

The Psychometric Structure of the Spanish Language Version of The Iowa-Netherlands Comparison Orientation Measure in Spain and Chile

Abraham Buunk^{1,2} , Dick Barelds¹, M. Alfonso Urzúa³, Rosario Zurriaga⁴, Pilar González-Navarro⁴, Pieter D. Dijkstra⁵ and Frederick Gibbons⁶

¹ University of Groningen (The Netherlands)

² Netherlands Interdisciplinary Demographic Institute (The Netherlands)

³ Universidad Católica del Norte (Chile)

⁴ Universidad de Valencia (Spain)

⁵ Private practice (The Netherlands)

⁶ University of Connecticut (USA)

Abstract. The present study investigated the structure of the Spanish version of the Iowa-Netherlands Comparison Orientation Measure (INCOM-E), an 11-item measure that assesses individual differences in *social comparison orientation* (SCO), i.e., the extent to which people compare themselves with others. Data came from samples from Spain ($n = 1,133$) and Chile ($n = 2,757$). Confirmatory Factor Analyses and Mokken Scale Analyses supported in both samples not the assumed two-factor structure, but a single factor structure, consisting of eight items. The resulting eight-item version of the INCOM-E was reliable in both samples, according to the Gutmann's λ_2 (.82 in Spain and .83 in Chile), and correlated very strongly with the full-length INCOM-E (.93 in Spain and .97 in Chile). In both samples, there were significant sex differences, $ps < .001$ with small effect sizes, η^2 in both samples = .01, but in the Spanish sample women scored higher, and in the Chilean sample men scored higher in SCO. The relationship with age was negative and significant ($ps < .001$) in both samples, albeit small ($r = .22$ in Spain and .13 in Chile). Based on the present research, it is advised to use the shortened eight-item version of the INCOM-E in Spanish speaking countries.

Received 13 September 2018; Revised 3 December 2019; Accepted 20 February 2020

Keywords: INCOM-E, psychometric analysis, social comparison orientation, Spanish

It has often been observed that many individuals are reluctant to admit engaging in social comparisons (e.g., Helgeson & Taylor, 1993). For example, in a series of studies, Wilson and Ross (2000) found that in open-ended descriptions, social comparisons were much less frequently used to characterize one-self than were comparisons with one's own self in the past. In fact, social comparisons are often considered inappropriate, especially when they result in socially undesirable emotions, such as envy and *Schadenfreude* (Buckley, 2014). However, according to Gilbert, Price, and Allan (1995) social comparison activity is phylogenetically very old, and can be recognized in many species. Gilbert et al. suggested that the process of social comparison stems from the need to assess one's power and strength compared to that of one's competitors. In the course of evolution, this tendency to compare oneself with others has undoubtedly become very important and elaborate for humans as they developed a complex social life. As suggested by Beach and Tesser (2000), as *Homo sapiens*

began to emerge as a distinct species, there was a shift toward more specialization within groups, and this required the ability to assess the domains in which one could specialize in order to enhance one's status and reproductive opportunities. Social comparison facilitates such an assessment. Also, in modern society the importance of social comparison for human activity cannot be overstated, and may even be higher than ever before. Due to modern technology, such as the Internet, mobile phones and television, individuals are more than ever confronted with information about others, offering them not only the opportunity to communicate with each other whenever they want, but also the opportunity to compare themselves more frequently and with a broader range of individuals than, for instance, half a century ago (Nesi & Prinstein, 2015).

How to cite this article:

Buunk, A., Barelds, D., Alfonso Urzúa, M., Zurriaga, R., González-Navarro, P., Dijkstra, Pieter D., & Gibbons, F. (2020). The psychometric structure of the Spanish language version of the Iowa-Netherlands comparison orientation measure in Spain and Chile. *The Spanish Journal of Psychology*, 23. e9. Doi:10.1017/SJP.2020.1

Correspondence concerning this article should be addressed to Abraham Buunk, University of Groningen, Groningen (The Netherlands).

E-mail: a.buunk@rug.nl

It was Festinger (1954) who in his classic paper introduced the term *social comparison* into social psychology. According to Festinger, people have a fundamental desire to evaluate their opinions and abilities through comparison with others when objective information is unavailable. As noted by Buunk and Gibbons (2006), while Festinger's original theory on social comparison had a restricted focus on the comparison of abilities and opinions, over the past decades work on social comparison has undergone numerous transitions and reformulations, and has developed into a lively, varied and complex area of research. This research encompasses many different paradigms, approaches, and applications (e.g., Suls & Wheeler, 2000), and has shown the importance of social comparison processes for human functioning with regard to, for instance, well-being, achievements, interpersonal relationships and work (e.g., Buunk & Dijkstra, 2017).

Social Comparison Orientation: English and Dutch Versions of the INCOM

Although social comparison is a basic social process, it has often been observed that individuals may differ in their disposition to compare themselves with others. For instance, Diener and Fujita (1997) suggested, "... making any comparisons at all, may often be a function of one's personality." (p. 349). According to Gibbons and Buunk (1999), the extent to which and the frequency with which people compare themselves with others varies indeed from one individual to the next, and this reflects an individual difference variable that was labeled *social comparison orientation* (SCO). In order to measure SCO, Gibbons and Buunk (1999) developed simultaneously a Dutch and English version of the Iowa-Netherlands Comparison Orientation Measure (INCOM), an 11-item scale including items such as 'I always like to know what others in a similar situation would do' and 'I often compare myself with others with respect to what I have accomplished in life'. The reliability and validity of the INCOM was examined in 22 studies in The Netherlands and the United States, showing that, in these countries, the INCOM is a valid and reliable measure of SCO (α 's ranging from .77 – .85). In an initial series of studies it was found that, relative to individuals with a low SCO, individuals with a high SCO were higher in public and private self-consciousness, higher in interpersonal orientation, somewhat higher in neuroticism, and somewhat lower in self-esteem. In addition, they tended to seek out more comparisons, spend more time engaging in comparisons, and experience stronger feelings from comparing themselves with others (Gibbons & Buunk, 1999). On the basis on these and subsequent studies, Buunk and Gibbons (2006) concluded over a decade ago that those with a high SCO

(i.e., those who score high on the INCOM) are characterized by a combination of a high accessibility and awareness of the self, an interest in what others feel and think, and some degree of negative affectivity and self-uncertainty. Various more recent studies have substantiated these conclusions (e.g., Neff & Vonk, 2009). For example, Vogel, Rose, Okdie, Eckles, and Franz (2015) found that individuals with a high SCO exhibited heavier Facebook use than those low in SCO, and showed lower self-esteem, poorer self-perceptions, and more negative affect than individuals with a low SCO after engaging in brief social comparisons on Facebook.

In addition, numerous studies in the past decades have shown that SCO as measured with the INCOM is an important moderator variable because those high in SCO are – often in complex ways – more and differently affected by upward and downward comparisons than those low in SCO. These moderating effects have been observed in many different domains, including the evaluation of one's attractiveness, the quality of life of patients, marital satisfaction, depression and the perception of health risks (Buunk & Gibbons, 2006; see for a review Buunk, Gibbons, Dijkstra, & Zlatan, *in press*). To give just a few examples, Gibbons and his colleagues found in various studies that only students high in SCO were in their risk perceptions and risky behavior affected by information about how others behaved (e.g., Gibbons, Lane, Gerrard, Pomery, & Lautrup, 2002; Stock, Gibbons, Beekman, & Gerrard, 2015). In a similar vein, with increasing levels of SCO, the quality of life of cancer patients improved over a period of three months as a consequence of being exposed to a tape with patients who showed good coping, but not in response to other tapes (Buunk et al., 2011).

The INCOM-E as a Spanish version of the INCOM

Over a decade ago, Buunk, Belmonte, Peiró, Zurriaga, and Gibbons (2005) reported the results of a psychometric evaluation of the INCOM-E, the Spanish language version of the INCOM. In two studies (among 212 students and 782 employees from primary health care centers), the reliability of the INCOM-E was found to be good ($\alpha = .80$). The INCOM-E did not correlate with social desirability at all ($r = .04$, $p = ns$; Buunk et al., 2005), and the correlations with other variables were very similar to those found by Gibbons and Buunk (1999). Individuals with a high SCO were higher in public and private self-consciousness, higher in interpersonal orientation, somewhat higher in neuroticism, although not lower in self-esteem. Particularly relevant for the present paper, moderating effects of SCO have also been observed in a number of studies in Spanish speaking countries, particularly in organizational settings. For example, in a study among Spanish nurses,

Buunk, Zurriaga, and Peíro (2010) found that especially among individuals with a high SCO, the frequency of social comparisons was a predictor of burnout nine to ten months later. Another study in Spain showed that especially among students high in SCO students in their final year of study exposed to a fictitious interview with a new graduate experienced more inspiration and proactive career behavior after exposure to fictitious interview with a former student who was successful in the job market (Buunk, Peiró, & Griffioen, 2007). A study among Spanish physicians showed that these individuals tended to interpret social comparisons at work in a more negative way as they were higher in SCO (Buunk, Zurriaga, Peiró, Nauta, & Gosálvez, 2005). In Argentina, it has been found that in work settings those high in SCO respond with more jealousy in response to a colleague who has a close relationship with one's supervisor (Buunk, Aan 't Goor, & Solano, 2010). In addition, though not a moderating effect, a study in Chile showed that SCO decreases with age (Urzúa, Zúñiga, & Buunk, 2012). All these findings illustrate that as in The Netherlands and the US, also in Spanish speaking countries SCO is a relevant construct that is in meaningful ways related to other variables.

The Psychometric Structure of the INCOM and INCOM-E

In the original studies on the INCOM and INCOM-E, similar factor structures were obtained. By means of exploratory principal components analyses, Gibbons and Buunk (1999) found two factors underlying the 11-item scale: One they labeled 'ability comparisons' and one they labeled 'opinion comparisons'. Confirmatory analyses indicated that the INCOM indeed comprises of two distinguishable factors that are very highly related (in their study the correlation between the two factors was $r = .79$). The factor structure of the INCOM-E was nearly identical to this one, with the same two factors, and similar amounts of variance explained by both factors.

Recently, Schneider and Schupp (2014) evaluated the psychometric properties of the INCOM in the German language, supporting the existence of the above mentioned two factors in the German language version of the INCOM. However, their analyses also pointed to some problems with the INCOM, more particularly with regard to the INCOM's two reversely coded items. According to Schneider and Schupp (2014) these reverse-coded items form either an independent component within social comparison orientation, or load with only minor values on the first factor ('ability comparisons'). However, they also suggested that opening up a third dimension of social comparison does not seem justified on the basis of theoretical reasoning,

and that also methodological arguments are not convincing enough to support a three factor structure. These authors suggested that the low factor loadings of these items may be the result of their indistinct wording and propose a shortened two-factor scale comprising six items.

The Present Study

The present study elaborates on previous studies, by investigating in more detail the structure of the Spanish version of the INCOM-E in two large samples from Spain and Chile. Results from large samples are quite robust and reliable. Special attention is given to the functioning of the two reversely coded items in the INCOM-E that have been proven problematic in the abovementioned analyses of the German version of the INCOM.

STUDY 1: Spain

Method

Participants

A total of 1,133 Spanish workers (44.6% men and 55.4 % women) in different organizations were surveyed. The mean age of respondents was 37.7 years ($SD = 11.7$, range 18–65). Of the participants 12.3% had only primary education, 14% had finished high school, 15.9% had completed vocational school and 57.1% had university qualifications such as bachelor, masters and PhD. degree. The respondents worked in different professional sectors, following the classification of the Spanish Ministry of Labour and Immigration: Administration and management (26.1%); trade and marketing (15%); cultural and communal services (15.4%); hotel and tourism (4.4%); computers and communications (5.9%); health (11.4%); installation and maintenance (9.4%); transport and maintenance of vehicles (1.7%); students (1.9%); unemployed (1.3%), and others sectors such as graphic arts, personal image, safety and environment (2.3%); missing values (5.2%). The overall sample therefore is large and heterogeneous with regard to several relevant background variables.

Procedure

Data collection was carried out by self-report questionnaires that were completed voluntarily by the participants in the presence of the researcher after providing their informed consent. Participants received instructions and information on the procedure in which the questionnaire had to be filled out. If necessary, the researcher provided additional explanations. The researchers emphasized that participation in the study was voluntary, and the confidentiality of the

information provided was guaranteed. No findings from this sample have been published before. The research was approved by the Ethics Committee of the Faculty of Psychology of the University of Valencia.

Instrument

Like the Dutch and English versions, the Spanish version of the INCOM (INCOM-E; Buunk et al., 2005), consists of 11 items, of which two (Items 6 and 10 in Table 1) are reversely coded. The items were answered on five-point scales (1 = *strongly disagree*, 5 = *strongly agree*). Note that the item order is slightly different from the original English language version of the INCOM (Gibbons & Buunk, 1999).

Data analysis

We examined the item characteristics, next executed a Confirmatory Factor Analysis (CFA), and finally conducted a Mokken Scale Analysis (MSA; Mokken, 1971).

Results

First, we calculated item characteristics (means, standard deviations, skewness and kurtosis; see Table 1). The items all appeared to be normally distributed. Next, we examined the fit of the a priori structure of the INCOM by means of a confirmatory factor analysis (CFA) in LISREL. For this purpose, we tested a hierarchical model with two factors (comparisons of abilities and comparison of opinions; cf. Gibbons & Buunk, 1999) and one higher order factor (social comparisons). The fit of this model was poor in this Spanish sample (see Table 2), with a RMSEA value well above and a CFI value well below the criteria commonly referred to as being indicative of a good model fit. Kline (2005), for example, suggests that a RMSEA value $\leq .05$ indicates a close fit, values between .05 and .08 a reasonable fit, and values $\geq .10$ a poor fit. For the CFI, Kline (2005) suggests that values $> .90$ might indicate a reasonably good fit. Hu and Bentler (1999) suggest that CFI values $\geq .95$ indicate a good fit. The SRMR should preferably be smaller than .10 (Kline, 2005), which was the case in this sample. Correlation between the two factors was $r = .39$. Closer examination of the factor loadings (see Table 1) revealed that both Item 6 and Item 10 (the reversely coded items of the INCOM-E) had very low loadings on the two a priori factors. Based on the poor fit indices, an alternative model was tested with just one factor (general SCO), but the fit of this model was similarly poor: $\chi^2(44) = 137.45$, $p < .001$, RMSEA = .163, 90% CI [.16, .17], CFI = .81, SRMR = .099. These results indicate a poor fit of both the two-factor solution and the one-factor solution, mainly because of the poor functioning

of the two reversely coded items. In addition, the modification indices suggest that Item 9 was also not functioning very well in the scale (see also Table 2; Item 9 had a relatively low factor loading compared to the other positively formulated items).

Because the CFA did not support the previously reported structure of the INCOM in Spanish, it was decided to examine if any other subscales might be present in the Spanish INCOM in the present sample (other than comparison of abilities and comparison of opinions). For this purpose, we conducted a Mokken Scale Analysis (MSA; Mokken, 1971). MSA is a nonparametric item-response theory-based method that can be used to exploratively examine the dimensionality of an instrument (e.g., Sijtsma, Meijer, & van der Ark, 2011; Sijtsma & Molenaar, 2002). Nonparametric methods, such as MSA, are generally based on less restrictive assumptions than parametric methods (e.g., Junker & Sijtsma, 2001), and are therefore more likely to fit empirical data than parametric methods (Smits, Timmerman, & Meijer, 2012). Another advantage of nonparametric methods such as MSA is that these are less sensitive to small sample sizes than parametric methods (Smits et al., 2012; Stout, 2001). The MSA was conducted in STATA, using a module to conduct MSA (see Hardouin, Bonnaud-Antignac, & Sébille, 2011).

The MSA revealed a moderate to weak Mokken scale (e.g., Sijtsma & Molenaar, 2002), $H = .37$, comprising of all items except the two reversely coded items. Of the nine remaining items in this scale, Item 9 had an H_i of exactly .30, which was used as a threshold for including items in the scale. The MSA suggested that there are no subscales within the INCOM-E, and confirmed the poor functioning of the two reversely coded items. Also, Item 9 was not functioning very well. Based on these results, it was decided to drop Item 6 and 10 (the reversely coded items), as well as Item 9 from the Spanish INCOM, consequently reducing the INCOM-E, to eight items. Reliability of this eight item scale was estimated by computing lambda-2 coefficients. Gutmann's lambda-2 is a reliability estimate that is similar to the more widely used Cronbach's alfa, but is a better estimate of the reliability of the test (Sijtsma, 2009). By definition, lambda-2 is greater than, or equal to, alfa. Gutmann's lambda-2 can easily be computed using statistical software such as SPSS. Lambda-2 of the shortened, eight item version of the INCOM was good in this Spanish sample ($\lambda_2 = .82$), with item-remainder correlations varying between $r = .44$ and $r = .61$. The correlation between the original INCOM total score, based on all eleven items, and the eight item short version was $r = .93$ ($p < .001$), which suggests that this shortened eight item version can still be considered a parallel test of the full scale INCOM-E.

Relations between the shortened INCOM-E and age and gender were also examined in the present sample.

Table 1. Item Characteristics and Correlations in Study 1 and Study 2 (Total Samples)

Item	Spain				Chile				Correlations										
	M	SD	Skewness	Kurtosis	M	SD	Skewness	Kurtosis	1	2	3	4	5	6	7	8	9	10	11
1	2.79	1.26	.03	-1.14	2.51	1.30	.25	-1.17	1.00	.19	.37	.39	.31	-.12	.39	.28	.26	-.15	.50
2	3.39	1.15	-.48	-0.49	3.08	1.39	-.20	-1.23	.31	1.00	.23	.18	.52	-.00	.29	.45	.30	-.01	.26
3	2.77	1.19	.03	-1.01	2.53	1.33	.31	-1.13	.48	.37	1.00	.40	.36	-.19	.51	.23	.16	-.08	.49
4	2.83	1.22	-.09	-1.09	2.65	1.35	.20	-1.19	.39	.28	.45	1.00	.28	-.11	.35	.25	.19	-.11	.40
5	3.28	1.21	-.30	-0.81	2.93	1.41	-.01	-1.29	.33	.32	.33	.41	1.00	-.06	.36	.58	.36	-.01	.36
6	3.17	1.34	-.08	-1.19	2.77	1.47	.19	-1.32	.09	.00	.10	.05	-.03	1.00	-.17	-.02	.06	.40	-.20
7	2.79	1.22	-.04	-1.10	2.81	1.33	.04	-1.20	.36	.32	.43	.37	.38	.02	1.00	.28	.18	-.10	.53
8	3.59	1.13	-.68	-0.21	2.98	1.36	-.12	-1.21	.33	.36	.33	.36	.42	.01	.46	1.00	.58	.01	.27
9	3.84	1.11	-.92	0.23	3.70	1.25	-.82	-0.32	.19	.28	.15	.17	.28	-.08	.24	.38	1.00	.07	.20
10	2.99	1.30	.10	-1.11	2.19	1.31	.72	-0.71	.40	.23	.41	.39	.30	.09	.40	.36	.10	1.00	-.13
11	2.80	1.25	.01	-1.08	2.52	1.38	.36	-1.16	.42	.27	.41	.41	.35	.07	.44	.39	.21	.56	1.00

Note. Above the diagonal the correlations for the Spanish sample, below the diagonal for the Chilean sample. Item scores all range from 1 to 5.

There was a significant negative effect of age in this sample ($r = -.22, p < .001$), indicating that older participants scored slightly lower than younger participants. Gender differences with regard to the mean score on the shortened INCOM-E were examined by means of an ANOVA. There was a small, but significant gender difference: $M_{\text{men}} = 23.51$ ($SD_{\text{men}} = 6.50$), $M_{\text{women}} = 24.82$ ($SD_{\text{women}} = 6.21$), $F(1, 1132) = 11.93, p < .001, \eta^2 = .01$. This means that, like what Gibbons and Buunk (1999) found, women scored slightly higher than men in this sample. Overall, it can be concluded that there are just small relations of the shortened INCOM-E with gender and age.

STUDY 2: Chile

Method

Participants

The data came from four studies conducted in the city of Antofagasta in Chile that examined the role of socio-cognitive processes in the evaluation of the quality of life in different age groups, i.e., children (10–14 years old; $n = 600$, 50.0% boys), adolescents (15–18 years old; $n = 530$, 45.1% boys), adults (19–60 years old; $n = 1,229$, 49.7% men), and senior citizens (over 60, with a maximum of 88 years old; $n = 398$, 48.5% men). The total samples included 2,757 participants (48.7% males). The samples of children and adolescents were similar in sex, age and type of educational establishment (public versus private). Parental consent was obtained for all children under the age of 18. The adults were civil servants and employees of public and social organizations, including universities, private companies, governmental institutions, waiting rooms of state services, and banks. Senior citizens were recruited from various government and non-governmental services including clubs for seniors, neighborhood councils, as well as public and private health centers. As in Study 1, the overall sample was large and quite heterogeneous with regard to a range of background variables.

Procedure

As in Study 1, data collection was carried out by self-report questionnaires that were completed voluntarily by the participants after providing their informed consent, and after receiving instructions on how the questionnaire had to be filled out. Additional explanations were provided, if necessary. The researchers emphasized that participation in the study was voluntary, and that the confidentiality of the information provided was guaranteed. Some findings from the adult sample, not related to the psychometric properties of the INCOM-E, have been published before in a Spanish language journal (Urzúa

Table 2. Factor Loadings Based on the CFA in Study 1 and Study 2 (Total Samples)

		Spain		Chile	
		Factor 1	Factor 2	Factor 1	Factor 2
1	I often compare myself with others with respect to what I have accomplished in life	.59		.64	
2	If I want to learn more about something, I try to find out what others think about it	.50		.50	
3	I always pay a lot of attention to how I do things compared with how others do things	.61		.68	
4	I often compare how my loved ones (boy or girlfriend, family members, etc.) are doing with how others are doing	.51		.64	
5	I always like to know what others in a similar situation would do	.64		.58	
6	I am not the type of person who compares often with others (R)	.18		.07	
7	If I want to find out how well I have done something, I compare what I have done with how others have done		.63		.66
8	I often try to find out what others think who face similar problems as I face		.57		.63
9	I often like to talk with others about mutual opinions and experiences		.42		.35
10	I never consider my situation in life relative to that of other people (R) /		.13		.65
11	I evaluate my life by comparing it with other people ¹				
	I often compare how I am doing socially (e.g., social skills, popularity) with other people		.66		.70
		$\chi^2(43) = 1,318.02, p < .001$ RMSEA = .162, 90% CI [.15, .17] CFI = .82 SRMR = .098		$\chi^2(43) = 956.09, p < .001$ RMSEA = .088, 90% CI [.083, .093] CFI = .95 SRMR = .050	

Note. ¹ This item (Item 10) was different in the two samples. The first item is the original item, and was used in the Spanish sample. The second item was used in the Chilean sample.

et al., 2012). The research was approved by the local ethics committee.

Instrument

As in Study 1, the INCOM-E was used, with one noticeable change. Based on a pilot study, one of the reversely coded items of the INCOM-E (Item 10 in Table 1) was replaced by a positively coded item. More specifically, the item "I never consider my situation in life relative to that of other people" was replaced by "I evaluate my life by comparing it with other people".

Data analysis

As in Study 1, we examined the item characteristics, next executed a Confirmatory Factor Analysis (CFA),

and finally conducted a Mokken Scale Analysis (MSA; Mokken, 1971). Given the large range in age, we also examined if the results differed between four subsamples that differed in age.

Results

First, we again calculated item characteristics (means, standard deviations, skewness and kurtosis; see Table 1). There were no clear deviations from normality in the Chilean sample (combined sample). CFA's were then conducted for the four age groups separately, as well as for males and females separately. The model that was tested was the a priori hierarchical model with two factors (see Study 1, comparisons of abilities and comparisons of opinions; cf. Gibbons & Buunk, 1999) and one higher factor. It was found that the fit indices were quite

Table 3. Model Fit for Four Age Groups, Males and Females in Study 2

	$\chi^2(43)$	RMSEA 90% CI	CFI	SRMR
Age 10–14	254.27	.091 [.080,.10]	.90	.065
Age 15–18	217.27	.088 [.076,.099]	.92	.059
Age 19–60	559.96	.099 [.092,.11]	.95	.055
Age 61–88	269.83	.12 [.10,.13]	.94	.058
Males	462.44	.085 [.078,.092]	.96	.048
Females	543.07	.091 [.084,.098]	.94	.054

comparable between all four age groups (see Table 3), although criteria for the examination of invariance (e.g., Cheung & Rensvold, 2002) indicated that these structures might not be invariant between samples (CFI's for example differ $> .01$)¹. The RMSEA was mostly reasonable (except for in the senior sample), the CFI's were all .90 or higher (but not all $> .95$), and the SRMR was well below .10 in all samples. This indicates a reasonable, but not very good fit of the tested model.

For comparability reasons, we also conducted the CFA for the total combined sample (as in Study 1). In this sample, the fit of the a priori model with two factors and one higher order factor was reasonable: RMSEA = .088, CFI = .95, SRMR = .050; see Table 2). Closer inspection of the factor loadings (for the total sample; see Table 2) revealed that Item 6 (a reversely coded item) again functioned poorly (low factor loading). The reformulated Item 10 on the other hand functioned quite well (in terms of its factor loading) in this Chilean sample. Item 9 had a relatively low factor loading compared to the other items, as was the case in Study 1. A model with just one factor functioned about equally well in the total sample as the model with two factors and one higher order factor: $\chi^2(44) = 1016.09$, $p < .001$, RMSEA = .090, 90% CI [.085, .094], CFI = .95, SRMR = .051. These results seem to support the distribution of the items over the two factors that have been reported previously (Gibbons & Buunk, 1999), although a unidimensional structure functioned about equally well. Correlation between the two factors was $r = .34$.

¹Invariance of the factor structure across age groups was also examined by extracting two factors in all four subsamples individually. Similarities between the factors were examined by calculating Tucker's phi coefficient (Tucker, 1951) between corresponding factors after Procrustes rotation (Kiers & Groenen, 1996). The first factor was equal in all four samples, with phi's ranging between .98 and 1.00. The second factor in the senior citizens sample differed somewhat from the second factor in both the sample of children ($\phi = .87$) and the sample of adolescents ($\phi = .91$). Other phi coefficients for the second factor ranged from .96 to .98, which is indicative of a good correspondence. The main difference between structures was that in the senior citizen sample, the second factor was characterized by high loadings from items 6 and 9 only.

As in Study 1, we also conducted MSA's in the present samples. The results were highly comparable between the four subsamples, and the MSA was therefore also conducted on the total sample. The MSA in the total Chilean sample resulted in a moderate Mokken scale ($H = .40$; see Sijtsma & Molenaar, 2002), consisting of nine items: All items except Items 6 and 9. These results are largely consistent with Study 1 and support the removal of these two items from the INCOM-E. Although Item 10 did function well in the present study, this item was not part of the original INCOM-E, and was also not present in Study 1. This item was therefore also dropped. In order to examine the invariance of the factor loadings in the four age groups, we calculated the factor loadings for a one factor solution in all four age groups separately, and compared these by means of Tucker's (1951) phi-coefficients. These were close to perfect (range .99 to 1.00). This means that the one-factor solution was highly comparable between all four subsamples. Reliability of the remaining eight items in the present study was $\lambda^2 = .83$, with item-remainder correlations ranging from $r = .46$ to $r = .59$. The correlation between the full scale and the reduced eight item version was $r = .97$ in this sample ($p < .001$), which indicates only a very marginal loss of information as a result of dropping these three items. We also compared the factor loadings of the eight item INCOM-E in the total Chilean sample with that found in the Spanish sample. Tucker's phi equaled 1.00, indicating that these factors can be considered identical.

As in Study 1, the relation between the shortened INCOM-E and age was examined. Consistent with Study 1, and in line with previously published findings for the adult sample (Urzúa et al., 2012), there was a small significant negative correlation with age in this sample ($r = -.13$, $p < .001$). Gender effects were again examined by means of an ANOVA. There was a small, but significant gender difference, but this time, men scored slightly higher than women: $M_{men} = 22.55$ ($SD_{men} = 7.40$), $M_{women} = 21.48$ ($SD_{women} = 7.20$), $F(1, 2,749) = 14.71$, $p < .001$, $\eta^2 = .01$. All in all, it can be concluded that the relations between age and gender and the shortened Spanish language version of the INCOM are overall quite small.

Discussion

The aim of the present paper was to examine in more detail the psychometric properties of the INCOM-E (the Spanish version of the INCOM; Buunk et al., 2005), in two large samples from Spain and Chile. In the Spanish sample, confirmatory analyses did not support the proposed two-factor structure of the INCOM-E that was found in a previous Spanish sample (Buunk et al., 2005), and in previous samples in The Netherlands (Gibbons &

Buunk, 1999), Germany (Schneider & Schupp, 2014) and the United States (Gibbons & Buunk, 1999). In the Chilean sample, the model fit was reasonable, and the structure was also found to be invariant across age groups and gender. Mokken Scale Analyses revealed a single scale in both samples, in which the reversely coded items were omitted. In addition, one of the remaining items was not functioning well in both samples. Based on these results, it was decided to remove three of the original eleven items of the INCOM-E, resulting in an eight-item scale with one underlying factor. This shortened unidimensional scale was found to be reliable in both samples, and correlated very strongly with the full-length INCOM-E. Relations between the shortened INCOM-E and age and gender were examined. Although some significant effects were found with regard to gender and age, these relations were rather small. Based on the present studies results, it is advised to use the shortened eight-item version of the INCOM-E in Spanish speaking countries.

Future studies may investigate why, in contrast to the Dutch and English versions of the INCOM, the underlying two factor structure, distinguishing between 'ability comparisons' and 'opinion comparisons', was not found in the Spanish version of the INCOM, at least not in the large composed samples that were examined in the present paper. It is possible that in Spanish-speaking cultures social comparisons with regard to abilities or opinions serve a somewhat different purpose than in Anglo-Saxon cultures. With regard to opinions, for instance, it has been found that individuals in relatively collectivistic cultures, such as Spain and Chile, are more likely to avoid disagreement with others than individuals in relatively individualistic cultures, such as The Netherlands and the US. As a consequence individuals from Spain and Chile may be more likely to seek opinion-related social comparisons that confirm consensus and agreement with others, in contrast to individuals from relatively individualistic cultures who may show less hesitation to seek opinion-related social comparisons that show disagreement and distinctiveness. Differences in social comparison tendencies such as these, may cause opinion- and ability related social comparisons to be less different in relatively collectivistic cultures as opposed to relatively individualistic cultures. The fact that a previous study with the INCOM-E (Buunk et al., 2005) did come up with an underlying two-factor structure distinguishing between opinion- and ability-related social comparisons may be explained by differences in samples. Compared to the samples collected by Buunk et al. (2005) the samples in the present study were larger and more heterogeneous in terms of country of residence (Chile and Spain) and age, and also included children and elderly people.

Because of the relatively large and heterogeneous samples that were used in the present research, the present study generates interesting and new insights into the structure of the INCOM-E. Our research suggests that the most reliable way to assess social comparison orientation in Spanish speaking countries is by means of a scale consisting of a selection of eight items from the INCOM-E (all items presented in Table 1, except the items 6,9 and 10). One could argue that a limitation of the present research is that the two studies used slightly different versions of the INCOM, as one badly functioning item was already replaced in the Chilean sample, which might have limited the direct comparison between the two datasets. However, precisely this item was also functioning poorly in the Spanish sample, which in fact confirmed its deletion in the Chilean sample. In addition, although the shortened version of the INCOM-E was reliable and highly comparable to the full-scale INCOM-E, we did not examine relations with other variables that might further support its validity. Future studies should primarily focus on the relations with other variables in order to further establish the construct validity of the shortened INCOM-E.

Financial support

This research was in part funded by Grant FONDECYT 11100090 from the National Commission for Scientific and Technological Research of the Government of Chile (CONICYT).

References

- Beach, S. R., & Tesser, A. (2000). Self-evaluation maintenance and evolution: Some speculative notes. In L. Wheeler & J. Suls (Eds.), *Handbook of social comparison: Theory and research* (pp. 123–140). Dordrecht, the Netherlands: Kluwer.
- Buckley, F. H. (2014). Schadenfreude and laughter. In W. W. van Dijk & J. W. Ouwerkerk (Eds.), *Schadenfreude: Understanding pleasure at the misfortune of others*. (pp. 219–226). New York, NY: Cambridge University Press. <https://doi-org.proxy-ub.rug.nl/10.1017/CBO9781139084246.018>
- Buunk, A. P., aan 't Goor, J., & Solano, A. C. (2010). Intrasexual competition at work: Sex differences in the jealousy-evoking effect of rival characteristics in work settings. *Journal of Social and Personal Relationships*, 27, 671–684. <http://doi.org/10.1177/0265407510368964>
- Buunk, A. P., & Dijkstra, P. (2017). Social comparisons and well-being. In M. Michael D. Robinson & M. Eid (Eds.), *The happy mind: Cognitive contributions to well-being* (pp. 311–330). New York, NY: Springer.
- Buunk, A. P., & Gibbons, F. X. (2006). Social comparison orientation: A new perspective on those who do and those who don't compare with others. In S. Guimond (Ed.), *Social comparison and social psychology: Understanding cognition*,

- intergroup relations and culture (pp. 15–33). Cambridge, UK: Cambridge University Press.
- Buunk, A. P., Belmonte, J., Peiró, J. M., Zurriaga, R., & Gibbons, F. X.** (2005). Diferencias individuales en la comparación social: Propiedades de la Escala Española de Orientation hacia la Comparación Social [Individual differences in social comparison: Characteristics of the Spanish scale for Social Comparison Orientation]. *Revista Latinoamericana de Psicología*, 37, 561–579.
- Buunk, A. P., Bennenbroek, F. T. C., Stiegelis, H. E., van den Bergh, A. C. M., Sanderman, R., & Hagedoorn, M.** (2011). Follow-up effects of social comparison information on the quality of life of cancer patients: The moderating role of social comparison orientation. *Psychology & Health*, 27, 641–654. <http://doi.org/10.1080/08870446.2011.613994>
- Buunk, A. P., Gibbons, F. X., Dijkstra, P., & Krizan, Z.** (in press). Individual differences in social comparison: The complex effects of Social Comparison Orientation. In J. Suls, R. L. Collins, & L. Wheeler, (Eds.). *Social comparison in judgment and behavior*. Oxford, UK: Oxford University Press.
- Buunk, A. P., Peiró, J. M., & Griffioen, C.** (2007). A positive role model may stimulate career-oriented behavior. *The Journal of Social Psychology*, 37, 1489–1500. <https://doi.org/10.1111/j.1559-1816.2007.00223.x>
- Buunk, A. P., Zurriaga, R., & Peiró, J. M.** (2010). Social comparison as a predictor of changes in burnout among nurses. *Anxiety, Stress & Coping*, 23, 181–194. <http://doi.org/10.1080/10615800902971521>
- Buunk, A. P., Zurriaga, R., Peiró, J. M., Nauta, A., & Gosálvez, I.** (2005). Social comparisons at work as related to a cooperative social climate and to individual differences in social comparison orientation. *Applied Psychology: An International Review*, 54, 61–80. <https://doi.org/10.1111/j.1464-0597.2005.00196.x>
- Cheung, G. W., & Rensvold, R. B.** (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. https://doi.org/10.1207/S15328007SEM0902_5
- Festinger, L.** (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140. <https://doi.org/10.1177/001872675400700202>
- Diener, E., & Fujita, F.** (1997). Social comparisons and subjective well-being. In A. P. Buunk & F. X. Gibbons (Eds.), *Health, coping, and well-being: Perspectives from social comparison theory*. (pp. 329–358). Mahwah, NJ: Erlbaum.
- Gibbons, F. X., & Buunk, A. P.** (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, 76, 129–142. <http://doi.org/10.1037/0022-3514.76.1.129>
- Gibbons, F. X., Lane, D. J., Gerrard, M., Pomery, E. A., & Lautrup, C. L.** (2002). Drinking and driving: A prospective assessment of the relation between risk cognitions and risk behavior. *Risk Decision and Policy*, 7, 267–283. <https://doi.org/10.1017/S1357530902000601>
- Gilbert, P., Price, J., & Allan, S.** (1995). Social comparison, social attractiveness and evolution: How might they be related? *New Ideas in Psychology*, 13, 149–165. [https://doi.org/10.1016/0732-118X\(95\)00002-X](https://doi.org/10.1016/0732-118X(95)00002-X)
- Hardouin, J. B., Bonnaud-Antignac, A., & Sébille, V.** (2011). Nonparametric item response theory using Stata. *The Stata Journal*, 11, 30–51. <https://doi.org/10.1177/1536867X1101100102>
- Helgeson, V. S., & Taylor, S. E.** (1993). Social comparisons and adjustment among cardiac patients. *Journal of Applied Social Psychology*, 23, 1171–1195. <https://doi.org/10.1111/j.1559-1816.1993.tb01027.x>
- Hu, L., & Bentler, P. M.** (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <https://doi-org.proxy-ub.rug.nl/10.1080/10705519909540118>
- Junker, B. W., & Sijtsma, K.** (2001). Nonparametric item response theory in action: An overview of the special issue. *Applied Psychological Measurement*, 25, 211–220. <https://doi.org/10.1177/01466210122032028>
- Kiers, H. A. L., & Groenen, P.** (1996). A monotonically convergent algorithm for orthogonal congruence rotation. *Psychometrika*, 61, 375–389. <https://doi.org/10.1007/BF02294345>
- Kline, R. B.** (2005). *Principles and practice of structural equation modeling* (2nd Ed.). New York, NY: Guilford.
- Mokken, R. J.** (Ed.). (1971). *A theory and procedure of scale analysis*. The Hague, the Netherlands: Mouton.
- Neff, K. D., & Vonk, R.** (2009). Self-compassion versus global self-esteem: Two different ways of relating to oneself. *Journal of Personality*, 77, 23–50. <https://doi-org.proxy-ub.rug.nl/10.1111/j.1467-6494.2008.00537.x>
- Nesi, J., & Prinstein, M. J.** (2015). Using social media for social comparison and feedback-seeking: Gender and popularity moderate associations with depressive symptoms. *Journal of Abnormal Child Psychology*, 43, 1427–1438. <http://doi.org/10.1007/s10802-015-0020-0>
- Schneider, S., & Schupp, J.** (2014). Individual differences in social comparison and its consequences for life satisfaction: Introducing a short scale of the Iowa-Netherlands Comparison Orientation Measure. *Social Indicators Research*, 115, 767–789. <http://doi.org/10.1007/s11205-012-0227-1>
- Smits, I. A. M., Timmerman, M. E., & Meijer, R. R.** (2012). Exploratory Mokken Scale analysis as a dimensionality assessment tool: Why scalability does not imply unidimensionality. *Applied Psychological Measurement*, 36, 516–539. <https://doi.org/10.1177/0146621612451050>
- Stock, M. L., Gibbons, F. X., Beekman, J. B., & Gerrard, M.** (2015). It only takes once: The absent-exempt heuristic and reactions to comparison-based sexual risk information. *Journal of Personality and Social Psychology*, 109, 35–52. <https://doi-org.proxy-ub.rug.nl/10.1037/a0039277>
- Stout, W.** (2001). Nonparametric item response theory: A maturing and applicable measurement modeling approach. *Applied Psychological Measurement*, 25, 300–306. <https://doi.org/10.1177/01466210122032109>
- Suls, J., & Wheeler, L.** (Eds.) (2000). *Handbook of social comparison: Theory and research*. New York, NY: Kluwer Academic Publishers.
- Sijtsma, K.** (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74, 107–120. <http://doi.org/10.1007/s11336-008-9101-0>
- Sijtsma, K., Meijer, R. R., & van der Ark, L. A.** (2011). Mokken scale analysis as time goes by: An update for scaling practitioners. *Personality and Individual Differences*, 50, 31–37. <https://doi.org/10.1016/j.paid.2010.08.016>

- Sijtsma, K., & Molenaar, I. W.** (2002). *Introduction to nonparametric item response theory*. Thousand Oaks, CA: Sage.
- Stock, M. L., Gibbons, F. X., Beekman, J. B., & Gerrard, M.** (2015). It only takes once: The absent-exempt heuristic and reactions to comparison-based sexual risk information. *Journal of Personality and Social Psychology*, 109, 35–52 <http://doi.org/10.1037/a0039277>
- Tucker, L. R.** (1951). *A method for synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington, DC: Department of the Army.
- Urzúa M., A., Zúñiga B., P., & Buunk, A. P.** (2012). La edad y el sexo en la orientación a la comparación social [Effects of age and gender on Social Comparison Orientation]. *Terapia Psicológica*, 30, 79–88. <http://doi.org/10.4067/S0718-48082012000300008>
- Vogel, E. A., Rose, J. P., Okdie, B. M., Eckles, K., & Franz, B.** (2015). Who compares and despairs? The effect of social comparison orientation on social media use and its outcomes. *Personality and Individual Differences*, 86, 249–256. <https://doi.org/10.1016/j.paid.2015.06.026>
- Wilson, A. E., & Ross, M.** (2000). The frequency of temporal-self and social comparisons in people's personal appraisals. *Journal of Personality and Social Psychology*, 78, 928–942. <http://doi.org/10.1037/0022-3514.78.5.928>