.org/10.1037/a0027403

Ehrhart, M. G., Holcombe Ehrhart, K. H., Roesch, S. C., Chung-Herrera, B. G., Nadler, K., & Bradshaw, K. (2009). Testing the latent factor structure and construct validity of the Ten-Item Personality Inventory. *Personality and Individual Differences, 8*, 900-905. http://dx.doi.org/10.1016/j.paid.2009.07.012

Furr, M. (2011). *Scale Construction and Psychometrics for Social and Personality Psychology.* London, UK: SAGE Publications Ltd.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mer-

vielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe, vol. 7* (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Gosling Lab Homepage (2012). *Scales we have developed.* Retrieved from: http://homepage.psy.utexas.edu/homepage/faculty/gosling

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. http://dx.doi.org/10.1016/S0092-6566(03)00046-1

Herzberg, P. Y., & Brahler, E. (2006). Assessing the Big-Five personality domains via short forms - A cautionary note and a proposal. *European Journal of Psychological Assessment, 3*, 139-148. http://dx.doi.org/10.1027/1015-5759.22.3.139

Hofmans, J., Kuppens, P., & Allik, J. (2008). Is short in length short in content? An examination of the domain representation of the Ten Item Personality Inventory scales in Dutch language. *Personality and Individual Differences, 8*, 750-755. http://dx.doi.org/0.1016/j.paid.2008.08.0044

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-139). New York: Guilford.

Kline, R. B. (2005). Principles and Practice of Structural Equation Modeling (2nd ed.). New York: Guilford. http://dx.doi.org/10.1177/1049731509336986

Muck, P. M., Hell, B., & Gosling, S. D. (2007). The construct validation of a short five-factor model instrument: A self-peer study on the German adaptation of the Ten-Item Personality Inventory (TIPI-G). *European Journal of Psychological Assessment, 23,* 166-175. http://dx.doi.org/10.1027/1015-5759.23.3.166

Oshio, A., Abe, S., & Cutrone, P. (in press). Development, reliability, and validity of the Japanese version of Ten Item Personality Inventory (TIPI-J). *The Japanese Journal of Personality, 1*, 40-51.

Rammstedt, B. (2007). The 10-Item Big Five Inventory - Norm values and investigation of sociodemographic effects based on a German population representative sample. *European Journal of Psychological Assessment, 3*, 193-201.

http://dx.doi.org/10.1027/1015-5759.23.3.193

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, 1*, 203-212. http://dx.doi.org/10.1016/j.jrp.2006.02.001

Romero, E., Villar, P., Gómez-Fraquela, J.A. &López-Romero, L. (2012). Measuring personality traïts with ultra-short scales: A study of the Ten Item-Personality Inventory (TIPI) in a Spanish sample. *Personality and Individual Differences, 3*, 289-293. doi: 10.1016/j.paid.2012.03.035

Schreiber, J. B., Stage, F. K., King, J., Nora, A., & Barlow, E. A. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A re-

view. *Journal of Education Research*, 99, 323-337. doi: 10.3200/JOER.99.6.323-338

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology, 84,* 1041-1053. http://dx.doi.org/10.1037/0022-3514.84.5.1041

Yarkoni, T. (2010). Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality, 44,* 363-373. http://dx.doi.org/10.1016/j.jrp.2010.04.001

Woods, S. A., & Hampson, S. E. (2005). Measuring the Big Five with single items using a bipolar response scale. *European Journal of Personality, 19,* 373-390. http://dx.doi.org/10.1002/per.542

Resumen

Traducción y validación del Ten-Item Personality Inventory al español y al catalán

Como respuesta a la creciente necesidad de contar con instrumentos cortos de personalidad, Gosling, Rentrow y Swann (2003) desarrollaron el *Ten Item Personality Inventory* (TIPI), que mide las dimensiones del modelo de los cinco grandes (*Five Factor Model*, FFM) utilizando diez ítems (dos para cada dimensión) y que puede administrarse aproximadamente en un minuto. En dos estudios, en los que hemos utilizado un diseño de multi-jueces (*self* y observadores) y de multi-instrumentos, hemos desarrollado una versión española y catalana del TIPI y las hemos evaluado en términos de consistencia interna, fiabilidad test-retest, validez convergente, discriminante y de contenido, así como en cuanto a las correlaciones con observadores. Las correlaciones test-retest fueron fuertes y la convergencia con los factores del NEO-PI-R fue significativa. Sin embargo, también hubo correlaciones fuertes entre las puntuaciones de los observadores y las de los mismos participantes. Aunque hay inconsistencias respecto a la escala de afabilidad, la versión catalana y las dos versiones españolas del TIPI original han demostrado poseer propiedades psicométricas suficientes para garantizar su uso como medida de la personalidad, cuando utilizar otros instrumentos más largos no es conveniente o posible. Además, como primera traducción al catalán de una medida breve de los cinco grandes puede facilitar futuras investigaciones en el campo de la personalidad en población catalanohablante.

Palabras clave: modelo de cinco factores; Ten Item Personality Inventory (TIPI); medida; validación; propiedades psicométricas

Appendix

TIPI-SPA (v2)

Aquí encuentra una serie de rasgos de personalidad que pueden o no referirse a Vd. Por favor, escriba un número al lado de cada par de expresiones para indicar en qué medida está de acuerdo o en desacuerdo con ella. Debería valorar el grado en qué cada par de rasgos corresponde a su persona, aunque una pueda corresponder en más grado que otra.

Me veo a mi mismo/a como a una persona:
1. Extravertida, entusiasta.
2. Colérica, discutidora.
3. Fiable, auto-disciplinada.
4. Ansiosa, fácilmente alterable.
5. Abierta a nuevas experiencias, polifacética.
6. Reservada, callada.
7. Comprensiva, amable.
8. Desorganizada, descuidada.
9. Serena, emocionalmente estable.
10. _Tradicional, poco imaginativa.

Puntuación ("R": invertir puntuación):

Extraversión: 1, 6R; Afabilidad: 2R, 7; Responsabilidad; 3, 8R; Estabilidad Emocional: 4R, 9; Apertura a la Experiencia: 5, 10R.

TIPI-CAT

Aquí trobes una sèrie de trets de personalitat que poden o no referir-se a tu. Si us plau, escriu un número al costat de cada expressió per a indicar en quina mesura està d'acord o en desacord amb ella. Hauràs de valorar el grau en que cada parell de trets correspon a la teva persona, encara que un ho pot fer en major grau que l'altre.

Em veig a mi mateix/a com a una persona:
1. Extravertida, entusiasta.
2. Crítica, combativa.
3. Fiable, auto-disciplinada.
4. Ansiosa, fàcilment alterable.
5. Oberta a noves experiències, polifacètica.
6. Reservada, callada.
7. Comprensiva, afectuosa.
8. Desorganitzada, descuidada.
9. Serena, emocionalment estable.
10. Tradicional, poc imaginativa.

Puntuació ("R": puntuar a l'inversa):

Extraversió: 1, 6R; Afabilitat: 2R, 7; Responsabilitat; 3, 8R; Estabilitat Emocional: 4R, 9; Obertura a la Experiència: 5, 10R.

Scoring the TIPI

Recode the reverse-scored items (2, 4, 6, 8, &10) (i.e. recode a 7 with a 1, a 6 with a 2, etc.); take the average of the two items (the standard item and the recoded reverse-scored item) that make up each scale.