

Cross-cultural validity of the scale for interpersonal behavior

Laura Nota^{1,*}, Willem A. Arrindell², Salvatore Soresi¹, Jan van der Ende³ and Ezio Sanavio⁴

¹ Department of Developmental Psychology and Socialization, University of Padova, Padova, Italy

² Department of Clinical Psychology, University of Groningen, Groningen, The Netherlands

³ Department of Child and Adolescent Psychiatry, Erasmus University, Rotterdam, The Netherlands

⁴ Department of General Psychology, University of Padua, Padua, Italy

Abstract

The Scale for Interpersonal Behavior (SIB) is a 50-item multidimensional measure of difficulty and distress in assertiveness. The SIB assesses negative assertion, expression of and dealing with personal limitations, initiating assertiveness and positive assertion. The SIB was originally developed in the Netherlands. The present study attempted to replicate the original factors with an Italian student sample ($n=995$). The four distress and four performance factors were replicable across two methods of analysis (the multiple group method of confirmatory analysis and Tucker's coefficient of congruence ϕ). The corresponding scales were internally consistent and showed predicted patterns of correlations with a measure of self-efficacy. Sex and age differences in assertiveness were generally negligible. Italian students had higher positive assertion-performance scores than the Dutch and comparable scores on other performance scales; by contrast, the Italian subjects had significantly higher scores on all SIB distress scales than their Dutch equivalents. This was ascribed to the stronger pressure on people in Italian society to behave assertively (Hofstede's National Masculinity score=70) as opposed to the Dutch society (National Masculinity score=14).

Keywords: assertiveness; cross-cultural assessment; factorial validity; multiple group method; reliability.

Introduction

In developing the Scale for Interpersonal Behavior (SIB), its constructors (1, 2) aimed at providing a multidimensional

measure of assertiveness that would yield two types of information for clinical applications and research purposes: the probability of response (performance) and the degree of discomfort (felt anxiety/distress) associated with attempts at self-assertion in a variety of social situations. Four factorially derived dimensions are measured with the 50-item SIB. Each dimension is scored for both distress and performance. The SIB dimensions include: Display of negative feelings or negative assertion (15 items); Expression of and dealing with personal limitations (14 items); Initiating assertiveness (9 items); Positive assertion (8 items). In addition to these subscales, General Assertiveness scores, for both distress and performance (item scores are summated across all 50 items), can be utilized as an indication of a person's level of assertiveness across various situations and various types of response classes.

The SIB has good to excellent test-retest and internal consistency reliabilities (2, 3). In addition, the construct validity of the SIB has been shown to be more than adequate. The SIB has been correlated with different measures that were administered to subjects from several distinct samples. Supporting convergent and divergent validity, associations between the SIB and analogous measures were more stronger in magnitude than those between the SIB and non-analogous scales (2, 3). Four confirmatory factor analyses (2, 4–6) yielded findings in favor of cross-sample invariance of the multidimensional structure underlying the SIB. Sensitivity to detect treatment outcome has also been demonstrated (2).

The SIB was originally developed in the Netherlands and the above-reported findings were mostly based on analyses of data collected with Dutch subjects. In recent years, however, the SIB has been introduced in several different countries where it is mostly used for research purposes. Findings obtained with the US (7), British (8), Turkish (9, 10), Swedish (10), Spanish (11), French (12) and Italian versions (13) of the SIB all support the internal consistency reliabilities of the relevant (sub)scales. Data to support the construct validity of the SIB are also available for the US (6), French (12) and Italian (13) versions of the SIB.

In the first psychometric study with the Italian SIB, Arrindell et al. (13), using data collected among 327 undergraduates enrolled at the University of Florence, studied the test-retest reliability (in a subsample of $n=75$; time-interval: 24–28 days), internal consistency reliability and correlates with the Rathus Assertiveness Scale (RAS) (14). Ranging from 0.70 to 0.91, test-retest correlations were shown to be excellent, as were the internal consistency coefficients which ranged from 0.73 to 0.91. In addition, as anticipated, each and every (sub)scale of the SIB correlated significantly with the RAS: 0.37–0.64 (with distress) and -0.27 to -0.57 (with performance) ($p<0.001$). To date, there are no studies available that have addressed the factorial validity of the SIB with Italian subjects. The present study was set up to extend the first

*Corresponding author: Laura Nota, Department of Developmental Psychology and Socialization, University of Padova, Via Belzoni 80, 35131 Padova, Italy

Fax: +39-49-8278451, E-mail: laura.nota@unipd.it

Submitted July 3, 2010. Revised August 5, 2010. Accepted

August 19, 2010. Previously published online March 1, 2011.

psychometric findings yielded with Italian students by tackling this and other psychometric issues specified below.

In this regard, the SIB was administered to a fresh sample of 995 Italian students, along with, among other measures, a reliable and valid measure of self-efficacy (described in the Methods section). It was hypothesized that (a) the dimensional structure underlying the SIB items would be similar to the structure found with the original Dutch SIB; (b) the different SIB scales would show evidence of internal consistency reliabilities similar to the ones obtained by Arrindell et al. (13) with a previous Italian student sample; (c) the patterns of correlations among the s-SIB (sub)scales would be similar to those observed with the original Dutch SIB (5); (d) in line with previous findings (3), the SIB (sub)scales would show evidence of significant correlations with self-efficacy (+ with performance; – with distress). In addition, the present study sought to explore sex and age differences in self-assessed difficulty and distress in assertiveness. Moreover, the Italian data were compared with those previously published for Dutch university students.

Methods

The Italian student sample comprised 637 female and 358 male subjects with a joint mean age of 17.09 years ($SD=0.18$). The original, Dutch SIB (1, 2), contains 50 items, 46 of which are distributed in a non-overlapping manner across the four subscales described above. The respondent evaluates each item on two separate 5-point [1–5] Likert-type scales, one for discomfort (response options range from ‘not at all’ to ‘extremely’) and the other for the probability of engaging in a specific behavior (‘I never do’ to ‘I always do’). The general/total scale scores are obtained by summing across all 50 items. The subscales comprise 15 (Negative assertion), 14 (Expression of and dealing with personal limitations), 9 (Initiating assertiveness), and 8 (Positive assertion) items (this holds for both distress and performance subscales).

The questionnaire “How much confidence do I have in myself?” (15) is made up of 20 items which gather information on: (a) Confidence in one’s own ability of decision-making (7 items; e.g., “If others knew me better they would say I am one who strongly believes in his/her decisional abilities”; possible score range was 5–35); (b) Confidence in one’s own ability of emotional self-control (6 items; e.g., “I can manage my emotions in an efficacious way”; possible score range was 5–30); (c) Confidence in one’s own ability of completing tasks and activities (3 items; e.g., “If I can’t solve an issue the first time I try, I tend to give it up”; possible score range was 5–15); (d) Confidence in one’s own ability of dealing successfully with different activities and situations (4 items; e.g., “I am so sure of my abilities that sometimes I like to devote myself to very difficult things”; possible score range was 5–20). This instrument requires respondents to indicate on a five-point scale how much the proposed statement describes the individual’s usual way of thinking and behaving (1=not at all; 5=very much). Scale scores were obtained summing item responses corresponding to each of the four factors, after reversing the scores of negatively worded items. Soresi and Nota (15) reported coefficient alpha values for the four factors, respectively, of 0.84, 0.74, 0.72, and 0.77, and test-retest correlation, respectively, of 0.81, 0.65, 0.78, and 0.76. Corresponding reliability estimates in the present study were 0.84, 0.73, 0.72, and 0.77. With regard to construct validity, a series of exploratory and confirmatory factor analyses provide support for this four-factor structure, accounting for 51.89% of the total variance. Concerning discriminant validity, the instrument allows differentiating individuals belonging to

different typologies of deciders: in particular, those individuals who are characterized by indecision and lack of school/career assurance are the ones who present the lowest levels of efficacy beliefs (15).

Each questionnaire set was completed anonymously. To ensure translation equivalence (16) of the Italian questionnaire versions, these were carefully translated and back-translated and checked for inconsistencies prior to being implemented.

Statistical analyses

A descriptive test for the existence of the four hypothetical difficulty and four corresponding distress in assertiveness factors was carried out using the multiple group method (MGM) of confirmatory analysis (17). The MGM is closely related to rotation of component weights to perfect congruence and the cross-validation of component weights (18). Basically, with this method factors obtained with the present sample are compared with a known theoretical structure. In doing so, the theoretical weight matrices which reflect the four primary difficulty or distress dimensions are cast as the hypothetical weight matrices in binary form. This implies that all items rationally defined as comprising a particular dimension (e.g., Negative assertion) are assigned a 1 for that dimension and a 0 for the other three dimensions, which results in a matrix in which each row has one non-zero entry only; the columns of weights thus conformed to item sets theoretically assigned to each difficulty or distress dimension. Next, the binary weight matrices (which are obviously identical) were used to compute two structure matrices: one from the 50×50 correlation matrix of distress items and the other from the 50×50 correlation matrix of performance (difficulty) items. Convergence between the patterns of factor loadings across binary weight matrix and structure matrix provides evidence of factorial invariance. The weights impose correlated factors on the data matrices, i.e., the MGM allows factors to be correlated. These factor correlations provide information with regard to the extent of the relative separate existence of factors. The factors are considered standardized weighted summations of item scores.

In the interpretative process, the strength of each factor should also be considered, requiring that this should be at least $(1/k) \times 100\%$, in which k = the total number of items involved in the analysis (in this case $n=50$) (19). Actual calculations involved were performed with the Simultaneous Components Analysis computer program (20).

For purposes of interpretation loadings with absolute values <0.20 are considered small, loadings from 0.20 to 0.39 as moderate, loadings from 0.40 to 0.70 as high, and loadings >0.70 as very high. In addition, we accepted moderate loadings as salient only if they reflected absolute values of at least medium effect sizes, i.e., ≥ 0.30 (21).

In addition to the MGM, a second method of analysis was employed for determining the cross-cultural invariance of the SIB factors. In doing so, the Varimax-rotated matrices of factor loadings obtained in the present Italian sample and the Dutch sample in which the distress and performance factors were originally identified (1, 2) were compared with one another. This was done with a method which enables the interpretation of similarity/identity between sets of factors derived for identical items from different population samples in terms of indices of factor comparisons. In the present study, Tucker’s coefficient of congruence ϕ , ϕ (22), was employed. Tucker’s ϕ ranges from -1.00 via 0 (and 0.70 which denotes poor similarity) to $+1.00$ (perfect similarity). The lower bound cut-off value for the ϕ -index was set at 0.80, as Haven and ten Berge (23) have empirically demonstrated that sets of factors for which calculated coefficients have been found to be ≥ 0.80 tend to be determined as similar. For essential identity of factors, ten Berge (19) requires that $\phi \geq 0.85$.

A statistical method, developed by Humphreys (24), enabled the computation of the difference, in a single sample, between two cor-

relations when three variables are involved [ra,c vs. rb,c], taking into account ra,b and n.

Results

Descriptive item-statistics indicated that no “difficulty factors” (25) could be expected to emerge from any of the analyses

performed below. In addition, there were no items with zero variance.

Table 1 gives the theoretically relevant component loadings which were obtained through the application of the MGM and the strength of each factor, whereas Table 2 displays the correlations between the components. The findings are obviously depicted for distress and performance separately.

Table 1 Confirmatory analysis using the Multiple Group Method (MGM).

Item no.	Component and item content SIB	Component loadings	
		Distress	Performance
	<i>Negative assertion: Factor strengths</i>	17.39%	16.34%
2.	Telling someone who interrupts you that you find this annoying	0.56	0.64
7.	Telling a friend or an acquaintance that he/she is doing something that annoys you	0.55	0.63
9.	Refusing a request made by a person in authority	0.57	0.56
10.	Asking people to make allowance for the fact that you are afraid of something	0.44	0.49
15.	Telling someone that you think he/she treated you unfairly	0.59	0.61
22.	Refusing a request made by someone you are fond of	0.48	0.40
25.	Asking an acquaintance to help you with a job	0.39	0.41
28.	Asking people to return things they have borrowed	0.52	0.48
31.	Refusing to give money to collections	0.46	0.40
33.	Asking a person to stop doing something that annoys you (for instance, in a train, a restaurant, or a cinema)	0.66	0.67
34.	Protesting when someone jumps the queue	0.60	0.62
39.	Refusing unsatisfactory goods or services (for instance, in a shop or in a restaurant)	0.52	0.51
41.	Discussing with someone your impression that they are trying to avoid you	0.54	0.55
48.	Refusing to lend something to a near acquaintance	0.47	0.43
50.	Insisting that someone does his/her share in a joint task	0.51	0.48
	<i>Expression of and dealing with personal limitations: Factor strengths</i>	16.60%	16.45%
4.	Asking someone to explain something you have not understood	0.44	0.51
11.	Maintaining your opinion in the face of a good friend who disagrees with you	0.49	0.53
12.	Admitting that you are wrong	0.58	0.65
13.	Looking at the person you are talking to	0.46	0.34
17.	Declining a drink, especially if it is offered to you repeatedly	0.43	0.41
18.	Telling someone who has justly criticized you that he/she is right	0.56	0.64
23.	Discussing someone's criticism of something you have done	0.56	0.57
30.	Accepting an offer for help	0.45	0.56
36.	Asking someone whether you have hurt him/her	0.55	0.61
40.	Telling someone who did something for you how pleased you are	0.51	0.60
42.	Saying that you are sorry when you have made a mistake	0.60	0.65
46.	Asking someone to show you the way	0.48	0.51
47.	Asking someone to criticize something you have made	0.55	0.53
	<i>Initiating assertiveness: Factor strengths</i>	17.69%	15.92%
1.	Starting a conversation with a stranger	0.68	0.62
3.	Telling a group of people about something you have experienced	0.62	0.58
5.	Offering an opinion that differs from that of the person you are talking to	0.60	0.58
26.	Putting forward your opinion during a conversation with strangers	0.75	0.72
27.	Joining in the conversation of a group of people	0.75	0.69
32.	Maintaining your own opinion against a person who has a very pronounced opinion	0.58	0.53
38.	Giving your opinion to a person in authority	0.73	0.71
44.	Explaining your philosophy of life	0.55	0.58
45.	Going up to someone in order to make their acquaintance	0.63	0.65
	<i>Positive assertion: Factor strengths</i>	15.28%	15.93%
6.	Acknowledging a compliment about your personal appearance	0.65	0.61
8.	Telling someone that you like him/her	0.57	0.69
16.	Telling someone that you are fond of him/her	0.64	0.64
19.	Acknowledging a compliment on something you have done	0.66	0.63
21.	Starting a conversation with a man/woman you find attractive	0.67	0.59
24.	Saying that you enjoy the experience of being told that you are liked	0.67	0.69
37.	Saying that you enjoy people telling you that they are very fond of you	0.63	0.69
43.	Telling someone that you are very pleased with something you have done	0.56	0.62

Table 2 Correlations between SIB distress components (lower-left triangle) and performance components (upper-right triangle) in the Italian student sample.

	I. Negative assertion	II. Personal limitations	III. Initiating assertiveness	IV. Positive assertion
I. Negative assertion	1.00	0.55	0.56	0.53
II. Personal limitations	0.62	1.00	0.50	0.56
III. Initiating assertiveness	0.63	0.60	1.00	0.57
IV. Positive assertion	0.54	0.57	0.57	1.00

All correlations are of the Pearsonian type and are statistically significant beyond $p < 0.001$ (one-tailed test), $n = 995$. SIB, Scale for Interpersonal Behavior.

From Table 1, it can be seen that each item which was hypothesized to load on its theoretically relevant factor was found to do so in an acceptable manner. Only two items, namely no. 25 ('Asking an acquaintance to help you with a job') and no. 13 ('Looking at the person you are talking to') had moderate but acceptable loadings on their respective distress and performance factors. In addition, all items hypothesized to load on their theoretically relevant factor shared a more variance with that factor than with a non-homologous factor, with the majority of the differences in loadings being statistically significant when applying the Humphreys method. There were no instances in which a loading on a non-homologous component exceeded a loading on a homologous component in magnitude.

The separate distress factors explained 15.28%–17.69% of the total variance. The corresponding data for the separate performance factors were 15.92%–16.45% (Table 1). Each factor strength exceeded the minimum required value of 2%. Cumulatively, the distress and the performance factors explained 33.61% and 35.13%, respectively, of the total variance.

The general patterns of correlations among the SIB factors for these are as follows: (a) significantly positive and strong intercorrelations between the distress factors; (b) significantly positive correlations between the performance factors, with the performance factors having somewhat less variance in common than corresponding distress factors; (c) significantly negative associations between corresponding distress and performance dimensions.

It can be seen from Table 2 that patterns (a) and (b) were more than clearly replicated. The usual pattern of correlations between corresponding distress and performance scales was also replicated: -0.48 (Negative assertion), -0.55 (Personal limitations), -0.73 (Initiating assertiveness), -0.64 (Positive assertion), and -0.58 (General assertiveness). These associations were all very similar across the sexes (all p -values < 0.001 , one-tailed tests).

Tucker's ϕ values for homologous Italian and Dutch distress components were excellent (≥ 0.85) for Negative assertion (0.95), Personal limitations (0.91) and Positive assertion (0.85), and satisfactory (0.80–0.85 range) for Initiating assertiveness (0.84). Indices for homologous Italian and Dutch performance components were excellent for Negative assertion (0.95), Personal limitations (0.94) and Initiating assertiveness (0.86), and satisfactory for Positive assertion (0.83). As one would expect, ϕ values for non-homologous components all pointed to construct dissimilarity (range: 0.50–0.68

for non-homologous distress factors; 0.50–0.60 for non-homologous performance factors).

Reliability findings were similar across the sexes and are therefore reported for the sexes combined (Table 3). Internal consistency reliabilities for the distress and performance scales were all more than satisfactory and ranged from 0.78 to 0.93. The mean inter-item r values (as measures of homogeneity) pertaining to the sets of distress and performance items ranged from 0.20 to 0.36. These data are of particular interest as they indicate that although the sample of items from the SIB scales would probably correlate 0.8–0.9 (based on the alpha values) with another sample of items drawn from the same population, at the same time we are dealing with homogeneous sets of scale items. The acceptable range for the mean inter-item r value has been reported to lie somewhere between 0.1 and 0.5, with the 0.2–0.4 range of intercorrelations offering an acceptable balance between bandwidth on the one hand and fidelity on the other hand (26). Moreover, item-remainder correlations were all satisfactory as they ranged from $+0.16$ to $+0.66$. Nunnally and Bernstein (17) (p. 285) proposed that the lowest acceptable bound for the item-remainder r value should be $+0.15$.

Table 3 Survey of outcome of reliability analyses in the Italian student sample ($n = 995$).

	Cronbach's α	Homogeneity	Item-remainder r values (range)
Distress			
Negative assertion	0.81	0.22	0.28–0.57
Personal limitations	0.79	0.21	0.31–0.51
Initiating assertiveness	0.83	0.36	0.41–0.66
Positive assertion	0.78	0.31	0.40–0.54
Total	0.93	0.21	0.28–0.62
Performance			
Negative assertion	0.81	0.22	0.28–0.59
Personal limitations	0.82	0.24	0.21–0.57
Initiating assertiveness	0.81	0.32	0.38–0.62
Positive assertion	0.80	0.33	0.43–0.58
Total	0.93	0.20	0.16–0.57

In line with predictions and with previous Dutch findings (3), and supporting construct validity, subjects with little confidence in themselves had higher distress scores on all SIB scales than their counterparts who had lots of confidence in themselves: -0.32 (Negative assertion), -0.30 (Personal limitations), -0.46 (Initiating assertiveness), -0.30 (Positive assertion), and -0.42 (General assertiveness). By contrast, subjects with great confidence in themselves had significantly higher scores on all performance scales than their equivalents who had only little confidence in themselves: 0.25 (Negative assertion), 0.23 (Personal limitations), 0.44 (Initiating assertiveness), 0.30 (Positive assertion), and 0.37 (General assertiveness). The associations reflected mostly medium effect sizes (all p -values <0.0001 , one-tailed test). Using the Humphreys method, Initiating assertiveness correlated most strongly ($p < 0.01$, two-tailed test) than any of the other SIB subscales with self-efficacy (this applied for both the distress and the performance subscales). These associations were not too high so as to obstruct the discriminant validity of the SIB measuring constructs.

Biological sex (male=1, female=2) correlated as follows with the distress scales: 0.13 ($p < 0.001$) with Negative assertion, 0.04 (NS) with Personal limitations, 0.12 ($p < 0.001$) with Initiating assertiveness, 0.00 (NS) with Positive assertion, and 0.10 ($p < 0.01$) with General assertiveness. The corresponding correlations with the performance scales were, respectively, -0.05 (NS), 0.04 (NS), -0.07 ($p < 0.05$), 0.07 ($p < 0.05$), and -0.01 (NS). Thus, where significant, the associations denoted small effect sizes.

Neither one of the SIB scales was statistically significantly associated with age (p -values >0.20 , NS). The associations with distress and performance were, respectively, -0.01 and 0.02 (with Negative assertion), -0.01 and 0.02 (with Personal limitations), 0.00 and 0.01 (with Initiating assertiveness), -0.02 and 0.00 (with Positive assertion), and -0.01 and 0.02 (with General assertiveness).

When the frequency distributions of the sexes were compared on the distress scales that were significantly associated with biological sex, these were found not to differ to any meaningful extent.

Arrindell et al. (6) (p. 1474, see Table 6) reported descriptive statistics for 409 Dutch university students. When these were compared with the Italian data, it was found that both groups had comparable scores on most performance scales. Reflecting a sizeable difference (Cohen's $d > 0.5$ SD), however, Italian students had significantly higher Positive assertion-performance scores than their Dutch equivalents ($p < 0.01$). In addition, Italian students had significantly higher scores ($p < 0.01$) on each and every distress scale than their Dutch counterparts. The largest differences were observed on General assertiveness, Initiating assertiveness, Negative assertion, and Personal limitations (large effect sizes).

Discussion

Using an Italian student sample, the SIB was demonstrated to be a factorially valid measure of difficulty and distress in assertiveness. The corresponding scales were all internally

consistent, contained homogeneous item sets, and showed patterns of correlations with one another and with self-efficacy, which were in general agreement with predictions and which supported their construct and discriminant validity.

Scores on the SIB were not meaningfully affected by sex or age. Probably in line with folklore, Italian students had significantly higher Positive affect-performance scores than their Dutch counterparts. Differences on the remaining four performance scales were not marked. However, Italian students had markedly higher scores on each and every distress scale. The higher distress scores of the Italian students as compared with the Dutch could be due to the fact that Italian society is more socially demanding than Dutch society with regard to interpersonal competence and assertive responding in social exchanges. Arrindell et al. (27) speculated that the perceived threats and cognitive appraisals associated with such demands could result in higher levels of fear and distress related to interpersonal functioning. The greater social demands with regard to interpersonal competence and assertive responding in Italy than in the Netherlands are in line with empirical findings by Hofstede (28) indicating that Italy can be typified as a masculine society, whereas the Netherlands can be characterized as a feminine society. Hofstede distinguished between masculine societies in which assertiveness, toughness, individual achievement, material success, and heroism are valued, as opposed to feminine societies in which attitudes of modesty, tenderness, and concern with quality of life are highly valued. For his sample of 53 countries and regions, Hofstede (28) reported the highest National Masculinity (MAS) scores for Japan, Austria, Venezuela, followed by Italy and Switzerland that together were ranked 4th and 5th. By contrast, the Netherlands was ranked 51st, with only Norway and Sweden having even lower MAS scores. Italy had a score on MAS=70, whereas the corresponding score for the Netherlands was MAS=15 (observed range: 5–95). Thus, the generally comparable scores on performance between the two national samples in conjunction with the sizably higher distress scores of the Italian students would lead one to speculate on the basis of the Hofstede findings that there could well be more anxious performers among Italian than among Dutch students. A similar line of reasoning was followed by Arrindell et al. (6), who observed comparable or only small differences in performance scores between United States and Dutch university students, as opposed to meaningfully higher distress scores for the US subjects. The US too, with a score on MAS=62 (ranked 15th), is among the masculine countries. Close inspection of the US SIB norms (6) and a general comparison with the Italian SIB norms revealed that, with exception of the clearly higher Positive assertion-performance scores for the Italian subjects, there are no substantial differences between students from Italy and the US on the 50-item SIB.

All in all, the findings presented here support the cross-cultural reliability and validity of the 50-item SIB. Similar findings were recently yielded with its shorter counterpart SIB comprising 25 items (29). With the availability of positive results such as these, investigations with both forms of the SIB in countries such as France and Taiwan can move ahead with greater confidence.

References

1. Arrindell WA, de Groot PM, Walburg JA. De Schaal voor Interpersoonlijk Gedrag (SIG); Handleiding Deel I [Scale for Interpersonal Behavior]. Lisse, Netherlands: Swets Zeitlinger, 1984.
2. Arrindell WA, van der Ende J. Cross sample invariance of the structure of self reported distress and difficulty in assertiveness. *Adv Behav Res Ther* 1985;5:205–43.
3. Arrindell WA, Sanderman R, Ranchor A. The Scale for Interpersonal Behavior and the Wolpe Lazarus Assertiveness Scale: a correlational comparison in a non clinical sample. *Pers Individ Dif* 1990;11:509–13.
4. Arrindell WA, Sanderman R, van der Molen H, van der Ende J, Mersch PP. The structure of assertiveness: a confirmatory approach. *Behav Res Ther* 1988;26:337–9.
5. Arrindell WA, van der Ende J, Sanderman R, Oosterhof L, Stewart R, Lingsma MM. Normative studies with the Scale for Interpersonal Behaviour (SIB): I. Nonpsychiatric social skills trainees. *Pers Individ Dif* 1999;27:417–31.
6. Arrindell WA, Bridges KR, van der Ende J, St Lawrence JS, Gray-Shellberg L, Harnish R, Rogers R, Sanderman R. Normative studies with the Scale for Interpersonal Behaviour (SIB): II. US students. A cross-cultural comparison with Dutch data. *Behav Res Ther* 2001;39:1461–79.
7. Bridges KR, Sanderman R, Breukers P, Ranchor A, Arrindell WA. Sex differences in assertiveness on the U.S. version of the Scale for Interpersonal Behavior. *Pers Individ Dif* 1991;12:1239–43.
8. Gilbert P, Allan S. Assertiveness, submissive behaviour and social comparison. *Br J Clin Psychol* 1994;33:295–306.
9. Eskin M. Reliability of the Turkish version of the Perceived Social Support from Friends and Family Scales, Scale for Interpersonal Behavior, and Suicide Probability Scale. *J Clin Psychol* 1993;49:515–22.
10. Eskin M. Cross-cultural gender differences in the psychosocial correlates of current adolescent suicidal ideation. *J Gend Culture Health* 1996;1:189–205.
11. Arrindell WA, Vergara AI, Torres B, Caballo VE, Sanderman R, Calvo MG, et al. Gender roles in relation to assertiveness and Eysenckian personality dimensions: replication with a Spanish population sample. *Sex Roles* 1997;36:79–92.
12. Bouvard M, Arrindell WA, Guérin J, Bouchard C, Rion AC, Ducottet E, et al. Psychometric appraisal of the Scale for Interpersonal Behavior (SIB) in France. *Behav Res Ther* 1999;37:741–62.
13. Arrindell WA, Bartolini C, Sanavio E. Versione italiana della 'Scale for Interpersonal Behavior': fedeltà, stabilità e validità concorrente [Italian version of Scale for Interpersonal Behavior: fidelity, reliability and concurrent validity]. *Psicoterapia Cognitiva e Comportamentale [Psychother Cogn Behav]* 1999;5:99–107 (in Italian).
14. Rathus SA. A 30-item schedule for assessing assertive behavior. *Behav Ther* 1973;4:398–406.
15. Soresi S, Nota L. Clipper: Autoefficacia e decision making [Clipper: self-efficacy and decision making]. Firenze, Italy: Giunti-Organizzazioni Speciali, 2003 (in Italian).
16. Lonner WJ, Berry JW. Field methods in cross-cultural research. Beverly Hills, CA: Sage, 1986.
17. Nunnally JC, Bernstein IH. Psychometric theory, 3rd ed. New York: McGraw-Hill, 1994.
18. ten Berge JM. The Kaiser, Hunka and Bianchini factor similarity coefficients: a cautionary note. *Multivar Behav Res* 1996;31:1–6.
19. ten Berge JM. Rotation to perfect congruence and the cross-validation of component weights across populations. *Multivar Behav Res* 1986;21:41–64.
20. Kiers HA. SCA: a program for simultaneous. Components analysis of variables measured in two or more populations. User's manual. Groningen, The Netherlands: iec ProGamma, 1990.
21. Cohen J. A power primer. *Psychol Bull* 1992;112:155–9.
22. Tucker LR. A method for synthesis of factor analysis studies. Personnel Research Section Report No. 984. Washington, DC: Department of the Army, 1951.
23. Haven S, ten Berge JM. Tucker's coefficient of congruence as a measure of factorial invariance: an empirical study. *Heymans Bulletins, HB 77-290 EX*. Groningen, The Netherlands: University of Groningen, 1977.
24. Humphreys LG. Differences between correlations in a single sample: a correction and amplification. *Psychol Rep* 1978;43:657–8.
25. Gorsuch RL. Factor analysis. Hillsdale, NJ: LEA, 1983.
26. Briggs SR, Cheek JM. The role of factor analysis in the development and evaluation of personality scales. *J Personality* 1986;54:106–48.
27. Arrindell WA, Kolk AM, Pickersgill MJ, Hageman WJ. Biological sex, sex role orientation, masculine sex role stress, dissimulation and self-reported fears. *Adv Behav Res Ther* 1993;15:103–46.
28. Hofstede G. Culture's consequences: comparing values, behaviors, institutions and organizations across nations. Thousand Oaks, CA: Sage, 2001.
29. Arrindell WA, Sanavio E, Sica C. Introducing a short form version of the Scale for Interpersonal Behaviour (s-SIB) for use in Italy. *Psicoterapia Cognitiva e Comportamentale (Psychother Cogn Behav)* 2002;8:3–18 (in Italian).