

This article was downloaded by: [84.126.115.248]

On: 10 February 2013, At: 08:18

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH,
UK



European Journal of Developmental Psychology

Publication details, including instructions for authors
and subscription information:

<http://www.tandfonline.com/loi/pedp20>

The structure and correlates of a measure of prosocial moral reasoning in adolescents from Spain

Gustavo Carlo ^a, Maria Vicenta Mestre ^b, Meredith
McGinley ^c, Ana Tur-Porcar ^b, Paula Samper ^b & Cara
Streit ^a

^a Human Development and Family Studies, University
of Missouri, Columbia, MO, USA

^b Department of Psychology, Universitat de Valencia,
Valencia, Spain

^c Department of Graduate Psychology, Chatham
University, Pittsburgh, PA, USA

Version of record first published: 29 Jan 2013.

To cite this article: Gustavo Carlo , Maria Vicenta Mestre , Meredith McGinley , Ana Tur-Porcar , Paula Samper & Cara Streit (2013): The structure and correlates of a measure of prosocial moral reasoning in adolescents from Spain, European Journal of Developmental Psychology, DOI:10.1080/17405629.2012.762909

To link to this article: <http://dx.doi.org/10.1080/17405629.2012.762909>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes.
Any substantial or systematic reproduction, redistribution, reselling, loan, sub-

licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

The structure and correlates of a measure of prosocial moral reasoning in adolescents from Spain

Gustavo Carlo¹, Maria Vicenta Mestre², Meredith McGinley³, Ana Tur-Porcar², Paula Samper², and Cara Streit¹

¹Human Development and Family Studies, University of Missouri, Columbia, MO, USA

²Department of Psychology, Universitat de Valencia, Valencia, Spain

³Department of Graduate Psychology, Chatham University, Pittsburgh, PA, USA

The study was designed to examine the structure and correlates of a measure of prosocial moral reasoning in a sample of young and middle-aged adolescents. Participants were 1,556 students (53% male, $M_{\text{age}} = 13.12$ years, $SD = 0.87$) from Valencia, Spain, who completed paper-and-pencil measures of prosocial moral reasoning (PROM), empathy, prosocial behaviours, and aggression. As expected, a series of confirmatory factor and structural equation modelling analyses revealed a four-factor solution of the PROM to have the best fit (as compared to alternative models) among Spanish youth (across gender and grade). Moreover, higher level and other-oriented forms of prosocial moral reasoning were generally positively related to empathy and prosocial behaviours, and negatively related to aggression. In contrast, generally, lower level and self-focused modes of prosocial moral reasoning were negatively related to such prosocial tendencies, and positively related to aggression. Discussion focuses on the usefulness of the PROM, its relations to theoretically relevant correlates, and its usefulness to study the development and universality of prosocial moral development.

Keywords: Morality; Prosocial behaviours; Aggression; Empathy; Culture.

Correspondence should be addressed to Gustavo Carlo, 410 Gentry Hall, University of Missouri, Columbia, MO, 65211, USA. Email: carlog@missouri.edu

Funding support was provided by an I+D Projects for Research Teams of Excellence, PROMETEO Program (Reference: PROMETEO/2011/2009), Department of Education, Generalitat Valenciana to MVM.

The authors greatly appreciate the assistance of Elisabeth Malonda Vidal and the cooperation of the educational centres in Valencia.

Humans have the ability to reason about care-based, helping situations in early childhood, and intra-individual stability and change in these abilities are evident across childhood, adolescence, and into adulthood (Carlo, 2006; Eisenberg, 1986). Such reasoning, referred to as prosocial moral reasoning, occurs in helping opportunity contexts where there is conflict between one's needs or desires and those of another, in the relative absence of formal laws or rules. Prior research suggests some similarities in prosocial moral reasoning development across cultures though most research has been conducted in Western (North American) industrialized societies (Carlo, 2006; Eisenberg, Fabes, & Spinrad, 2006). However, limited research exists in some European countries (e.g., Boehnke, Silbereisen, Eisenberg, Reykowski, & Palmonari, 1989; Mestre, Frías, Samper, & Tur, 2002; Skoe et al., 1999), in Brazil (Carlo, Koller, Eisenberg, DaSilva, & Frohlich, 1996; Eisenberg, Zhou, & Koller, 2001), in Turkey (Kumru, Carlo, Mestre, & Samper, 2012), and in Papua New Guinea (Tietjen, 1986). These researchers generally find that there are many commonalities in the forms of prosocial reasoning across culture groups though the developmental emergence and frequency of some forms of prosocial reasoning may differ (see Eisenberg et al., 2006). However, studies of prosocial moral reasoning in other cultures are needed to further examine the development and universality of prosocial morality.

Spain has similarities to other Western, industrialized, social democracies, and is a member of the European Union. Although one must be cautious in over generalizing the characteristics of people from any society (Turiel, 2006), researchers often do so to contextualize their findings. For example, Spain scores higher on individualism and lower on masculinity than other Latino cultures (e.g., Brazil), but scores substantially lower on individualism than the United States (Hofstede, 1984; Oyserman, Coon, & Kemmellmeier, 2002). Thus, Spain is generally considered relatively moderate on indices of collectivism (Basabe et al., 2000; Fernández-Berrocal, Salovey, Vera, Ramos, & Extremera, 2001). Moreover, as in other Latino cultures, two prosocial socialization influences, the family and religion (primarily Christianity), are highly valued and play important major roles in Spanish customs and traditions (Centro de Investigaciones Sociológicas, 2004; Elzo, 2004). Given the interest in understanding the universality of moral development theories (Malti & Keller, 2010; Turiel, 2006), the present study examined prosocial moral reasoning in a sample of early adolescents from a relatively collectivist, and strongly religious and family-oriented society.

With regard to prosocial moral reasoning, prior developmental research shows that such reasoning is characterized by hedonistic, needs-oriented, and approval-oriented considerations in early and middle childhood, and global stereotyped, and empathic and internalized principled considerations in late childhood, adolescence, and adulthood (Eisenberg et al., 2006). Furthermore, prosocial moral reasoning is hypothesized and found to develop across time as a result of changes in cognitive developmental abilities (e.g., perspective taking) and socialization processes (Eisenberg, 1986). However, empathic and needs-oriented reasoning are

not prevalent in responses to harm-based, prohibition-oriented dilemmas (i.e., Kohlberg dilemmas) and references to punishment are not prevalent in responses to care-based dilemmas (Eisenberg, 1986). Moreover, the emphasis on caring for others and interpersonal relationships in prosocial moral reasoning is distinct from the emphasis on issues of justice and rights in harm-based moral reasoning, and gender differences in prosocial moral reasoning are not uncommon (Carlo, 2006; Eisenberg et al., 2006). These findings suggest that prosocial moral reasoning is somewhat distinct from harm-based moral reasoning (see Eisenberg, 1986; Gilligan, 1982; Skoe et al., 1999).

Despite the important distinctions regarding prosocial and harm-based moral reasoning, moral development scholars have asserted the need for integrative theories of morality that account for development in moral cognitions, and associated emotions and behaviours (Arsenio & Lemerise, 2004; Gibbs, 2003; Hoffman, 2000; Malti, Gummerum, & Keller, 2008). For example, reasoning about moral dilemmas is expected to promote other-oriented sociocognitive thinking (e.g., perspective taking) and emotions (e.g., empathy), tendencies that have been shown to facilitate prosocial behaviours, and mitigate hurtful, aggressive behaviours. Furthermore, higher levels of prosocial moral reasoning often contains references to reducing human suffering and caring for others— aspects of empathic and prosocial behavioural responding. Conversely, empathy might induce or prime more elaborated forms of moral cognitions and vice versa (Hoffman, 2000; Turiel, 2006). Indeed, there is ample evidence that moral reasoning is linked to moral emotions such as empathy (i.e., feeling the same as another), and to sociomoral behaviours such as prosocial (i.e., actions intended to benefit others) and aggressive (i.e., actions that harm or injure others) behaviours (Carlo, 2006; Eisenberg et al., 2006; Lemerise & Arsenio, 2000; Malti & Keller, 2009). However, further research directly examining these correlates in non-North American populations is necessary to examine the feasibility of universal, integrative theories of morality (Malti & Keller, 2010; Turiel, 2006).

The present study addresses existing gaps in prosocial moral reasoning research. First, fundamental questions regarding the development and correlates of prosocial moral reasoning in non-North American cultures remain. Within-culture research in other societies is important in its own right to understand prosocial development among youth in those cultures. Second, our ability to address questions regarding the universal or culture-specific nature of moral development is hampered given the relative scarcity of studies outside of North America. And third, there are uncertainties regarding the reliability and validity of existing measures of prosocial moral reasoning to use across different cultures and within-culture groups (e.g., across different ages, gender). Indeed, this latter concern undermines our ability to infer with confidence any cross-cultural research findings. The present study addressed these gaps by examining the psychometric properties of a commonly used, measure of prosocial moral reasoning.

Rigorous studies of the psychometric properties of measures of moral reasoning are rare. One reason is that few standardized, objective measures of moral reasoning exist. Traditionally, semi-structured interview measures were considered the strongest and most reliable instruments to assess moral reasoning; though there was considerable debate on this issue (Rest, 1979). Although there are advantages and disadvantages to the use of specific assessment instruments, there are some important advantages to paper-and-pencil, standardized response measures over open-ended, interview measures. For example, standardized response measures can be more easily subjected to stringent confirmatory tests of the internal factor model fit of the measure, to direct comparisons of the hypothesized latent model to alternative models, and to equivalence tests of the factor structure, and relations of the measure to other theoretically relevant constructs, across different groups (e.g., gender, age). Moreover, such measures have relatively minimal interviewer and coding biases as compared to interview measures. Based on Eisenberg's interview measure of prosocial moral reasoning, a paper-and-pencil measure of prosocial moral reasoning (PROM) was developed (Carlo, Eisenberg, & Knight, 1992). Several studies have demonstrated adequate reliability (e.g., test-retest) and validity (e.g., convergent, discriminant) properties of the PROM (see Carlo, 2006). However, only one previous study examined the factor structure fit of the PROM.

Carlo, McGinley, Roesch, and Kaminski (2008) examined the latent factor structure of a measure of prosocial moral reasoning (the PROM) in a sample of Brazilian and European-American adolescents. The investigators found evidence that a four-factor model of the PROM had adequate fit, had significantly better fit than alternative factor models, and that the PROM demonstrated measurement and functional equivalence (to theoretically relevant constructs) across nationality groups, age, and gender. These findings suggest that research findings using the PROM in Brazilian and European-American adolescents are relatively unlikely to be due to measurement artefact of the PROM. Thus, those findings enhance our ability to interpret with confidence research findings using the PROM with Brazilian and European-American adolescents. However, the absence of similar research on the psychometric properties of the PROM limits our ability to interpret research findings regarding prosocial tendencies in other cultures (e.g., Spain).

Based on the above cited theory and research, several hypotheses were developed. First, the four-factor model of prosocial moral reasoning (i.e., hedonistic, approval-oriented, needs-oriented, and stereotyped/internalized) was expected to be a good fitting model for Spanish adolescents. Second, the four-factor model was expected to be significantly better fit than alternative models (i.e., one-, two-, three-, or five-factor). Third, the four-factor model was expected to demonstrate equally good fit across gender, and across grades. And fourth, developmentally lower level hedonistic and approval-oriented prosocial moral reasoning should be positively associated with aggressive

behaviour, and negatively associated with prosocial behaviours and empathy. In contrast, needs-oriented and developmentally higher level, stereotyped and empathic/internalized prosocial moral reasoning ought be positively linked to prosocial behaviours and empathy, and negatively linked to aggressive behaviours.

METHOD

Participants

The sample was 1,556 students (53% male, $M_{\text{age}} = 13.12$ years, $SD = 0.87$) from schools in Valencia, Spain (7th graders, $n = 799$, $M_{\text{age}} = 12.60$ years, $SD = 0.68$; 396 girls; 8th graders, $n = 757$, $M_{\text{age}} = 13.66$ years, $SD = 0.70$; 341 girls). The 36 participating schools had either more than 30% ($n = 13$), between 20 and 30% ($n = 13$), or less than 20% ($n = 10$), immigrant children in classes. All students were recruited via letters sent to homes. Mothers' education was well distributed (38% less than high school diploma, 27% high school diploma or equivalent, 23% at least some university education). The majority of the sample self-identified as Spanish (83%; 10% from Latin America, e.g., Colombia, Argentina; 4% from Eastern European countries, e.g., Bulgaria, Romania). Surveys were administered to students (all instruments were self-reports) in classrooms (approximately 45 minutes completion time).

Measures

All measures have been adapted for use and validated in samples of adolescents from Spain (e.g., Del Barrio, Moreno, & López, 2001; Mestre, Pérez-Delgado, Frías, & Samper, 1999; Mestre et al., 2002; Pereña & Seisdedos, 1997).

Prosocial moral reasoning. The Prosocial Reasoning Objective Measure (PROM; Carlo et al., 1992) was used to assess prosocial reasoning. The PROM contains stories designed to invoke a conflict between the actor's needs, wants, and desires and those of others. Due to time limitations, only three stories were used. The stories depicted situations in which characters: (i) help a peer who is being teased versus incurring rejection from peers; (ii) donate blood to a needy other at the cost of losing money at work and school; and (iii) go to the beach with friends or help a peer study for a math exam. A sample story is as follows:

Sandy (Begoña) was a student at school. One day Sandy was walking into her new class early and saw an older girl teasing and making fun of another girl's clothes. The girl was crying. There was no one else around and Sandy did not know the girls very well, but she had heard that the girl that was being teased was very poor and the older girl had a lot of friends. Sandy thought that maybe she

should try to stop the older girl but she was afraid that the older girl and her friends might pick on her and tease her also.

For each story, youth indicated whether the protagonist should or should not help and then indicated the importance of five different reasons (on a 5-point scale, anchored by 1 = *Not at all* and 5 = *Greatly*). Based on the order of progression from less to more mature forms of prosocial moral reasoning (see Eisenberg, 1986), each story included reasons reflecting hedonistic reasoning (e.g., “It depends whether Sandy can find other friends to do things with in school”), needs-oriented reasoning (e.g., “It depends whether the other girl is crying a lot”), approval-oriented reasoning (e.g., “It depends whether Sandy’s classmates would approve of what she does”), stereotyped reasoning (e.g., “It depends whether Sandy thinks the older girl is mean or not”), and internalized reasoning (e.g., “It depends whether Sandy thinks that she is doing what she believes she should do”).

To compute PROM scores, ratings that corresponded to each of the five types of prosocial moral reasoning were summed and averaged to obtain a raw score. Then, to obtain an individual’s preference for one type relative to the other types (see, e.g., Carlo et al., 1992), a proportion score was calculated by dividing each of the raw scores by the sum of the PROM scores. Other research has shown adequate reliabilities and validity of the PROM to use with Spanish adolescents (e.g., Carlo, Mestre, Samper, Tur, & Armenta, 2010; Mestre et al., 2002).

Empathy. The Inventory of Empathy for Children and Adolescents (Bryant, 1982; Mestre et al., 1999) assessed their tendency to feel sorrow for others. Adolescents responded to 15 items (e.g., “I feel sad when I see a girl who has no one to play with”) on a yes/no scale. Items were summed and averaged ($\alpha = .70$).

Prosocial behaviours, aggression. A multidimensional instrument (Caprara & Pastorelli, 1993; Del Barrio et al., 2001) assesses prosocial behaviours and overt aggression. Items are on a 3-point scale (anchored by 1 = *Never* and 3 = *Frequently*). After dropping some items due to low inter-item correlations, the 15 prosocial behaviour items (“I console those who are sad”; $\alpha = .70$) and 15 overt aggression items (“I kick and punch others”; $\alpha = .83$) were summed and averaged separately.

RESULTS

Data analytic plan

Overall model. Using confirmatory factor analysis (CFA) in Mplus version 6.1 (Muthén & Muthén, 1998–2010), several initial competing models of the PROM were tested and compared for fit (see Table 1 for model descriptions).

TABLE 1
Tests of alternative latent factor models for the overall sample

	<i>S-B</i> χ^2 (<i>df</i>)	<i>CFI</i>	<i>SRMR</i>	<i>RMSEA</i>	<i>AIC</i>	<i>BIC</i>
Five-factor model	492.242 (144)	.933	.039	.041	94653.946	95231.722
Four-factor model	517.581 (148)	.929	.040	.040	94676.526	95232.913
Three-factor model (a)	544.858 (151)	.924	.042	.041	94705.481	95245.818
Three-factor model (b)	866.101 (151)	.863	.055	.055	95093.295	95633.632
Two-factor model	865.858 (153)	.863	.055	.055	95089.904	95619.541

Notes: The five-factor model is Eisenberg's (1986) original model. The hypothesized four-factor model is based on prior PROM findings that combined stereotyped and internalized reasoning (Carlo et al., 2008). The three-factor models indicate: (a) the model in which needs-oriented was collapsed with the combined internalized/stereotype factor; and (b) reflect Eisenberg's (1986) alternative three developmental levels of prosocial moral reasoning. The two-factor model tested the notion that there are self- and other-oriented modes of prosocial moral reasoning.

Factorial invariance. Factorial invariance of the PROM was assessed by using multi-group confirmatory factor analysis (CFA) to fit a series of hierarchically nested factor structures (Millsap & Kwok, 2004). These analyses are designed to determine if individual items are functioning similarly across groups (e.g., gender, grade). Configural invariance allows the same set of items to form a factor in each group, but freely estimating all model parameters in each group. If the configural invariance model fits well, the similarity of the factor loading across groups is tested. Weak factorial invariance exists if the factor loadings between each item and the respective latent construct are invariant across groups. Strong factorial invariance exists if the item intercepts associated with each item are invariant across groups. Finally, strict factorial invariance exists if the unique error variances associated with each item are invariant across groups.

Construct validity equivalence. Scalar and functional equivalence can be assessed by examining whether group membership moderates the correlations between the PROM and related factors (Hui & Triandis, 1985; Knight & Hill, 1998). The first set of construct validity analyses examined group mean differences on the four PROM factors, and in the relations among the four PROM factors (i.e., scalar equivalence). Further tests were conducted using a series of analyses correlating the PROM factors with each of the construct validity variables (i.e., prosocial behaviours, empathy, or aggression; i.e., functional equivalence). These analyses first compared an unconstrained model to a model that constrained the correlations between the PROM factors and the construct validity scale, as well as the construct validity scale intercept, across groups. Similarity of correlations and intercepts across groups demonstrates functional equivalence (Knight & Hill, 1998).

Assessing model fit. The following indices were employed to evaluate model fit: (1) the Satorra–Bentler Scaled χ^2 (S–B χ^2 ; Satorra & Bentler, 1988); (2) the comparative fit index (CFI; Bentler, 1990), with values greater than .95 indicating reasonable model fit; (3) the root mean square error of approximation (RMSEA; Steiger, 1990), with values less than .06 indicating reasonable model fit; (4) the Akaike information criterion (AIC; Akaike, 1987) and the BIC Bayesian information criterion (Schwarz, 1978) to compare non-nested models (lower values suggesting better model fit; Brown, 2006); and (5) the S–B χ^2 difference test (Δ S–B χ^2 ; Satorra, 2000) to statistically compare the relative fit of nested models.

When chi-square difference tests demonstrated no significant difference between a nested model (e.g., metric invariance) and the initial model (e.g., configural invariance), this suggests no significant differences between groups (Byrne, Shavelson, & Muthén, 1989). Conversely, if a significant difference between groups occurred, modification indices (i.e., the Lagrange multiplier) were used to identify which model paths were significantly different.

Establishing the overall model

A test of a five-factor model showed four multivocal and/or poorly loading items that were subsequently dropped. Furthermore, the five-factor model indicated that early adolescents did not distinguish between the stereotype and internalized scales; indeed, these two factors were highly correlated and produced a Heywood case (i.e., correlation higher than one; see Satorra, 2000). The two- and three-factor (combining needs-oriented/stereotyped/internalized as one factor) model tests showed worse fit than the four-factor model. Therefore, we found the four-factor model to fit best (Table 1). Item loadings on the PROM ranged from .31 to .76 (standardized values). Cronbach's alpha coefficients were .60, .76, .60, and .72 for hedonistic (three items), approval-oriented (five items), needs-oriented (four items), and stereotyped/internalized (nine items), respectively.

Factorial invariance analyses

Gender. The four-factor model was then tested for gender invariance (Table 2). Configural and weak invariance (Models 1 and 2) was achieved. However, tests for Model 3 (strong invariance) and Model 4 (strict invariance) resulted in significantly poorer model fit. Model fit indices suggested that five item intercepts should be freed in Model 3, and 12 item unique variances should be freed in Model 4. These models (3a and 4a) were found to have acceptable fit.

Grade. The four-factor model was tested for invariance across grade (Table 2). Configural and weak invariance (Models 7 and 8) was achieved. However, Model

TABLE 2
Summary of the factorial invariance tests across gender groups and grade groups

	χ^2	df	$\Delta\chi^2$	Δdf	CFI	RMSEA
<i>Gender group</i>						
1. Configural invariance model (Model 1)	663.616	296	–	–	.930	.040
2. Model 1 + all factor loadings invariant	682.171	313	15.144	17	.930	.039
3. Model 2 + all item intercepts invariant	733.906	330	50.834**	17	.923	.040
3a. Model 3 – 5 item intercepts freed	698.105	325	15.768	12	.929	.038
4. Model 3 + all item unique variances invariant	831.955	346	104.380**	21	.907	.042
4a. Model 4 – twelve unique variances freed	713.845	334	14.404	9	.928	.038
5. Model 5 + all latent factor means invariant	775.191	338	52.658**	4	.917	.041
6. Model 6 + all latent factor correlations invariant	736.939	340	18.07	6	.924	.039
6a. Model 6a—one latent factor correlation freed	728.152	339	11.61*	5	.926	.038
<i>Grade group</i>						
7. Configural invariance model (Model 7)	724.855	296	–	–	–	.043
8. Model 7 plus all factor loadings invariant	744.537	313	15.75	17	.918	.042
9. Model 8 plus all item intercepts invariant	802.688	330	57.84**	17	.910	.043
9a. Model 9 – 6 item intercepts freed	761.614	324	17.02	11	.917	.042
10. Model 9 plus all item unique variances invariant	810.283	345	37.76**	21	.912	.057
10a. Model 10 – one unique variance freed	792.234	344	23.97	20	.915	.041
11. Model 10 plus all factor means invariant	825.958	348	33.52**	4	.909	.042
11a. Model 11 – two latent means freed	795.034	346	2.78	2	.915	.041
12. Model 11a + all latent factor correlations invariant	813.771	352	15.14*	6	.913	.041
12a. Model 12 – one latent factor correlation freed	805.961	351	9.27	5	.914	.041

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

TABLE 3
Summary of the construct equivalence tests across gender groups and grade groups

<i>Gender group</i>		χ^2	<i>df</i>	$\Delta\chi^2$	Δdf	<i>CFI</i>	<i>RMSEA</i>
Prosocial behaviour	1. Model freely estimating the intercept and correlations	797.12	373	–	–	.923	.038
	2. Model 1 + the intercept and correlations constrained	904.37	378	99.36**	5	.904	.042
	2a. Model 2 – prosocial behaviour mean freed	803.11	377	5.65	4	.922	.038
Empathy	1. Model freely estimating the intercept and correlations	779.883	373	–	–	.926	.037
	2. Model 1 + the intercept and correlations constrained	1096.662	378	275.22**	5	.869	.049
	2a. Model 2 – empathy mean freed	786.034	377	5.82	4	.925	.037
Aggressive behaviour	1. Model freely estimating the intercept and correlations	784.509	373	–	–	.924	.038
	2. Model 1 + the intercept and correlations constrained	833.791	378	45.32**	5	.915	.039
	2a. Model 2 – aggressive behaviour mean freed	788.019	377	3.02	4	.924	.037
Prosocial behaviour	1. Model freely estimating the intercept and correlations	880.290	385	–	–	.911	.041
	2. Model 1 + the intercept and correlations constrained	892.107	390	11.70*	5	.910	.041
	2a. Model 2 – correlations freed	884.615	389	4.17	3	.911	.041
Empathy	1. Model freely estimating the intercept and correlations	866.992	385	–	–	.914	.040
	2. Model 1 + the intercept and correlations constrained	872.276	390	5.25	5	.914	.040
Aggressive behaviour	1. Model freely estimating the intercept and correlations	843.283	385	–	–	.915	.039
	2. Model 1 + the intercept and correlations constrained	852.780	390	8.65	5	.915	.039

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

9 (strong invariance) and Model 10 (strict invariance) tests resulted in significantly poorer model fits. Model fit indices suggested that six item intercepts should be freed in Model 9, and one item unique variance should be freed in Model 10. These models (9a and 10a) were found to have good fit.

Construct equivalence analyses

Gender. Constraining all latent factor means to be invariant resulted in significantly poor fit (Model 5; see Table 2). Model fit indices suggested that all four means be freed across gender. Girls had significantly higher stereotyped/internalized and needs-oriented means, and lower hedonistic and approval-oriented means, than boys. Next, we constrained the latent factor correlations (Model 6), but the $\Delta\chi^2$ was significant. Model fit indices suggested that one factor correlation, between stereotyped/internalized and needs-oriented, was significantly different across gender and should be freely estimated. This correlation was significantly higher among boys than girls, though both correlations were significant and positive. The chi-square difference test of Model 6a was still significant, but model fit indices were comparable to Model 4a and no additional modifications were suggested.

Construct validity analyses of the PROM across gender groups were then conducted (Table 3). For prosocial behaviours, the fully constrained model yielded a significant $\Delta\chi^2$. Modification indices revealed that mean scores on prosocial behaviours were significantly ($p < .05$) higher for girls. However, the correlations across gender groups remained equivalent. Hedonistic reasoning was negatively related, and needs-oriented and internalized reasoning were positively related, to prosocial behaviours. Similarly, the $\Delta\chi^2$ test revealed a significant gender difference in empathy. Modification indices revealed that empathy was significantly higher for girls. However, the correlations across gender groups remained equivalent. Hedonistic reasoning was negatively, and needs-oriented and internalized reasoning were positively, related to empathy. Finally, the $\Delta\chi^2$ test revealed a significant gender difference in aggression. Modification indices revealed that aggression was significantly higher for boys. However, the correlations across gender remained equivalent. Hedonistic reasoning was positively, and needs-oriented and internalized reasoning were negatively, related to aggression.

Grade. Constraining all latent factor means to be invariant resulted in a model that had significantly poor fit (Model 11; see Table 2). Model fit indices suggested that two means be freed across grade groups. Older students had significantly ($ps < .05$) higher stereotyped/internalized and hedonistic means than younger students. The partial latent means model (Model 11a) had acceptable fit. Next, we constrained the latent factor correlations (Model 12), but the $\Delta\chi^2$ was significant. Model fit indices suggested that one factor correlation, internalized/needs, was

significantly different across grade and should be freely estimated. This correlation was significantly higher among younger students than older students, though both correlations were significant and positive. This partially constrained correlation model had acceptable fit.

Construct validity analyses of the PROM across grade groups were then conducted (Table 3). For prosocial behaviours, the fully constrained model yielded a significant $\Delta\chi^2$. Modification indices revealed the correlations between prosocial behaviours and both needs-oriented and stereotyped/internalized factors to be significantly different. The correlation with needs-oriented moral reasoning was significantly weaker for the older grade, whereas the correlation with stereotyped/internalized moral reasoning was significantly stronger for the older grade. However, both sets of correlations were positive and significant. Equivalently across grades, hedonistic reasoning was significantly and negatively related to prosocial behaviours. The construct equivalence tests for empathy and aggression both yielded acceptable fit. The correlations (all significant) between empathy and hedonistic and approval-oriented reasoning were negative, whereas the correlations between empathy and needs-oriented and stereotyped/internalized reasoning were positive. For aggression, there were significant correlations between aggression and hedonistic reasoning (positive), and between aggression and both needs-oriented and stereotyped/internalized moral reasoning (negative).

DISCUSSION

The present study was designed to examine the structure and equivalence of a measure of prosocial moral reasoning across Spanish young and middle-aged adolescents, across gender, and its relations to empathy, aggression, and prosocial behaviours. As expected, the findings demonstrate strong evidence of a four-factor model of the PROM as the best fit model of prosocial moral reasoning. The four-factor model also showed adequate fit across young and middle-aged adolescents and across boys and girls. Moreover, as expected, higher level and other-oriented forms of prosocial moral reasoning were generally positively related to empathy and prosocial behaviours, and negatively related to aggression. In contrast, generally, lower level and self-focused modes of prosocial moral reasoning were negatively related to such prosocial tendencies, and positively related to aggression (generally equivalent across grade and gender). The findings suggest that the PROM is a valid instrument to study prosocial moral reasoning development and its correlates in Spanish adolescents. Moreover, consistent with prior similar studies (e.g., Carlo et al., 2008; Eisenberg et al., 2001), the structure and correlates of prosocial moral reasoning (using the PROM) is similar across youth from Spain, Brazil, and the United States.

The fact that the four-factor model of the PROM showed the strongest psychometric properties provides growing support for a model of prosocial moral reasoning that reflects hedonistic, approval-oriented, needs-oriented, and stereotyped/empathic/internalized as distinct levels of prosocial moral reasoning among youth from Spain (as well as from Brazil and the United States). In addition, the PROM demonstrated good within-culture fit across young and middle adolescence and gender. Such findings indicate support for the notion that, at least among these Western, industrialized societies, a four-factor model best characterizes prosocial moral reasoning among adolescents despite the prior evidence that these specific societies may differ somewhat on their orientation towards individualism, role of family and religion, and traditional gender roles. However, future research on the psychometric properties of the PROM in much more distinct societies (e.g., Non-Western, non-industrialized) is needed to better examine the universality of prosocial morality.

Although the original proposed model of prosocial moral reasoning suggested that stereotyped was distinct from empathic and internalized reasoning (see Eisenberg, 1986), youth from Spain (and Brazil and the United States) do not seem to make such distinctions using the PROM. These findings may be due to differences between the use of an interview, open-ended measure (Eisenberg's original measure) and a paper-and-pencil, multiple-choice measure (the PROM) of prosocial moral reasoning. As noted by several scholars (see Rest, 1979), responding to multiple-choice measures of moral reasoning requires respondents to recognize moral reasons rather than to spontaneously recall and substantiate such reasons (as in interview measures). The use of a preference measure may minimize the distinction between stereotyped and internalized reasoning. Alternatively, prior research demonstrates that stereotyped and internalized modes of prosocial moral reasoning emerge during late childhood and early adolescence, and that both modes require higher level, sociocognitive skills (e.g., perspective taking; see Eisenberg et al., 2006). Thus, the similar cognitive prerequisites in these higher levels of prosocial moral reasoning may exacerbate the relations among these variables. Perhaps future longitudinal studies that examine a wider age range (with older adolescents and adults) might result in stereotyped and internalized factors as distinct factors.

Several interesting age and gender differences in the PROM were found. For example, older students scored higher than younger students on stereotyped/internalized and hedonistic reasoning. The findings for stereotyped/internalized are consistent with prior research that suggests such reasoning types to be more cognitively advanced (see Carlo, 2006). The age group difference in hedonistic reasoning is consistent with prior research that shows an increase in such reasoning from early adolescence to young adulthood (see Eisenberg et al., 2006). As in prior research (see Carlo, 2006), girls scored higher than boys on needs-oriented and stereotyped/internalized reasoning findings whereas boys score higher than girls on hedonistic and approval-oriented reasoning. These

findings, in general, are in accord with gender theories that suggest that girls are socialized to be more caring and nurturing and that boys are oriented towards instrumentalism (Gilligan, 1982).

Several limitations and concerns should be noted. First, the fact that the PROM shows similar structure across some cultures does not necessarily mitigate possible culture group differences in the emergence, the rate of development, and the frequency of use of specific types of prosocial moral reasoning. Culture group differences in prosocial moral development may still exist and further research that directly compares cultures is needed. Second, all the instruments were self-reports, which raises concerns regarding shared method variance and self-presentational demands. Although prior research suggests minimal social desirability bias in the PROM (Carlo et al., 1992), the use of multiple measures or methods that maximize anonymity (such as computer-assisted instruments) and minimize shared method variance is desirable. And third, we did not examine specific forms of aggressive (e.g., relational, proactive) and prosocial (e.g., emergency, volunteerism) behaviours. Future research on the relations between prosocial moral reasoning and specific forms of social behaviours could yield interesting findings.

Nonetheless, the present findings suggest that a four-factor PROM model is the best fitting model among Spanish adolescents, that the PROM is useful and valid to use with this population, and adds to the mounting evidence that the structure and correlates of prosocial moral reasoning is similar across Western, industrialized societies. Such findings lend more evidence to further examine the development and universality of morality (see Malti & Keller, 2010; Turiel, 2006).

Manuscript received 21 June 2012

Revised manuscript accepted 27 December 2012

First published online 29 January 2013

REFERENCES

- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*, 317–332.
- Arsenio, W. F., & Lemerise, E. A. (2004). Aggression and moral development: Integrating social information processing and moral domain models. *Child Development*, *75*(4), 987–1002.
- Basabe, N., Paez, D., Valencia, J., Rimé, B., Pennebaker, J., Diener, E., & González, J. L. (2000). Sociocultural factors predicting subjective experience of emotion: A collective level analysis. *Psicothema*, *12*, 55–69.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246.
- Boehnke, K., Silbereisen, R. K., Eisenberg, N., Reykowski, J., & Palmonari, A. (1989). Developmental pattern of prosocial motivation: A cross-national study. *Journal of Cross-Cultural Psychology*, *20*, 219–243.

- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.
- Bryant, B. K. (1982). An index of empathy for children and adolescents. *Child Development*, 53, 413–425.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456–466.
- Caprara, G. V., & Pastorelli, C. (1993). Early emotional instability, prosocial behavior, and aggression: Some methodological aspects. *European Journal of Personality*, 7, 19–36.
- Carlo, G. (2006). Care-based and altruistically based morality. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 551–579). Mahwah, NJ: Erlbaum.
- Carlo, G., Eisenberg, N., & Knight, G. P. (1992). An objective measure of prosocial moral reasoning. *Journal of Research on Adolescence*, 2(4), 331–349.
- Carlo, G., Koller, S. H., Eisenberg, N., DaSilva, M., & Frohlich, C. (1996). A cross-national study on the relations among prosocial moral reasoning, gender role orientations, and prosocial behaviors. *Developmental Psychology*, 32, 231–240.
- Carlo, G., McGinley, M., Roesch, S. C., & Kaminski, J. M. (2008). Measurement invariance in a measure of prosocial moral reasoning to use with adolescents from the USA and Brazil. *Journal of Moral Education*, 37, 485–502.
- Carlo, G., Mestre, M. V., Samper, P., Tur, A., & Armenta, B. E. (2010). Feelings or cognitions? Moral cognitions and emotions as longitudinal predictors of prosocial and aggressive behaviors. *Personality and Individual Differences*, 48, 872–877.
- Centro de Investigaciones Sociológicas (CIS). (2004). *Revista Española de Investigaciones Sociológicas*. Madrid: Ministerio de la Presidencia.
- Del Barrio, V., Moreno, C., & Lopez, R. (2001). Evaluación de la aggression e inestabilidad emocional en niños españoles y sur elación con la depression. *Clínica y Salud*, 13, 33–50.
- Eisenberg, N. (1986). *Altruistic emotion, cognition and behavior*. Hillsdale, NJ: Erlbaum.
- 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: Wiley.
- Eisenberg, N., Zhou, Q., & Koller, S. (2001). Brazilian adolescents' prosocial moral judgment and behavior: Relations to sympathy, perspective taking, gender-role orientation, and demographic characteristics. *Child Development*, 72, 518–534.
- Elzo, J. (2004). La educación del futuro y los valores. En *Debates de educación (2004: Barcelona)* [On-line article]. Fundación Jaume Bofill: UOC.
- Fernández-Berrocal, P., Salovey, P., Vera, A., Ramos, N., & Extremera, N. (2001). cultura, inteligencia emocional percibida y ajuste emocional: Un estudio preliminar. *Revista Electrónica de Motivación y Emoción*, 4, 8–9.
- Gibbs, J. C. (2003). *Moral development and reality: Beyond the theories of Kohlberg and Hoffman*. Thousand Oaks, CA: Sage.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. New York, NY: Cambridge University Press.
- Hofstede, G. H. (1984). *Culture's consequences: International differences in work related values*. Beverly Hills, CA: Sage.
- Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology*, 16, 131–152.
- Knight, G. P., & Hill, N. E. (1998). Measurement equivalence in research involving minority adolescents. In V. McLoyd & L. Steinberg (Eds.), *Study minority adolescents: Conceptual, methodological and theoretical issues* (pp. 183–210). Hillsdale, NJ: Erlbaum.

- Kumru, A., Carlo, G., Mestre, V., & Samper, P. (2012). Prosocial moral reasoning and prosocial behavior among Turkish and Spanish adolescents. *Social Behavior and Personality*, *40*(2), 205–214.
- Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development*, *71*(1), 107–118.
- Malti, T., Gummerum, M., & Keller, M. (2008). Moral emotions and moral cognitions. *European Journal of Developmental Science*, *2*(3), 203–205.
- Malti, T., & Keller, M. (2009). The relation of elementary school children's externalizing behaviour to emotion attributions, evaluation of consequences, and moral reasoning. *European Journal of Developmental Science*, *6*, 592–614.
- Malti, T., & Keller, M. (2010). The development of moral emotions in a cultural context. In W. F. Arsenio & E. A. Lemerise (Eds.), *Emotions, aggression, and morality in children: Bridging development and psychopathology* (pp. 177–198). Washington, DC: American Psychological Association.
- Mestre, V., Frías, M. D., Samper, P., & Tur, A. (2002). Adaptación y validación en población española del PROM: Una medida objetiva del razonamiento moral prosocial. *Acciónpsicológica*, *1*, 221–232.
- Mestre, V., Pérez-Delgado, E., Frías, M. D., & Samper, P. (1999). Instrumentos de evaluación de la empatía. In E. Pérez-Delgado & V. Mestre (Eds.), *Psicología moral y crecimiento personal* (pp. 181–190). Barcelona: Ariel.
- Millsap, R. E., & Kwok, O.-M. (2004). Evaluating the impact of partial factorial invariance on selection in two populations. *Psychological Methods*, *9*, 93–115.
- Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*, 3–72.
- Pereña, J., & Seisdedos, N. (1997). *ACS. Escalas de afrontamiento para adolescentes*. Madrid: TEA.
- Rest, J. R. (1979). *Development in judging moral issues*. Minneapolis, MN: University of Minnesota Press.
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock & A. Satorra (Eds.), *Innovations in multivariate statistical analyses: A festschrift for Heinz Neudecker* (pp. 233–247). London: Kluwer Academic.
- Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *Proceedings of the American Statistical Association of the Business and Economic Section* (pp. 308–313). Alexandria, VA: American Statistical Association.
- Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.
- Skoe, E. E., Hansen, K. L., Morch, W., Bakke, I., Hoffman, T., Larsen, B., & Aasheim, M. (1999). Care-based moral reasoning in Norwegian and Canadian early adolescents: A cross-national comparison. *Journal of Early Adolescence*, *19*, 280–291.
- Steiger, J. S. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*(2), 173–180.
- Tietjen, A. (1986). Prosocial moral reasoning among children and adults in a Papua New Guinea society. *Developmental Psychology*, *22*, 861–868.
- Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 7–35). Mahwah, NJ: Erlbaum.