Instituto de Investigaciones en Ciencias del Comportamiento "IICC"

Prosociality Assessment Across Cultures

Gian Vittorio Caprara, Carlo Tramontano, Laura Di Giunta
University of Rome "La Sapienza"

Patrizia Steca University of Milan
Nancy Eisenberg
Arizona State University

Erick Roth*
Universidad Católica Boliviana "San Pablo"- La Paz

Instituto de Investigaciones en Ciencias del Comportamiento

"IICC"

Prosociality Assessment Across Cultures

Gian Vittorio Caprara, Carlo Tramontano, Laura Di Giunta
University of Rome "La Sapienza"

Patrizia Steca University of Milan
Nancy Eisenberg
Arizona State University

Erick Roth*
Universidad Católica Boliviana "San Pablo"- La Paz

(*) Dr. Erick Roth U.
eroth@ucb.edu.bo

2009

Prosociality Assessment Across Cultures

Gian Vittorio Caprara, Carlo Tramontano, Laura Di Giunta
University of Rome "La Sapienza"

Patrizia Steca
University of Milan

Nancy Eisenberg
Arizona State University

Erick Roth
Universidad Católica Boliviana "San Pablo"- La Paz

ABSTRACT

In the present study, the measurement properties of a revised scale of adults' prosociality were tested by using Classical Test Theory and Item Response Theory approaches. Participants were young adults from three countries: Italy, U.S., and Bolivia. The goal was to examine whether the scale items were equally effective in discriminating people and equally informative, and if there were cultural differences in the items' discrimination and difficulty. Results strongly supported the measurement effectiveness and sensitivity of the scale in the three countries and showed some cultural differences in the items' functioning.

Introduction

Prosociality refers to individuals' tendencies to undertake voluntary actions aimed at benefiting others, such as sharing, donating, caring, comforting, and helping (Batson, 1998; Eisenberg, Fabes, & Spinrad, 2006; Penner, Dovidio, Piliavin, & Schroeder, 2005). Findings from developmental research have shown that prosocial responding becomes relatively stable during childhood and early adolescence and that it often arises from complex processes involving evaluative mechanisms, moral reasoning, perspective taking, and self-regulatory capacities (Caprara & Pastorelli, 1993; Eisenberg et al., 2006; Eisenberg et al., 1999; Krebs & Van Hesteren, 1994). In addition, prosocial tendencies are correlated with psychosocial adjustment in children and adolescents (see Eisenberg et al., 2006, for a review). Hence, prosociality may represent a protective factor that fosters self-enhancement, self-acceptance and successful psychosocial adaptation, as it is associated with one's own integration in the community, positive mood, health, and life satisfaction (Caprara & Steca, 2005; Keyes, 1998; Piliavin, 2003; Van Willigen, 2000).

Prosociality Measurement

Despite its importance, prosociality is still an elusive topic in regard to conceptualization and assessment, especially in adulthood. There has been a lively debate concerning its psychological components and measurement, especially in adulthood (Batson, 1991; Eisenberg et al., 2006; Schroeder, Penner, Dovidio, & Piliavin, 1995), resulting in difficulty in reaching an agreement on its content and the most appropriate indicators (Eisenberg et al., 2002).

Various measures of prosocial tendencies were developed in the 1980s and thereafter, focusing on altruistic tendencies and behaviors (Midlarsky, Hannah, & Kahana, 1984; Rushton, Chrisjohn, & Fekken,1981) and different types of prosociality (Carlo & Randall, 2002). Caprara and Pastorelli (1993) constructed a measure to assess individual differences in prosociality among children that has been used in different studies (e.g., Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000; Bandura, Caprara, Barbaranelli, Gerbino, & Pastorelli, 2003; Caprara, Barbaranelli, Pastorelli, Cermak, & Rosza, 2001). More recently, Caprara and colleagues (Caprara, Steca, Zelli, & Capanna, 2005) derived from the original measure a 16-item scale designed to assess prosociality even among adolescents and adults. Whereas the focus of the initial scale was the frequency of prosocial behaviors such as helping, caring and sharing, the more recent scale also included items that reflect evidence regarding the importance of empathic feelings and sympathetic concern in altruism (Batson, 1991; Eisenberg et al., 2006). The scale for adults is, consequently, characterized by a broader sampling of prosocial tendencies in comparison to previous measures and it has been widely validated in the Italian context via the Classical Test Theory (henceforth CTT) and the Item Response Theory (henceforth IRT) approach (Caprara et al., 2005).

Cultural Differences in Prosociality

Prior research provides support for the relevance of the social environment and specific contexts in shaping prosocial responding—perhaps especially cultural norms and values stressing solidarity and cooperation (e.g., Eisenberg et al., 2006).

2

Cross-cultural comparisons have shown significant variations in prosocial actions (Caprara et al., 2001), particularly in sharing behavior (Carlo, Roesch, Knight, & Koller, 2001; Pilgrim & Rueda-Riedle, 2002).

In studying cultural differences, it is very common to refer to the dimensions of individualism and collectivism (Hofstede, 1980; Triandis, 1990). Generally, individualism pertains to an attitude of autonomy, a focus on the self, and the preference for competition and personal goals (Hofstede 1980; Triandis, 1990); collectivism refers to the prevalence of in-group goals and values and the preference for interdependence, solidarity and cooperation (Hui, 1988; Shkodriani & Gibbons, 1995). Western countries, including the U.S. and Italy, are generally considered individualistic cultures, whereas Latin American as well as Asian countries are generally considered collectivistic cultures, even though this kind of taxonomy risks to be overly simplistic and should be generalized with great attention, as suggested by Oyserman and colleagues (Oyserman, Coon, & Kemmelmeier, 2002). Further the prevalence of the individualistic values could appear antithetic to prosocial actions and behaviors, which could be expected to be more common in collectivistic societies (Bolivia, in our case), but this hypothesis is not completely true (Allik, & Realo, 2004; Kemmelmeier, Jambor, Letner, 2006), perhaps because collectivistic and individualistic values coexist in varying degree within the same culture (Triandis, 1990; Triandis & Gelfand, 1998). Moreover, in individualistic societies, prosociality can derive from an emphasis on a universalistic ethic that promotes social justice and human rights (Kemmelmeier, Jambor, Letner, 2006; Schwartz, 1990), and prosocial actions could be a consequence or a reflection of personal responsibilities and social commitment (Waterman, 1981).

Difficulties in investigating and comparing the relevance and meaning of prosocial actions in different cultures can result from the lack of instruments that have been validated across cultures. Consequently, research examining the psychometric properties of an instrument designed to measure prosociality in different cultures could foster more in-depth research on between- and within-country differences in prosociality.

Aims

The main purpose of this research was to examine the psychometric properties of Caprara and colleagues' (Caprara et al., 2005) measure of adolescents' and adults'prosociality using both the CCT and IRT in three different cultural contexts: two Western countries, United States and Italy, and a Latin American country, Bolivia. To our knowledge, no analogous empirical study, comparing three countries from three continents (i.e., Europe, North America and South America) via IRT, is available.

Method

Participants and Procedure

The Italian participants were 844 students, 307 men and 537 women, ranging in age from 20 to 24 years ($M = 21.89$ years; $SD = 1.26$). They were all college students, but varied widely in demographic and socioeconomic backgrounds. Participants were recruited in Rome and Milan on voluntary basis. They completed a set of questionnaires that included the measure of interest for this study during specially scheduled group sessions, coordinated by trained researchers.

The American participants were 1671 college students at a large south-western university, including 746 men and 915 women (6 individuals did not indicate their sex). Participants ranged in age from 18 to 22 years old (M = 18.85 years; SD = 1.00). The majority of the participants were Caucasian (70.9%), 10% were Hispanic, 5% were Asian, 2.9% were African American, 1.4% were Middle-Eastern, and the remaining participants declared none (1.4%) or two (5.3%) or three (.5%) of the previous ethnicities. A group of assistants administered a set of questionnaires that included the prosociality scale during class time. Participation was voluntary, although it was one of many activities that students could select to fulfill a requirement for their introductory psychology class.

The Bolivian participants were 937 students, 409 men and 510 women, ranging in age from 18 to 24 years old (M = 20.33 years; SD=1.98). Fifteen percent of the participants were enrolled in fourth and fifth years of high school whereas 85% were college students. Seventy-two percent of the youth were residents in the urban center of La Paz and the remaining 28% lived in the rural area of North Yungas. Participation was voluntary. An Italian researcher and at least one Bolivian researcher were in class during the scale's administration and offered the students a preliminary brief training in order to ensure their understanding of the comprehension of the instructions.

The Prosociality Scale

Students rated (1 = *never/almost never true*; 2 = *occasionally true*; 3 = *sometimes true*; 4 = *often true;* 5 = *almost always/always true)* the 16 items, including 4 items on helping, 4 on sharing, 4 on caring behaviour, and 4 on empathic concern. The scale was translated from Italian into English and Spanish by bilingual experts, paying a great attention to the adaptation of the item content to the socio-cultural context. The scale was then back translated to verify the correspondence to the original version.

CTT vs. IRT Application to the Assessment of Adults' Prosociality

The psychometrics literature of the past two decades has shown that the CTT approach to evaluating scales' properties has some limitation that the IRT analytical methods and procedures overcome (Hambleton, Swaminathan, & Rogers, 1991; Salzberger, Sinkovics, & Schlegelmilch, 1999). First, IRT analyses can determine a test's measurement precision at any value of the latent trait. Second, IRT parameters are invariant with respect to the sample characteristics from which they are generated. Third, IRT methods can quantify the information value of both individual items and the overall test, and this information can be evaluated at any level of the latent trait. Finally, IRT methods permit a direct comparison of different sets of items or entire scales referring to the same construct domain on the basis of their discriminative power.

In the present study, we were interested in answering four specific questions that we believe could not be adequately addressed via traditional assessment procedures: 1) Do different prosociality items behave similarly, that is, are they equally discriminating and similar in their difficulty? 2) Are the items equally informative across the different levels of the estimated latent trait of prosociality? 3) What is the information provided by the full scale along the trait continuum? and 4) Are there differences among youth from different cultures in the items' difficulty and discrimination parameters?

IRT characteristics

IRT refers to a family of different models designed to represent the relation between an individual's item response and an underlying latent trait, theta (θ), to which the items refer and which is assumed to influence the way in which a person responds to items (Embretson, & Reise, 2000; Reise, Ainsworth, & Haviland, 2005; Van der Linden & Hambleton, 1997). One core element of IRT is the Item Response Function (IRF), that "for item on a rating scale an IRF is a mathematical function describing the relation between where an individual falls on the continuum of a given construct and the probability that he or she will give a particular response to a scale item designed to measure that construct" (Reise et al., 2005, p. 95). In the two-parameter logistic item response model (2PLM), usually used with personality scales, the probability of endorsing an item in a trait-consistent manner is a function of a person parameter (i.e., the level of θ) and two item parameters, the difficulty and the discrimination parameters, respectively. The item difficulty or location parameter (β_i) is defined as the point on the θ scale at which the probability that the item is endorsed in a trait-consistent manner is greater than .5 (Hambleton, Swaminathan, & Rogers, 1991; Reise, Widman, & Pugh, 1993). This parameter is expressed in the same metric of θ; typically it varies between -2 and 2, where higher values correspond to more difficult items (Hambleton et al., 1991; Reise & Henson, 2003). The item discrimination parameter (α) represents the item's ability to differentiate between people with contiguous θ levels. It is proportional to the slope of IRF and theoretically ranges from - ∞ to + ∞ (Hambleton et al., 1991). The higher the slope parameter, the higher the capability of the item to discriminate between respondents with different θ levels and the

more the IRF looks like a step function (Hambleton et al., 1991; Reise & Henson, 2003).

The relation of the person's θ level with the item difficulty and discrimination parameters is expressed by the item information. Hence, it is possible to plot an item information curve (IIC) to represent the information as a function of θ level (Hambleton et al., 1991). Item information curve is additive across items; hence, it is possible to produce an information curve for the full scale, called the test information curve (TIC), that expresses the relative precision of the scale across the different levels of the trait continuum. The inverse square root of the information value at a particular θ level is equal to the standard error of measurement (Reise & Henson, 2003).

IRT methods allow to identify items that are biased for or against a particular group of subjects through the analysis of the "differential item functioning" (DIF), that is, the difference in performance on a test item across groups that is not explainable by the difference in the target trait among groups (Bingenheimer, Raudebush, Levental, & Brooks-Gunn, 2005). However, in DIF analyses it is necessary to combine different information in order to demonstrate the equivalence of an instrument: parameters and DIF indices, group specific ICCs and IIF, and item content (Bingenheimer et al., 2005).

We explored DIF in the prosociality scale, comparing the three groups of youth from Italy, the U.S., and Bolivia two at a time, looking for invariant items.

Results

Classical Test Theory: Item and Scale Statistics

Before performing IRT analyses, CTT item and scale statistics were computed on the data from the three countries. In particular, prosociality items' means, standard deviations, skewness, and kurtosis were calculated. Exploratory factor analyses for each sample also were conducted to test the monodimensionality of the scale, which is a fundamental assumption for IRT models. In evaluating the dimensionality we considered a multi-criteria approach: the analysis of the eigenvalues; the replicability of the factorial structure across the samples, through the Tucker coefficient; and the analysis of the reproduced correlation matrix. Then Cronbach alphas and mean-corrected item-total correlations were computed for the entire set of items within each sample.

Looking across countries, the mean responses ranged from 3.00 to 4.16, with an overall mean of 3.67 and standard deviation of .92. The average skewness was -.36 and the average kurtosis was .18, indicating a good approximation of the items' distribution to the normal curve.

Although the eigenvalues of the second and sometimes third factors were above 1.00, they were quite low compared to the first factor and accounted for much less variance than the first factor in all three groups (see Table 1). In addition, the items on the second and third factors varied considerably across countries and did not make conceptual sense.

Moreover, the one-factor solution could be easy replicated in each sample and the Tucker coefficient indicated an optimal similarity among the factorial structure: the coefficient was equal to 1 in the comparisons U.S. vs. Italy and Italy vs. Bolivia, and to .99

in the comparison U.S. vs. Bolivia. For each sample, in the reproduced correlation matrix, only few residuals (less than 15%) were greater than .10, attesting to the appropriateness of the suggested monofactorial solution (McDonald, 1985). As shown in Table 2 (See Annex), across countries, the loadings of all items on the first factor ranged from .47 to .80 and the total variance explained by the first factor was uniformly high: 41.93%, 40.33%, and 41.78%, for Italy, U.S., and Bolivia, respectively. The corrected item-total correlations varied from .45 to .70, with an overall mean of .59, whereas at the scale level, the α for the entire set of items were

.91, .90, and .90, for Italy, U.S. and Bolivia, respectively, indicating high reliability of the scale.

Table 1. Factor analyses results: Initial eigenvalues (first column) and percentage of explained variance (second column), separately for the three countries.

	Italy		U.S.		Bolivia	
1	7.27	45.43	6.45	40.33	6.68	41.78
2	1.32	8.27	1.16	7.28	1.22	7.63
3	1.14	7.14	1.12	7.03	.974	6.09
4	.79	4.95	.89	5.58	.86	5.36
5	.73	4.55	.79	4.93	.81	5.06

Item Response Theory: Item and Scale Statistics

Item parameters. Next, we performed IRT analyses separately for each country sample, using the PARSCALE program (Muraki, & Bock, 1993). We implemented the Generalized Partial Credit Model (G-PCM; Muraki, 1996) that is based on an item response model potentially useful when the item responses can be conceptualized as ordered categories. Table 3 (See Annex) presents the G-PCM estimated item slope and location parameters for the three countries.

With respect to item discrimination parameters, in all three countries, items 3 and 13 showed the highest values. Location parameter for all the item had negative values, indicating that they are relatively easy to endorse; nevertheless, in all three countries, the easiest items were 3 (helping), 2 (sharing), and 10 (caring).

On the basis of the information derived jointly from the CTT and IRT, we decided to drop some items. In particular, we dropped items 2, 4 and 11 because they had very low values for both location and slope in each country. Items 2 and 11 also had low values for IRT parameters, as well as relatively low factor loadings and item-scale correlations.

After dropping the aforementioned three items, we replicated the IRT analyses on the reduced set of 13 items (see Table 3). In the Italian sample, all 13 items had slope values higher than 1, indicating a good ability to detect slight differences in young adults' prosociality. In particular, items 3, 6, 7, 10, 12 and 13 were the most informative and had the highest slope parameters. Thus, the items referring to helping and caring behaviors were the most effective in detecting slight variations in the prosociality trait. The discrimination parameters were also quite acceptable for items 1, 5, 8, 9, 14, 15 and 16, all of which had a slope parameter greater than 1. As to the items' location parameters, all of the items had a negative value, and the "easiest" items pertained to caring behavior and empathic feelings.

In the American sample, 10 of the 13 items were characterized by discrimination parameters higher than 1. In particular, items 3, 10 and 13, referring to helping and caring, were the most informative and had the highest slope parameters. The discrimination parameters were also good for items 5, 6, 8,

9, 12, 14 and 15, all of which had a slope parameter greater than 1. Three of the items with good discrimination (5, 8, and 12) referred to young adults' empathic concern toward others, three items (9, 14, 15) were concerned with sharing behaviors, and the last item (6) referred to helping behavior.

With regard to items' location parameters, as for the Italian sample, all of the items had a negative value; the "easiest" items pertained to sharing and helping behaviors.

In the Bolivian sample, all items except for item 12 were quite good in detecting differences in young adults' prosociality (above 1.0). Item 13, referring to caring behavior, was the most informative and had the highest slope parameter, equal to 1.59. Again, all of the items had a negative location parameter (difficulty value) and the "easiest" items pertained to caring.

Item and Scale Information Functions

The estimated item parameters were then used to produce the information functions for each item and for the overall scale. The items were, overall, quite informative, even though they showed some differences in the degree to which they could detect slight variations in people's prosociality and discriminate along the entire trait continuum, within and between countries. Some of the items were particularly good in discriminating slight differences in prosociality among people with either low or moderate levels of the trait; they were items 3, 5, 6, 7, 8, 9, 10, 12, and 13 in the Italian sample; items 3, 5, 10, 12, 13, and 15 in the American sample; and items 3, 9, 10, and 13 in the Bolivian sample. Overall, these results indicate that items measuring people's helping and caring behaviors (3, 10, and 13) had the highest measurement precision, especially at relatively moderate levels of prosociality, in each of the three

countries. Items referring to individuals' empathy (5 and 12) showed a high level of precision in both U.S. and Italy. Likewise, the level of information provided by items 5, 7, 8, 9, 10, 12 for the Italian sample; 5, 10, 12 for the American sample; and 6, 7, 9, 10, 14, 15 for the Bolivian sample was also quite good and, again, especially high at relatively low and moderate levels of prosociality. The remaining items (1, 14, 15, 16 for the Italian sample; 1, 6, 7, 8, 9, 14, 16 for the American sample; and 5, 8, 12, and 16 for the Bolivian sample) provided relatively little information.

Figures 1 (See Annex) present the IIC of sample items that were highly and lowly informative in the three countries. The information and standard error curves for the entire prosociality scale are presented for the Italian, American and Bolivian samples in Figure 2 (See Annex). For the three countries, the greatest information is concentrated at the low and middle levels of the prosociality trait continuum; accordingly, the standard error is greatest at relatively high trait levels.

Differential Item Functioning

The last IRT analysis was performed to examine whether the items' functioning varied across countries, that is, whether a differential item functioning (DIF) existed across country groups. In particular, we analyzed the non-uniform DIF, that is, DIF concerning the items' location and slope parameters of the three different pairs of comparisons (Italy vs. U.S.; Italy vs. Bolivia; and U.S. vs. Bolivia). Due to the large number of tests implemented in the method, a *p* level of .001 was used as the criterion to evaluate the standardized DIF statistic (SDIF) and the –2 log likelihood statistic was used to evaluate estimation improvement from a no-group difference model (i.e., one-group model) to a model that hypothesized differences in item parameters between the

countries (i.e., the two-group model). This statistic is defined as the difference between the –2 log likelihood values (G2) of the two models (one-group model and two-group model). Under certain conditions, this difference is distributed as a chi-square with degrees of freedom equal to the difference in degrees of freedom between the more and the less constrained models. Large and significant chi-square values indicate that imposing constraints across two countries on item locations and slopes worsened the model significantly. The test was performed with 26 degrees of freedom, corresponding to the 26 parameters that were freed in the two-group model.

Results of the three comparisons are reported in Table 4 (See Annex). In regard to the comparison between Italy and U.S., the slope parameter for item 7, referring to helping behavior, was significantly higher in Italy. The location parameters of items 8, 10, 12, 13 and 15, referring to empathic feelings and caring behaviors, were significantly lower in Italy, as well as item 6, referring to helping. In contrast, the locations for items 1, 3, 7 and 14, mainly referring to helping behaviors, were lower in the U.S. Items 5, 9 and 16 did not show statistically significant differences in *b* parameters. A worsening in G2 statistics was found when the item parameters were freely estimated in Italian and American samples (the two-group model), indicating the one-group model was better than the two-group model: G2 = 73607.737 for the two-group model and G2 = 71465.12 for the one-group model.

In regard to the comparison of Italy vs. Bolivia, the slope parameters for items 12 and 13, referring to empathic feelings and caring behaviors, were higher in Italy. The location parameters for items 8, 12 and 16, referring to empathic feelings, were significantly lower in Italy. In contrast,

parameters for items 6, 7, 13 and 14, mainly referring to sharing and helping behaviors, were lower in Bolivia. The remaining items did not show statistically significant differences in b parameters. As was found for the Italy vs. U.S. comparison, there was a worsening in G2 statistics when the item parameters were freely estimated for Italian and Bolivian samples, indicating the one-group model was better than the two-group model: G2 = 52275.555 for the two-group model and G2 = 51117.944 for the one-group model.

Finally, for the comparison U.S. vs. Bolivia, the slope parameter for item 7, referring to helping behavior, was significantly higher in Bolivia, whereas the slope parameter for item 12, referring to empathic feelings, was significantly higher in the U.S. The location parameters for items 1, 3, 5, 7, 12 and 16, referring to empathic feelings and helping behavior, were significantly lower in U.S. than Bolivia. On the other hand, location parameter for items 6, 8, 10, 13, 14 and 15, mainly referring to sharing behaviors, were significantly lower in Bolivia. The location b parameter was not different in the two samples for item 9. In contrast to the other comparisons, an improvement in G2 statistics was found when the item parameters were freely estimated in the two countries: G2 = 77025.916 for the two-group model and G2 = 76068.786 for the one-group model.

Discussion

In this article, we examined the generalizability of a new scale for assessing dispositional prosociality in three different cultural contexts. Previous work in Italy demonstrated the adequate psychometric properties of the scale, showing its reliability and its effectiveness in discriminating individual differences in prosociality,

especially at intermediate and low levels of prosociality. Furthermore, items referring to helping and caring behaviors were the most discriminative and relatively easy to endorse, whereas the most informative items were the ones referring to youths' feelings of empathy and caring behaviors (Caprara et al., 2005).

Building on this work, we used both the Classical Test Theory and Item Response Theory approaches to analyze the prosociality scale's psychometric properties in the U.S. Bolivia, and Italy. Findings with both CCT and IRT approaches attest to the good properties of a reduced 13-item scale. In particular, IRT analyses performed in each country on the reduced set of items demonstrated the effectiveness of the scale in measuring individual differences in prosociality, particularly for individuals characterized by an intermediate or low level of prosociality.

Items' information varied widely across the three countries, however. Items measuring adults' attempts to take care of others were the most informative in Italy, as well as in the U.S. and Bolivia, as found in a previous study by Caprara et al. (2005) in an Italian sample. Information curves of these items were generally peaked, with the highest degree of information associated with item 13. Only item 16 referring to empathic feelings provided uniformly low information across the three countries

The negative values of the difficulty parameters for all the items suggested that they were relatively easy to endorse. In the Italian sample, the easiest items referred to helping and caring behaviors, whereas in the Bolivian and American samples, items referring to helping behaviors and empathic feelings were most readily endorsed. The low difficulty for items pertaining to helping behavior in all three countries reflects the

high endorsement that people expressed for this component of prosocial responding, probably due to its highly normative value and, as a consequence, its high social desirability.

We also used IRT to investigate potential cultural differences in the items' functioning. Individuals' tendencies to act prosocially appear to arise from a host of socializing experiences. As a consequence, specific behavioral indicators may have somewhat different meanings in diverse cultures because each culture favors different socialization processes that may confer different meanings to the same behavior. Hence, helping, sharing, caring or empathic feelings can play a different role in discriminating people prosociality. Differential item functioning analyses are especially desirable for comparing any measure across very diverse countries.

We implemented three pairs of comparisons (Italy vs. U.S.; Italy vs. Bolivia; and U.S. vs. Bolivia) and, for each item, compared the difficulty and discrimination values. There were few differences in the items' discriminative value. Item 7, referring to helping behavior, was less discriminative in the U.S. than in both Italy and Bolivia; item 14, referring to empathic feelings, discriminated less in Bolivia than in both Italian and U.S. samples. Finally, item 13, referring to caring behavior, discriminated less in Bolivia than in Italy.

A larger number of significant differences were found in regard to the items' difficulty parameter. Items concerning empathic feelings and caring behavior were more readily endorsed by Italian young adults, whereas items concerning immediate helping behavior were more easily endorsed by American youth. Italians, compared to Bolivians, also more easily endorsed items

concerning empathic feelings; in contrast, items concerning helping and sharing behaviors were more readily endorsed by Bolivian adults than Italians. Finally, items concerning empathic feelings and helping behavior were more readily endorsed by American than Bolivian youths, whereas items referring to sharing and caring behavior were more easily endorsed by Bolivian adults.

These differences suggest that the prosocial behaviors assessed by various items occur and are valued differently in the three countries. Nevertheless, the dichotomy between individualism and collectivism seems to reduce the complexity of each culture in a overly simple classification (Turiel & Wainryb, 1994) and cannot completely explain the results. As living conditions may make certain needs more salient than others, the pursuit of others' well-being may imply different priorities and, consequently, different actions. It is likely that in societies where material resources are still scarce, sharing is particularly valued as a sign of prosociality. In post-materialistic societies, where self-actualization needs become more salient, and the satisfaction of material needs is easily discounted, empathy and caring may be particularly valued (Inglehart, 1996). Because different priorities may make people more sensitive to certain needs than others and certain aspects of prosocial behavior can be more easily accessible than others, social desirability and self-presentation in concert may induce respondents to endorse easily items that refer to specific facets of prosociality that are particularly valued in their culture.

American young adults generally seemed to have more facility in enacting direct helping behaviors than youth from the other two countries; this may be at least partially due to the strong emphasis on

voluntarism that youth often encounter when they are at school due to voluntarism being valued at both the economical and political levels. Both Italian and American youth seemed to have more facility in dealing with vicarious emotions and in taking others' perspective, whereas Bolivians more readily endorsed items concerning sharing behaviors. Perhaps living in very large families and in conditions of scarce material resources, as it typical for most of Bolivian young adults, leads people to assign special value to sharing. In contrast, living in conditions of affluence, as in Italy and U.S., may lead to appreciate the importance of behaviors aimed to meet needs of acceptance, comprehension and emotional support, beyond the mere satisfaction of material needs. In accordance with the reasoning of Inglehart (1996), it is likely that the transition from materialistic values to post-materialistic ones, and the consequent priority given to personal and interpersonal well-being, reciprocity and empathic understanding are more characteristic of people in countries such as the U.S. and Italy than in Bolivia, which is still is one of the poorest and politically instable countries in South America.

Despite the numerous differences, it was still possible to identify one invariant item across groups that can be used as an "anchor item" to rescale the values for respondents from different countries on a common metric and make comparisons among them (Reise et al., 1993). Moreover, the proposed scale could be used in the three countries since it showed good psychometrics properties cross-culturally both in CTT and IRT models. Thus, the findings attest to the validity and reliability of a 13-item reduced version of the proposed scale for evaluating young adults' prosociality. We believe that the scale is a unique instrument that was needed since long for comparisons within and across countries along a substantial portion of life span

There are a number of limitations that should be addressed by future studies. It is desirable that future studies ascertain the generalizability of our findings at different stages of late adolescence and adulthood, as well as in other cultural contexts. Whereas we have no reason to expect significant differences between young adults, middle-aged and old people in countries like U.S. and Italy, it is possible that there are larger distances between the young and old in more traditional cultures and in countries that just recently started the modernization process.

REFERENCES

Allik, J., Realo, A. (2004). Individualism-collectivism and social capital. *Journal of Cross-Cultural Psychology*, 35, 29-49.

Bandura, A., Caprara, G. V., Barbaranelli, C., Gerbino, M., & Pastorelli, C. (2003). Impact of affective self–regulatory efficacy on diverse spheres of functioning. *Child Development,74*, 1–14.

Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of*

social psychology (Vol. 2, pp. 282–316). Boston: McGraw Hill.

Batson, D. (1991). *The altruism question: Toward a social-psychological answer*. Hillsdale, NJ: Erlbaum.

Bingenheimer, J. B., Raudenbush, S. W., Leventhal, T., & Brooks-Gunn, J. (2005). Measurement equivalence in two dimensions of children's home environments. *Journal of Family Psychology, 19*, 441-455.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Cermak, I., & Rosza, S. (2001). Facing guilt: Role of negative affectivity, need for reparation, and fear of punishment in leading to prosocial behavior and aggression. *European Journal of Personality, 15*, 219-237.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. (2000). Prosocial foundations of children's academic achievement. *Psychological Science, 11*, 302–306.

Caprara, G. V., & Pastorelli, C. (1993). Early emotional instability, prosocial behavior, and aggression: Some methodological aspects. *European Journal of Personality, 7*, 19–36.

Caprara, G. V., & Steca, P. (2005). Self-efficacy beliefs as determinants of prosocial behavior conducive to life satisfaction across ages. *Journal of Social & Clinical Psychology, 24*, 191-217.

Caprara, G. V., Steca, P., Zelli, A., & Capanna, C. (2005). A new scale for measuring adults' prosocialness. *European Journal of Psychological Assessment, 21*, 77-89.

Carlo, G., & Randall, B. A. (2002). The development of a measure of prosocial behaviors for late adolescents. *Journal of*

Youth and Adolescence, 31, 31-44.

Carlo, G., Roesch, S. C., Knight, G. P., & Koller, S. H. (2001). Between- or within-culture variation? culture group as a moderator of the relations between individual differences and resource allocation preferences. *Journal of Applied Developmental Psychology, 22*, 559-579.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial Development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds). *Handbook of child psychology*,Vol.3, Social, emotional, and personality development (6th ed.). (pp. 646-718). Hoboken, NJ, US: John Wiley & Sons Inc.

Eisenberg, N., Guthrie, I.K., Cumberland, A., Murphy, B.C., Shepard, S.A., Zhou, Q., & Carlo, G. (2002). Prosocial development in early adulthood: A longitudinal study. *Journal of Personality and Social Psychology, 82,* 993–1006.

Eisenberg, N., Guthrie, I. K., Murphy, B.C., Shepard, S.A., Cumberland, A., & Carlo, G. (1999). Consistency and development of prosocial dispositions: A longitudinal study. *Child Development, 70*, 1360–1372.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Applications to psychological measurement*. Homewood, IL: Dow Jones-Irwin.

Hambleton, R. K., Swaminathan, H., & Rogers H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Hofstede, G.H. (1980). *Culture Consequences: International Differences in*

Work-related Values. Sage Publications, London.

Hui. C.H. (1988). Measurement of individualism-collectivism. *Journal of Research in Persona1ity*, 22, 17-36.

Inglehart, R. (1996). *Modernization and post-modernization: Cultural economic and political change in 43 societies.* Princeton, NJ: Princeton University Press

Kemmelmeier, M., Jambor, E. E., & Letner, J. (2006). Individualism and Good Works: Cultural Variation in Giving and Volunteering Across the United States. *Journal of Cross-Cultural Psychology*; 37(3): 327-344.

Keyes, C. L. M. (1998). Social well-being. *Social psychology quarterly, 61*(2), 121-140.

Krebs, D. L., & Van Hesteren, F. (1994). The development of altruism: Toward an integrative model. *Developmental Review, 14,* 103–158.

McDonald, R. P. (1985). Comments on D. J. Bartholomew, "Foundations of factor analysis: Some practical implications". *British Journal of Mathematical and Statistical Psychology, 38*, 134-137.

Midlarsky, E., Hannah, M. E., & Kahana, E. (1984). *The Altruism Orientation Scale*. Working paper, Center for the Study of Development and Aging, University of Detroit.

Muraki, E. (1996). A generalized partial credit model. In W. J. Van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 153–164). New York: Springer.

Muraki, E., & Bock, R.D. (1993). *PARSCALE: IRT based test scoring and itemanalysis for graded open-ended exercises and performance tasks*. Chicago: Scientific Software Int.

Oyserman, D., Coon, H.M., & Kemmelmeier, M. (2002). Rethinking Individualism and Collectivism: Evaluation of Theoretical Assumptions and Meta-Analyses. *Psychological Bulletin, 128*, 3-72.

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology, 56*, 365-392.

Pilgrim, C., & Rueda-Riedle, A. (2002). The importance of social context in cross-cultural comparisons: First graders in colombia and the united states. *Journal of Genetic Psychology, 163*, 283-295.

Piliavin, J. A., (2003). Doing well by doing good: Benefits for the benefactor. In C. L. M. Keyes, & J. Haidt (Eds.), *Flourishing: Positive psychology and the life well-lived.* (pp. 227-247). Washington, DC, US: American Psychological Association.

Reise, S. P., Ainsworth, A. T., & Haviland, M. G. (2005). Item response theory: Fundamentals, applications, and promise in psychological research. *Current Directions in Psychological Science, 14*, 95-101.

Reise, S. P., & Henson, J. M. (2003). A discussion of modern versus traditional psychometrics as applied to personality assessment scales. *Journal of personality assessment, 81*, 93-103.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*, 552-566.

Rushton, J. P., Chrisjohn, R. D., & Fekken, G. C. (1981). The altruistic

personality and the self-report altruism scale. *Personality and Individual Differences, 2*, 293-302.

Salzberger, T., Sinkovics, R. R., & Schlegelmilch, B. B. (1999). Data Equivalence in Cross-Cultural Research: A Comparison of Classical Test Theory and Latent Trait Theory Based Approaches, *Australian Marketing Journal*, 7(2), 23-38.

Schroeder, D. A., Penner, L. A., Dovidio, J. F., & Piliavin, J. A. (1995). *The psychology of helping and altruism: Problems and puzzles*. New York: McGraw Hill.

Shkodriani, G., & Gibbons, J. (1995). Individualism and collectivism among university students in Mexico and the United States. *The Journal of Social Psychology*, 135, 765-772.

Schwartz, S. H. (1990). Individualism-collectivism: Critique and

Van Willigen, M. (2000). Differential benefits of volunteering across the life course. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, 55, 308-318.

proposed refinements. *Journal of Cross-Cultural Psychology*, 21(2), 139-157.

Triandis, H. C. (1990). Cross-cultural studies of individualism and collectivism. In J. Berman (Ed.), *Nebraska symposium on motivation*, 1989, 41-133. Lincoln, Nebraska: University of Nebraska Press.

Triandis, H.C., & Gelfand, M.J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology*, 74, 118-128.

Turiel, E., & Wainryb, C. (1994). Social reasoning and the varieties of social experiences in cultural contexts. In H. W. Reese (Ed.), *Advances in Child Development and Behavior*, 289–326. San Diego, CA: Academic Press.

Van der Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern item response theory*. New York: Springer.

Waterman, A. S. (1981). Individualism and interdependence. *American Psychologist, 36*(7), 762-773.

ANNEX

Table 2. Factor loadings for the one-factor solution, separately
for the three countries.

	Ital	U.S.	Bolivia
1) I am pleased to help my friends/colleagues in their activities	.60	.55	.63
2) I share the things that I have with my friends	.54	.52	.55
3) I try to help others	.71	.67	.70
4) I am available for volunteer activities to help those who are in need	.50	.54	.58
5) I am emphatic with those who are in need	.65	.64	.61
6) I help immediately those who are in need	.71	.62	.61
7) I do what I can to help others avoid getting into trouble	.69	.54	.65
8) I intensely feel what others feel	.67	.61	.66
9) I am willing to make my knowledge and abilities available to others	.66	.59	.66
10) I try to console those who are sad	.68	.66	.65
11) I easily lend money or other things	.53	.52	.47
12) I easily put myself in the shoes of those who are in discomfort	.68	.66	.50
13) I try to be close to and take care of those who are in need	.80	.73	.70
14) I easily share with friends any good opportunity that comes to me	.64	.62	.60
15) I spend time with those friends who feel lonely	.63	.64	.62
16) I immediately sense my friends' discomfort even when it is not directly communicated to me	.59	.51	.61

15

Table 3. Items' Slope (first column) and Location (second column) parameters estimated under the generalized-partial credit model, separately for the three countries.

Item	16 Item solution						13 Item solution					
	ITALY		U.S.		BOLIVIA		ITALY		U.S.		BOLIVIA	
1	1.13	-1.20	1.02	-1.76	1.15	-.98	1.04	-1.25	.94	-1.83	1.07	-1.02
2	.97	-1.28	.92	-1.73	.91	-1.47	-	-	-	-	-	-
3	1.97	-1.29	1.66	-1.60	1.60	-1.11	1.87	-1.32	1.54	-1.66	1.48	-1.14
4	.61	-.06	.73	-.37	.87	-.85	-	-	-	-	-	-
5	1.34	-.87	1.21	-1.14	1.06	-.79	1.43	-.86	1.20	-1.16	1.05	-.80
6	1.81	-.74	1.13	-.62	1.09	-1.04	1.81	-.74	1.09	-.63	1.09	-1.04
7	1.57	-.35	.87	-1.22	1.19	-.58	1.56	-.35	.87	-1.23	1.24	-.58
8	1.39	-.82	1.01	-.40	1.22	-.58	1.48	-.81	1.07	-.39	1.25	-.58
9	1.53	-1.10	1.12	-1.44	1.33	-.94	1.47	-1.12	1.11	-1.45	1.35	-.94
10	1.73	-1.52	1.41	-1.34	1.30	-1.32	1.78	-1.52	1.50	-1.32	1.31	-1.32
11	.73	-.27	.66	-.61	.57	-.75	-	-	-	-	-	-
12	1.49	-.88	1.27	-.70	.65	-.13	1.54	-.87	1.32	-.70	.65	-.13
13	2.78	-.67	1.77	-.74	1.53	-.81	2.79	-.68	1.79	-.74	1.59	-.80
14	1.34	-1.09	1.20	-1.39	1.06	-1.32	1.23	-1.14	1.14	-1.43	1.05	-1.33
15	1.39	-1.37	1.26	-1.22	1.16	-1.21	1.34	-1.40	1.28	-1.22	1.16	-1.21
16	1.12	-1.37	.81	-1.47	.98	-.80	1.14	-1.37	.84	-1.45	1.02	-.79

Table 4: Item Slope and Location Parameters and Differential Item Functioning Tests for the comparisons: Italy (I) vs. U.S. (U), Italy vs. Bolivia (B) and U.S. vs. Bolivia.

	Slope			Location			Slope			Location			Slope			Location		
Item	I	U	X^2	I	U	X^2	I	B	X^2	I	B	X^2	U	B	X^2	U	B	X^2
1	2.85	3.22	1.67	-.40	-.68	-12.95**	2.59	3.11	2.24	-.47	-.44	1.07	1.00	1.07	1.04	-1.98	-1.18	13.32**
3	5.05	5.35	.87	-.46	-.56	-6.37**	4.46	4.33	-.41	-.52	-.50	.61	1.61	1.47	-1.41	-1.65	-1.37	6.34**
5	3.85	4.09	.88	-.31	-.36	-3.19	3.00	3.34	1.39	-.38	-.31	3.05	1.14	1.19	.55	-1.25	-.94	6.39**
6	4.62	3.88	3.05	-.30	-.14	9.66**	3.97	3.40	-2.38	-.30	-.43	-6.53**	1.10	1.13	.38	-.60	-1.34	14.25**
7	3.98	3.08	4.52**	-.12	-.36	-13.42**	3.57	3.76	.72	-.10	-.27	-8.05**	.92	1.24	3.97**	-1.11	-.86	4.90**
8	3.78	3.78	.02	-.32	-.06	15.77**	3.22	3.90	2.41	-.33	-.24	4.30**	1.09	1.27	2.34	-.37	-.86	10.58**
9	3.86	4.02	.63	-.42	-.44	-.88	3.44	4.09	2.23	-.45	-.41	1.66	1.18	1.32	1.69	-1.36	-1.22	2.89
10	4.89	5.14	.71	-.49	-.43	3.87**	4.05	3.93	-.39	-.57	-.59	-1.04	1.49	1.39	-.97	-1.32	-1.52	-4.34**
12	4.14	4.61	1.61	-.30	-.18	7.77**	3.15	2.19	-5.97**	-.34	-.08	10.73**	1.26	0.78	-9.11**	-.71	-.44	5.07**
13	7.18	6.32	2.18	-.26	-.20	5.49**	6.06	4.91	-3.30**	-.26	-.36	-6.53**	1.78	1.68	-.94	-.70	-1.10	10.47**
14	3.60	3.90	1.16	-.38	-.46	-4.80**	3.08	2.96	-.53	-.40	-.63	-9.33**	1.18	1.07	-1.54	-1.35	-1.64	-5.09**
15	3.86	4.43	1.95	-.51	-.36	8.81**	3.33	3.39	.21	-.53	-.54	-.58	1.32	1.16	-1.96	-1.13	-1.56	-8.45**
16	3.16	2.85	1.50	-.45	-.48	-1.47	2.71	3.01	1.31	-.52	-.35	6.93**	.86	1.06	2.62	-1.47	-1.00	8.09**

** $p<.001$