2017-02-10

How do men and women help? Validation of a multidimensional measure of prosocial behavior

Matthew G. Nielson
Arizona State University

Laura Padilla-Walker
Brigham Young University - Provo

Erin K. Holmes
Brigham Young University - Provo, erin_holmes@byu.edu

CrossMark

How do men and women help? Validation of a multidimensional measure of prosocial behavior

Matthew G. Nielson [a, *], Laura Padilla-Walker [b], Erin K. Holmes [b]

[a] *Arizona State University, USA*
[b] *Brigham Young University, USA*

ARTICLE INFO

ABSTRACT

The current study sought to address gender differences in prosocial behavior by creating and validating a multidimensional measure of prosocial behavior that more fully captures the ways that men help others. The new measure is directed toward family, friend, and strangers, and has five factors: defending, emotional support, inclusion, physical helping, and sharing. In Study 1, CFA analyses performed on a sample of 463 emerging adults online (mean age 23.42) revealed good model fit and divergent validity for each of the five factors. Study 2 replicated the analyses on a sample of 453 urban adolescents in the Northwest (mean age 18.37). Results established that all factors had good model fit, construct validity, and convergent validity. The discussion focuses on implications of this measure for future prosocial research including an increased diversity in how people (particularly men) help others and developmental differences toward different targets of prosocial behavior.

© 2017 The Foundation for Professionals in Services for Adolescents. Published by Elsevier Ltd. All rights reserved.

In most Western cultures, men are expected to be emotionally reserved and to avoid appearing feminine, which often prohibits showing others that they care about them. One consequence of this culture is that men describe themselves as less helpful than women in most studies despite research that indicates similar levels between boys and girls in the first few years of life (Eisenberg, Spinrad, & Knafo-Noam, 2015). Social scientists refer to the ways that people help others as prosocial behavior, or voluntary behavior primarily aimed at benefitting another (Eisenberg et al., 2015). Commenting on this trend, Hastings, McShane, Parker, and Ladha (2007) noted that rather than preschool boys being less prosocial than girls, perhaps they are just prosocial in different ways. They proposed that masculine prosocial behavior might appear more active, agentic, and physical than the ways that girls help each other — more in-line with cultural masculine standards of behavior. Subsequent studies have generated masculine prosocial items in focus groups of middle schoolers (Bergin, Talley, & Hamer, 2003) and explored gender-typing in prosocial behavior (Hine & Leman, 2013), but these new insights have not yet been actualized in an improved measure of prosocial behavior. Are men naturally less prosocial than women, or are those findings based on bias in measurement? Can "masculine prosocial behavior" be measured by these instrumental items, or is masculine behavior structurally different from "feminine prosocial behavior"? Can existing items be rephrased in such a way that they reflect the pressures that men likely feel to avoid feminine presentation? The purpose of this study was to answer these questions by creating and validating a multidimensional measure of prosocial behavior suited to both men and women, to compare it with existing prosocial measures, and to build upon previous ideas about the ways men help others.

---

 * Corresponding author. 2011 N. 51st St. #4, Phoenix, AZ 85008, USA.
    E-mail addresses: matthew.nielson@asu.edu (M.G. Nielson), laura_walker@byu.edu (L. Padilla-Walker), erin_holmes@byu.edu (E.K. Holmes).

1. Theoretical foundation

Hegemonic masculinity, as proposed by Carrigan, Connell, and Lee (1985), states that masculine stereotypes of emotional stoicism, avoidance of anything feminine, and physical strength (among others) are intrinsically and extrinsically perpetuated in Western culture. Therefore, although boys and girls may have the same models for prosocial behavior, they are likely being intentionally instructed, and innately developing in intricately different gendered ways (Crouter, Whiteman, McHale, & Osgood, 2007; Hastings et al., 2007). The gender intensification theory explains an increase in the pressure that teens likely feel from hegemonic masculine culture as they proceed through adolescence. This theory posits that as boys and girls age, gender differences in psychological, behavioral, and attitudinal processes increase, and this increase is due to heightened social pressure to conform to established gender stereotypes as their bodies mature into adult forms (Hill & Lynch, 1983). Consequently, the differences in the gendered socialization that boys and girls receive continues to increase throughout adolescence (Hill & Lynch, 1983). Key to this paper, prosocial behavior is a social interaction that is likely affected by pressure to conform to cultural gender stereotypes. Eisenberg, Fabes, and Spinrad (2006) found that when girls from preschool to college-age were aware they were being assessed and they knew what they were being assessed on, they projected a socially desirable image of themselves. That is, the respondents knew that their society valued women who are sweet, helpful, and kind, so they modified their answers accordingly. Unsurprisingly, the same results were not found for male respondents because hegemonic masculine gender roles mandate that men be emotionally aloof rather than empathic (Jaffee & Hyde, 2000), thus men would not feel the same pressure to live up to expectations of sweetness and kindness that have been the focus of traditional prosocial measures. A review of the studies on how gender and masculinity affect prosocial behavior depicts how this theoretical foundation is realized in adolescent behavior.

Western masculine culture can be broken down into a set of rules that one must live by in order to access the power and privilege afforded men in patriarchal societies. Important rules in this code of behavior include not showing emotional sensitivity, being physically hard/strong, and displaying obvious heterosexuality (Pleck, 1983). Breaking these rules result in a loss of masculine capital. Individuals it his culture accordingly become gender police to ensure the worthiness of those who are to be called men, and those who fail to live up to the masculine standard are penalized in order to enforce adherence to masculine rules (Pascoe, 2007). The masculine tenet of avoiding emotional sensitivity is particularly important in the current study because the majority of the items comprising current prosocial measures rely on emotional sensitivities that prompt empathic response. Men wishing to strengthen their masculine image and avoid the punishments for breaking it are not likely to want to relate to "feminine" characteristics of concern and sympathy. Indeed, Hine and Leman (2013) found that adolescent boys and girls were particularly unified in defining masculine prosocial behaviors as direct or physical, feminine prosocial behaviors as involving emotion or relationships, and that peers of both gender discourage cross-gendered prosocial behaviors. Differential socialization continues throughout adolescence and emerging adulthood with Carlo and Randall (2002) indicating college-aged men more likely to participate in "heroic" types of prosocial behavior and Kimmel (2008) documenting the very strong pressures that men still face in college to uphold masculine ideals.

2. Gender differences in prosocial behavior

As the amount of prosocial research grows, an increasing number of articles highlight gender differences in the amount of prosocial behavior between men and women, though patterns and replicability are weak. Many studies show that women display more empathy (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992), more kindness (Hastings, Utendale, & Sullivan, 2007) and more desire to help others (Eisenberg et al., 2006), yet Eisenberg et al.'s (2015) most recent meta-analysis of prosocial behavior shows gender differences in prosocial behavior fluctuating over time and across prosocial constructs. Although women were higher on most indicators of prosocial behavior, these results were clearest when self-report measures were used (Chapman et al., 2006). These results are not surprising given the feminine stereotype of care and the masculine stereotype of stoicism. These deeply ingrained stereotypes are difficult to avoid and likely color the measurement taken by impartial observers; observers from within this culture tend to interpret participant actions according to their gendered expectations, with actions of women regarded as more prosocial than actions of men (Shigetomi, Hartmann, & Gelfand, 1981).

Indeed, the early years of life show no gender difference in natal levels of empathic response (Eisenberg et al., 2006). However, the development of children increasingly reflect the ideas of the culture that is socializing them. Gretarsson and Gelfand (1988) found that in the US people see women as intrinsically caring and prosocial, whereas men are prosocial as the result of careful socialization. Physiological data tells a different story: Michalska, Kinzler, and Decety (2013) found that although girls *reported* feeling more upset than boys when viewing others that were hurt, there were no gender differences in physiological response. In a meta-analysis of gendered prosocial behavior, Eagly (2009) claimed that it is incorrect to say women are more helpful. Rather, she found that prosocial behavior research yields patterns of gender specialization with women focused on emotionally supportive behavior and men focused on agentic, social, stranger-oriented behavior.

Several studies have questioned whether or not a potential bias is created when using primarily empathic/feminine items in traditional measurements of prosocial behavior (Bergin et al., 2003; Carlo & Randall, 2002; Zarbatany, Hartmann, Gelfand, & Vinciguerra, 1985). "Feminine" prosocial items measure caring behaviors like tending to another's needs (Skoe, Cumberland, Eisenberg, Hansen, & Perry, 2002), kindness (Söchting, Skoe, & Marcia, 1994), and empathic responses (Gretarsson & Gelfand, 1988); items from which many men are trying to disassociate. Eagly (2009) proposed that masculine

prosocial behavior might be more agentic, collectively oriented, and strength intensive, which are not often mentioned in current measurement tools of prosocial behavior (e.g., Carlo & Randall, 2002). Items like physical assistance (Zarbatany et al., 1985), aiding in an emergency (Eisenberg et al., 2015), and taking turns (Hastings et al., 2007) are all still valid prosocial behaviors and are reasonably more attractive to men as well as descriptive of their behavior.

Zarbatany et al. (1985) conducted a study in which fifth-graders identified the real-life prosocial actions of their peers and found that boys were more likely to nominate other boys for prosocial behaviors, and that "masculine-related items do elicit endorsements of boys for prosocial behavior" (Zarbatany et al., 1985, p. 97). Bergin et al. (2003) led a focus group of adolescents in a discussion of prosocial behaviors with the aim of identifying some masculine prosocial behavior. Standing up for others, including others, and coaching others to develop skills (i.e. giving pointers during games) were some of the original categories of masculine prosocial behavior they validated across different focus groups.

Accordingly, there are two forces working to distort our current view of the way men help others: first, that men are socialized to avoid performing any behavior that might be construed as feminine, and second, the condition that current measures only asses the kind of prosocial behaviors that men are likely intent on avoiding. We believe that men and women have the same empathic response capabilities (which are manifest in young children), but that men are socialized throughout their lives to display less of the kinds of empathic behaviors that are captured in current measures of prosocial behavior. An adjusted measure of prosocial behavior that reflects the actions that men in Western cultures are socialized to preform will yield more accurate portrayals of their prosocial levels.

3. Multidimensional prosocial research

While many of the conceptualizations and empirical explorations of prosocial behavior have treated this behavior as a general, unidimensional construct, more recent studies have begun to hone in on the multidimensional nature of this complex set of behaviors (Padilla-Walker & Carlo, 2014). This multidimensional approach to studying prosocial behavior has emphasized different types of prosocial behavior (Carlo & Randall, 2002), as well as how prosocial behavior differs as a function of the target of the behavior (e.g., strangers, friends, and family; Padilla-Walker & Christensen, 2011), the effect of relational and dispositional influences on prosocial behavior, and gender differences in prosocial behavior (Eisenberg et al., 2015). A key outcome of multidimensional prosocial research is that a person's willingness to help others is likely to be affected greatly by their relationship with the recipient. The same task that might be performed for a sibling or parent without a second thought might cause consternation if it were for a stranger (Eisenberg, 1983). As the case is strong to look at prosocial behavior toward each of these targets, this study utilized two separate models with prosocial behavior directed at three different targets: family, friend, and stranger.

Previous research indicates few gender differences in prosocial behavior toward family members during early adolescence (Padilla-Walker & Christensen, 2011), and we were interested to see whether this trend continues through emerging adulthood. Additionally, men have been shown to prefer "heroic" actions, prosocial behavior in a public setting (Carlo & Randall, 2002), and are less motivated to serve others in order to maintain relationships (Eisenberg et al., 2015). Given this and the potential for physical danger, we expected that men might display higher levels of prosocial behavior toward strangers than women.

In addition to the variance associated with different targets, different types of prosocial behavior also elicit different patterns. For example, Caprara and colleague's (2005) scale assesses four types of prosocial responding, and the Prosocial Tendencies Measure by Carlo and Randall (2002) focused on six different situational contexts (e.g., emotional, dire, anonymous), along with their unique outcomes. This measure seeks to build on these previous measures in an attempt to consider type *and* target, as well as specifically focusing on reducing gender differences that might be overestimated due to self-reported biases. In line with the theoretical basis of hegemonic masculinity, social learning, and gender intensification, the types of prosocial behavior included in the current study move beyond emotional behaviors to describe behaviors that might be more native to the ways that men think about and perform prosocial actions. Building off the categories of behavior given as typical to the ways men help others in Bergin et al. (2003), the measure included in this study contained five potential types of prosocial behavior that might be more common among men including defending behavior, a reworded version of emotional support, inclusion, physical helping, and sharing.

The defending behavior items were generated by the participants in Bergin and colleague's (2003) focus group and introduce novel ways of measuring prosocial behavior that capture the protective ways boys might act to help others. The emotional support items are similar to items found in Peterson and Seligman's prosocial scale (2004), but adjustments were made such as "I try to cheer up people who seem sad" to "If someone is sad, I try to make that person laugh" in order to fit more closely the ways boys are socialized to interact with others. Inclusion items have been utilized in previous prosocial work (Hine & Leman, 2013), but are not common and generally never more than one or two items. The physical helping items capitalize on Hastings and colleague's (2007) suggestions that the prosocial behavior of boys might be more active and agentic than are currently captured. This type of behavior includes items such as "I do physical acts of service for others [lifting heavy things, yard work, cleaning]". Finally, much past prosocial research has focused exclusively on sharing behaviors, and the items we included are found in many prosocial measures. We have few studies on which to base specific hypotheses regarding how these five different types of prosocial behavior might vary as a function of gender and age toward the three different targets, but masculinity scholars such as Pascoe (2007) and Pleck (1983) would not be surprised to find the largest gender differences in prosocial behavior toward friends as boys' peers might be their most stringent gender police.

4. Current study

The focus of the current study was to analyze existing literature, evaluate the need to more accurately identify masculine prosocial behavior, develop and validate a multidimensional prosocial measure, and compare it to an existing framework of prosocial behavior and its corollary behaviors. Two studies were carried out with different samples in order to examine the measure across various ages and types of individuals. The sample for the first study was comprised of emerging adults recruited online from Amazon Mechanical Turk (Amazon Turk), while the second study was conducted with adolescents from the Flourishing Families Project. In Study 1, a confirmatory factor analysis (CFA) was conducted on the five factor, 20-item measure created for friend and stranger oriented prosocial behavior. Study 2 repeated the CFA on a different sample, compared mean differences between men and women's prosocial behavior on the new measure, and examined patterns of relation with an extant measure and cross-sectional correlates of prosocial behavior.

4.1. Study 1

A multidimensional prosocial measure composed of items that both men and women relate to is likely to give us a more accurate understanding of how men perform prosocial behavior. The current study sought to test this theory by assessing various types of prosocial behaviors that are not stereotypically feminine, as well as by re-wording traditional prosocial items in a more gender-neutral style. According to the Hegemonic Masculine theory, we expected that the pressures men face to be masculine would produce more active prosocial behaviors that may be different from the limited range of empathic prosocial behaviors represented in extant prosocial measures. This study seeks to validate a new measure of prosocial behavior that includes new categories such as defending behavior, emotional support, inclusion, physical helping, and sharing behaviors.

5. Method

5.1. Participants and procedures

Participants included 463 adolescents and emerging adults from the Amazon Mechanical Turk online workforce database. Amazon Turk is a global crowdsourcing internet marketplace through which various tasks can be performed, including social science research. General Amazon Turk demographics show that most of the workers live in the US (57%), that they are between the ages of 18 and 24, most hold a bachelor's degree, and earn less than $10,000 annually (Goodman, Cryder, & Cheema, 2013). Additional studies have validated the characteristics of Amazon Turk workers compared to more traditional samples (Holden, Dennie, & Hicks, 2013), and replicated extant behavioral science research such as outcomes on the Big Five Personality questionnaire (Suri & Watts, 2011). The participants from Amazon Turk in this sample were from the US (ages 18–25, M_{age} = 23.42, SD = 1.79), 60% male, 69% Caucasian, 11% black, 9% Asian, 6% Latino, and 4% other/mixed race. Forty-eight percent were currently employed, and 75% lived on their own rather than with their parents.

5.2. Item selection

The first step in constructing a multidimensional measure of prosocial behavior was to collect potentially gender-neutral items from the prosocial literature body. A general search yielded a few important studies that aided in the development of this measure, particularly Bergin et al. (2003), a focus group conducted with students to generate different ways that boys help others. From this study, we borrowed concepts from several of the types of behavior they generated including stands up for others, peacemaker, humorous, emotionally supportive, helps others develop skills, inclusive, provides physical assistance, and shares. Further items were generated based on theoretical directions from Hastings et al. (2007), Eagly (2009), Carlo, McGinley, Hayes, Batenhorst, and Wilkinson (2007), and Hine and Leman (2013) including: "I do physical acts of service for others [lifting heavy things, yard work, cleaning]," "I share my personal belongings with people," and "If I notice someone who is lonely, I try to include that person."

This process of item collection and creation resulted in a measure comprised of 68 items measuring various types of prosocial behavior. With parsimony in mind, we consolidated similar items (i.e. standing up for others and physical helping behaviors) which removed 10 items. Next, we rejected behavior that was simply avoiding negative behavior (avoids fights, honesty, humorous, keeps confidences, does not make fun of others, expresses happiness) which narrowed the list down to 37 items.

Following the target-specific model of Padilla-Walker and Christensen (2011), the same items were used for each target, but unique stems were created to direct participants to answer about only one particular target (i.e. "The following questions will ask you about your behavior with strangers"). In this framework, items that are specific to one target, such as community service, could not be adapted toward the three different targets of our study, thus they were removed. After these theoretically based reductions, we ended up with five types of behavior (defending behavior, emotional support, inclusion, physical help, and sharing) and a total of 20 items.

6. Measures

6.1. Prosocial behavior

We expected the new measure to be composed of five factors of four items each: defending behavior (i.e. "If someone is being made fun of, I stick up for that person"), emotional support (i.e. "If someone is upset, I listen to that person"), inclusion (i.e. "If someone is new to a group, I make an effort to include that person"), physical help (i.e. "I do physical acts of service for others [lifting heavy things, yard work, cleaning]"), and sharing (i.e. "I share my personal belongings with people"). Items were measured on a 5-point Likert scale ranging from 1 (not like me at all) to 5 (very much like me) (see Appendix A for the full measure).

7. Data management

Maximum Likelihood (MLR) feature of Mplus 7 (Muthén & Muthén, 1998–2015) was used to handle the less than 6% missing data for each variable. Logistic regression analyses were conducted on the cases with missing data and SES variables including income, ethnicity, and gender. The results revealed a non-significant association for all three tests (Income: β $(1) = 0.01$, $p = 0.067$; Ethnicity: β $(1) = 0.43$, $p = 0.057$; Gender: β $(1) = 0.16$, $p = 0.097$). Descriptive analyses conducted in SPSS 21 include an estimation of means, standard deviations, and bivariate correlations for each factor from the three different target models (see Table 1).

8. Planned analyses

8.1. Step one

Content validity of the measure produced in the literature review was tested using a confirmatory factor analysis (CFA) conducted on the Amazon Turk sample using structural equation modeling with Mplus. Because of the highly correlated nature of the prosocial factors, the basic models were compared with two hierarchical models: a second-order latent variable model (in which an overarching "prosocial behavior" factor was composed of the five of prosocial factors), and a bi-factor model (in which observed variables compose a first-order global prosocial factor along with the five first-order factors of the types of prosocial behavior) (Rindskopf & Rose, 1988). As these models are not nested, comparisons were conducted based on the Bayesian Information Criteria (BIC) (with lower scores indicating better model fit), and on factor loadings for the first and second latent factors. For the final model, acceptable fit was assessed by observing the chi-square significance test (significant outcomes should be noted, but allowed to continue), a comparative fit index (CFI) score above 0.90, a Tucker Lewis Index (TLI) above 0.90, and a root mean square error of approximation (RMSEA) score below 0.08 (Little, 2013). After the model fit was deemed acceptable, the factor analysis was performed, and factor loadings were assessed. Factor loadings were deemed acceptable if $\beta > 0.50$ (Little, 2013).

Table 1
Descriptive statistics and correlations.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	M	SD
Family																	
1. Def		0.83*	0.85*	0.78*	0.65*	0.62*	0.56*	0.58*	0.56*	0.43*	0.27*	0.25*	0.40*	0.40*	0.13*	4.28	0.83
2. Emo	0.74*		0.88*	80*	0.68*	0.52*	0.64*	0.60*	0.55*	0.41*	0.19*	0.27*	0.43*	0.38*	0.10*	4.35	0.82
3. Inc	0.72*	0.74*		0.84*	0.70*	0.55*	0.60*	0.66*	0.57*	0.42*	0.23*	0.29*	0.48*	0.41*	0.13*	4.31	0.77
4. Phys	0.66*	0.61*	0.62*		0.73*	0.51*	0.56*	0.58*	0.63*	0.40*	0.17*	0.20*	0.40*	0.41*	0.05	4.46	0.76
5. Shar	0.55*	0.53*	0.57*	0.56*		0.38*	0.43*	0.46*	0.49*	0.57*	0.15*	0.18*	0.32*	0.29*	0.21*	4.24	0.88
Friend																	
6. Def	0.54*	0.45*	0.48*	0.58*	0.40*		0.78*	0.77*	0.77*	0.55*	0.48*	0.39*	0.54*	0.56*	0.20*	4.15	0.77
7. Emo	0.45*	0.55*	0.46*	0.53*	0.35*	0.64*		0.80*	0.73*	0.54*	0.30*	0.41*	0.58*	0.46*	0.13*	4.33	0.72
8. Inc	0.42*	0.44*	0.58*	0.43*	0.38*	0.66*	0.64*		0.73*	0.51*	0.35*	0.39*	0.66*	0.50*	0.18*	4.21	0.72
9. Phys	0.49*	0.43*	0.44*	0.72*	0.41*	0.70*	0.63*	0.62*		0.57*	0.33*	0.30*	0.49*	0.60*	0.15*	4.24	0.73
10. Shar	0.38*	0.38*	0.36*	0.42*	0.58*	0.49*	0.53*	0.46*	0.56*		0.32*	0.33	0.38*	0.39*	0.45*	3.90	0.89
Stranger																	
11. Def	0.46*	0.39*	0.39*	0.34*	0.28*	0.58*	0.37*	0.47*	0.38*	0.26*		0.74*	0.55*	0.64*	0.62*	3.18	0.99
12. Emo	0.40*	0.49*	0.37*	0.29*	0.29*	0.41*	0.57*	0.44*	0.38*	0.33*	0.63*		0.64*	0.61*	0.63*	3.35	0.99
13. Inc	0.36*	0.39*	0.52*	0.35*	0.31*	0.52*	0.48*	0.75*	0.48*	0.33*	0.57*	0.56*		0.64*	0.41*	3.80	0.74
14. Phys	0.40*	0.35*	0.36*	0.55*	0.32*	0.54*	0.45*	0.46*	0.72*	0.36*	0.56*	0.55*	0.59*		0.51*	3.68	0.83
15. Shar	0.29*	0.29*	0.26*	0.22*	0.41*	0.26*	0.29*	0.28*	0.31	0.58*	0.44*	0.56*	0.39*	0.47*		2.85	0.99
Mean	4.07	4.21	4.20	4.40	3.97	0.403	4.37	4.15	4.20	3.91	3.46	3.95	3.96	3.95	3.48		
SD	0.82	0.80	0.72	0.68	0.91	0.74	0.67	0.67	0.71	0.83	0.82	0.81	0.71	0.71	0.95		

Notes: Amazon Turk scores on top half, Flourishing Family scores on bottom half, * indicates $p < 0.05$, Def = Defending behavior, Emo = Emotional support, Inc = Inclusion, Phys = Physical helping, Shar = Sharing.

8.2. Step two

Models were tested for invariance across gender using Little's (2013) Fixed Factor method of invariance. In this method, nested models of configural, weak, and strong gender invariant models are compared successively. Due to the hierarchical nature of the model, the first order latent variables were tested for invariance followed by the second order variable (Rindskopf & Rose, 1988). Model fit was determined by log likelihood values obtained with the MLR estimator which were then converted to a chi-square difference value using the Satorra-Bentler Scaled Chi-Square adjustment. Satorra-Bentler Scaled Chi-Square values were used rather than the traditional chi-square difference tests because when MLR is used, traditional chi-square difference tests become inapplicable (Satorra & Bentler, 2010). The Satorra-Bentler adjustment produces a chi-square value that can be interpreted as a regular chi-square value, thus acceptable fit was determined in the comparison model by a non-significant change in the chi-square value from the nested model.

Configural models include the same items held constant across constructs. If configural invariance was achieved, the model underwent weak invariance testing in which the configural model was compared to a model where factor loadings were constrained to be equal for each group. If weak invariance was achieved, the model was compared to a model in which factor intercepts and loadings were constrained to be equal across groups. If strong variance was achieved, the model was considered invariant across constructs and mean levels could be compared across those constructs.

The second-order latent variables were tested on two levels: unconstrained and constrained (Chen, Sousa, & West, 2005). At the unconstrained level, the strong constraints from the first-order level variables were maintained, but second order level factor loadings were free to vary. If adequate model fit was achieved, the model was compared to a model in which second-order level factor loadings were constrained to be equal across gender. If the constrained level of variance was achieved, mean levels could be compared across constructs at the second-order level.

8.3. Step three

Tests of divergent validity were conducted on the five-factor measure by constraining estimated correlation parameters between factors to 1.0 and assessing $\chi^2\Delta$ (Anderson & Gerbing, 1988).

8.4. Step four

After establishing gender invariance, mean level differences between men and women were analyzed by interpreting Mplus betas and significance levels for mean level differences between the male and female models. In the fixed factor method, female factor means are constrained at 0 whereas male factor means are allowed to vary; thus significant male scores indicate by how much the male means differ from zero (Little, 2013).

9. Results

9.1. Step one: CFAs

9.1.1. Family

Using Mplus 7 software (Muthén & Muthén, 1998–2015), the five factor model for family-oriented prosocial behavior was analyzed on the Amazon Turk sample and revealed an acceptable fit to the data after adding inter-item correlations according to modification indexes suggested by Mplus: $\chi^2(344) = 584.987$, $p < 0.001$, CFI = 0.958, RMSEA (90% CI) = 0.047 [0.040, 0.053]. When hierarchical models were compared, the second-order model proved to have the best BIC (original model BIC = 21890.445, second-order BIC = 21782.101, and bi-factor BIC = 21783.416), and achieved adequate model fit ($\chi^2(356) = 590.866$, $p < 0.001$, CFI = 0.958, RMSEA (90% CI) = 0.047 (0.042, 0.055). The factor structure included a second-order latent variable model with prosocial behavior composed of the five factors with items that loaded above $\beta > 0.50$. Modification indices were examined and there was no evidence of cross-loading. This same procedure was followed for all subsequent CFA analyses.

9.1.2. Friend

The five factor model for friend-oriented prosocial behavior revealed an acceptable fit to the data after adding inter-item correlations according to modification indexes suggested by Mplus: $\chi^2(342) = 608.071$, $p < 0.001$, CFI = 0.948, RMSEA (90% CI) = 0.049 [0.043, 0.055]. When hierarchical models were compared, the second-order model proved to have the best BIC (original model BIC = 26020.350, second-order BIC = 9109.969, and the bi-factor model would not converge), and achieved adequate model fit ($\chi^2(356) = 626.818$, $p < 0.001$, CFI = 0.947, RMSEA = 0.049 (0.042, 0.055).

9.1.3. Stranger

The five factor model for friend-oriented prosocial behavior revealed an acceptable fit to the data after adding inter-item correlations according to modification indexes suggested by Mplus: $\chi^2(338) = 703.967$, $p < 0.001$, CFI = 0.945, RMSEA (90% CI) = 0.057 [0.051, 0.063]. When hierarchical models were compared for stranger-oriented behavior, the original model had the best BIC (original model BIC = 32213.603, second-order BIC = 32232.328, and bi-factor would not converge), but the

factor correlations were very high ($\beta > 0.90$, $p < 0.05$), thus the second-order model was chosen ($\chi^2(352) = 788.445$, $p < 0.001$, CFI = 0.934, RMSEA = 0.061(0.056, 0.067)).

9.2. Step two: invariance testing across gender

Multiple group analyses were conducted to determine whether invariance could be established across gender on both factor loadings (weak invariance) and intercepts (strong invariance), which is what is needed to compare means across groups (Little, 2013). First-order and second-order latent variable models were analyzed for invariance sequentially (Rindskopf & Rose, 1988). Using the Satorra-Bentler Chi-Square correction factor, all models tested were revealed to be strongly invariant by gender, thus enabling mean level comparisons across male and female models (see Table 2).

9.3. Step three: divergent validity

Constraining all five factors for each of the three models (one at a time) to 1.0 each resulted in a decrease in model fit, suggesting that the five factors were distinct (no constraint resulted in less than $\chi^2\Delta = 186.814$, $p < 0.001$). Cronbach's alphas were calculated, resulting in additional evidence of adequate internal reliability for each 4-item scale: defending behavior (α's ≥ 0.85), emotional support (α's ≥ 0.88), inclusion (α's ≥ 0.75), physical help (α's ≥ 0.83), and sharing (α's ≥ 0.82).

9.4. Step four: mean level differences between gender

9.4.1. Family
To determine mean level differences between men and women, Betas from the strong invariant male model were interpreted. Results indicated that women were significantly higher than men for all five factors of prosocial behavior (defending behavior $\beta = -0.28$, $p < 0.001$; emotional helping $\beta = -0.35$, $p < 0.001$; inclusion $\beta = -0.26$, $p < 0.001$; physical helping $\beta = -0.24$, $p < 0.05$; and sharing $\beta = -0.29$, $p < 0.05$).

9.4.2. Friend
For the friend-oriented model, results indicated that women were significantly higher than men for three factors of prosocial behavior including emotional support ($\beta = -0.45$, $p < 0.001$), inclusion ($\beta = -0.29$, $p < 0.05$), and sharing ($\beta = -0.29$, $p < 0.05$). There was no significant difference between men and women for defending behavior ($\beta = -0.22$, $p = 0.085$) and physical helping ($\beta = 0.04$, $p = 0.752$).

9.4.3. Stranger
For the stranger-oriented model, results indicated that women were only significantly higher on inclusion ($\beta = -0.24$, $p < 0.05$) and that all other factors had no significant gender difference (defending behavior $\beta = -0.06$, $p = 0.591$; emotional support $\beta = -0.19$, $p = 0.078$; physical helping $\beta = 0.19$, $p = 0.128$; and sharing $\beta = 0.04$, $p = 0.74$).

Table 2
Model fit statistics for measurement invariance (Amazon Turk sample).

Model tested	χ^2	df	p	CFI	CFIΔ	RMSEA (90% CI)	SRMR
1st Order Family:							
Configural	585.809	334	<0.001	0.957	–	0.049 (0.042, 0.055)	0.059
Factor Loadings	599.010	349	<0.001	0.957	0.00	0.048 (0.041, 0.054)	0.063
Intercepts	584.987	344	<0.001	0.958	0.001	0.047 (0.040, 0.053)	0.047
2nd Order Family							
Unconstrained	641.866	358	<0.001	0.951	–	0.050 (0.044, 0.056)	0.050
Constrained	643.733	362	<0.001	0.951	0.00	0.050 (0.043, 0.056)	0.055
1st Order Friend:							
Configural	615.477	332	<0.001	0.945	–	0.051 (0.045, 0.058)	0.068
Factor loadings	630.254	347	<0.001	0.945	0.00	0.050 (0.044, 0.056)	0.070
Intercepts	608.071	342	<0.001	0.948	0.003	0.049 (0.043, 0.055)	0.056
2nd Order Friend:							
Unconstrained	626.817	356	<0.001	0.947	–	0.049 (0.042, 0.055)	0.057
Constrained	629.397	360	<0.001	0.948	0.001	0.048 (0.042, 0.054)	0.060
1st Order Stranger:							
Configural	715.422	328	<0.001	0.941	–	0.060 (0.054, 0.066)	0.059
Factor loadings	734.884	343	<0.001	0.941	0.00	0.059 (0.053, 0.065)	0.061
Intercepts	703.967	338	<0.001	0.945	0.004	0.057 (0.051, 0.063)	0.056
2nd Order Stranger:							
Unconstrained	886.261	356	<0.001	0.920	–	0.067 (0.062, 0.073)	0.072
Constrained	888.206	360	<0.001	0.920	0.00	0.067(0.061, 0.072)	0.074

10. Study 1 discussion

The CFAs indicated appropriate model fit, and mean differences between women and men suggested that including a broader range of activities that extended beyond empathic response was more likely to capture the ways that men are prosocial. Altering existing prosocial questions to reflect hegemonic masculine pressure, as well as creating new "masculine" prosocial questions, led to higher levels of male prosocial behavior. Although men still seem more prone to physical helping than emotional support, given the pressures that men face to be emotionally stoic (Hill & Lynch, 1983; Hine & Leman, 2013), these outcomes were not unexpected. However, the present research shows less gender differences for emotional support after altering items to avoid potential language biases. By adding items like "I would enter a risky situation in order to help someone", and rephrasing items from "I always listen to my friends talk about their problems" to "If my friend is upset, I listen to them", we expected to tap into a desire to feel/appear masculine that is strongly manifest in men (Hill & Lynch, 1983). As noted in the introduction, prosocial behavior levels and motivations change with developmental stages, and we were interested to see how the trends produced with this measure would compare with a younger sample. Hine and Leman (2013) found that the pressure for boys to avoid feminine prosocial behaviors was still intense in their early teen years, but how might these pressures change near the end of adolescence?

11. Study 2

The purpose of Study 2 is to re-test the measures of multidimensional prosocial behavior from Study 1 on a different, younger sample. Along with conducting CFAs on the Amazon Turk sample, an important part of measure validation is convergent validity − testing how the new measure relates with theoretical and established correlational behaviors of the construct. Personality characteristics like sympathy form complicated interactions that impact how and when a person is prosocial. For example, an individual's ability and inclination to help others depends on the native levels of sympathy they feel (Davis, 1983). For this measure, we expected that sympathy would be positively correlated with prosocial behavior toward all targets. Similarly, it has been theorized that self-regulation, an individual's ability to control emotional and behavioral outcomes, is particularly associated with stranger-oriented prosocial behavior because the process of helping an unknown person might require a little more self-control than helping family or friends (Padilla-Walker & Christensen, 2011).

In accordance with the relational grounding of this study by delineating prosocial behavior toward different targets, a measure of friendship connection is an important correlate of prosocial behavior toward friends. Friendship connection is understood to be a central indicator of friendship quality, and previous research indicates that children displayed more prosocial behavior toward their friends than toward acquaintances. Padilla-Walker, Dyer, Yorgason, Fraser, and Coyne (2015) found that this pattern strengthens throughout adolescence thus we would therefore expect these patterns to emerge in our own work. Accordingly, measures of friendship connection are included in the influential correlates to prosocial behavior in this study.

The relation between prosocial behavior toward family and parental connection is similar to the relation between prosocial behavior toward friends and friendship connection. Closeness with mothers and fathers has been found to be particularly strongly associated with adolescent prosocial behavior (Eberly & Montemayor, 1999; Padilla-Walker et al., 2015). The effects of parental connection are not limited to prosocial behavior to the family; research indicates that positive parental connection positively affects prosocial behavior toward all targets (Carlo et al., 2007; Eisenberg et al., 2015). Accordingly, measures of friendship connection and parenting connection are included in the influential correlates to prosocial behavior in this study.

12. Method

12.1. Participants and procedures

Participants were 453 late adolescents who participated in the Flourishing Families Project, a longitudinal study of family life, now entering its eleventh wave. Data for this study were taken from Wave 8 (ages 16−21, M_{age} Wave 8 = 18.37, SD = 1.04) gathered during the summer months of 2014. The participants in this sample were 51% female, 67% European American (12% black, 4% Asian, 1% Latino, 15% other/mixed race), and 30% single parent families. 43% of the sample reported living at home. Average monthly income as reported by mothers was $5800 (approximately $70,000 per year), but approximately 28% of the sample reported a family income below $40,000 per year.

At the initial wave, families were randomly selected from targeted census tracts that mirrored local school districts using a purchased national telephone survey database. In an attempt to more closely mirror the demographics of the local areas, a limited number of families were recruited into the study through other means (e.g., referral, fliers; n = 77, 15%). Of the eligible families contacted, those agreeing to participate (N = 500) resulted in a 61% response rate. Of the families who participated at Wave 1, 90% had complete data at Wave 8. For wave 8 of data collection, questionnaires were administered via an online data collection service, Qualtrics, to enable data collection for participants who had graduated high school and no longer lived at home.

12.2. Data management

At Wave 8, data were gathered utilizing a planned missing design, consequently, the data for each of these items has been imputed, and the only missing data (approximately 10 percent due to longitudinal attrition) was handled using Maximum Likelihood. Logistic regression analyses were conducted on the cases with missing data at Time 8 and SES variables including income, ethnicity, and gender. The results revealed a non-significant association for all three tests (Income: β (1) = 0.00, $p = 0.089$; Ethnicity: β (1) = 0.59, $p = 0.051$; Gender: β (1) = −0.015, $p = 0.93$). Descriptive analyses conducted in SPSS 21 include an estimation of means, standard deviations, and bivariate correlations for each factor from the three different target models (see Table 1).

12.3. Measures

12.3.1. Sympathy

Adolescents reported on their sympathy using a 7-item measure (Davis, 1983; e.g., "When I see someone being taken advantage of, I feel kind of protective towards them"). Based on a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) with higher scores were indicative of greater sympathy ($\alpha = 0.77$).

12.3.2. Self-regulation

Adolescents reported on their self-regulation using a revised 13-item version of the Novak and Clayton (2001) self-regulation measure (e.g., "I have a hard time controlling my temper"). Items were reverse scored as necessary. Responses ranged from 1 (*never true*) to 5 (*always true*) with higher scores were indicative of greater self-regulation ($\alpha = 0.78$).

12.3.3. Parenting

Adolescents reported on their parent's parenting style using the Parenting Styles and Dimensions Questionnaire-Short Version (PSDQ; Robinson, Mandleco, Olsen, & Hart, 2001). Using a revised 15-item scale, adolescents were asked how often both their mothers and their fathers did certain behaviors characteristic of authoritative parenting which included subscales for warmth and support (5 items; e.g., "My mom/dad is responsive to my feelings and needs"), reasoning/induction (5 items; e.g. "My parent emphasizes the reasons for rules"), and autonomy granting (5 items; e.g. "My parent allows me to give input into family rules"). Responses ranged on a five point Likert-type scale from 1 (*never*) to 5 (*always*), with higher scores indicating authoritative parenting (mothers: $\alpha = 0.89$; fathers: $\alpha = 0.89$).

12.3.4. Friendship

Adolescents' relationship quality with a best friend was assessed using a 9-item short form of the 14 item scale from Barber and Olsen (1997). The scale assessed the overall relationship, utilizing questions about friend connection (7 items; e.g. "If you needed help with something, how often could count on this friend to help you?"), friend companionship (3 items; e.g. "How often do you and this friend go places together, like a movie, shopping or a sports event?"), and friend psychological control (4 items; e.g. "When this friend disagrees with you, how often does he/she make you feel like your ideas aren't as good as his or hers?") that were reverse coded. Adolescents responded on a 5-point scale ranging from 0 (*never*) to 4 (*every day*) ($\alpha = 0.73$) with higher scores indicative of better friendship quality.

12.3.5. Global prosocial measure

Adolescents' global prosocial behavior was measured using a modified version of the Kindness and Generosity subscale of the Values in Action Inventory of Strengths (Peterson & Seligman, 2004). The original measure was designed to assess behaviors toward strangers, but for this study it was modified to capture different prosocial targets including friends and family. The scales assess prosocial behavior toward friends (5 items e.g., "I go out of my way to cheer up my friends") and strangers (5 items e.g., "I go out of my way to cheer up people I do not know") on a 5-point Likert scale ranging from 1 (*not like me at all*) to 5 (*very much like me*). Mean scales were created for each prosocial measure in order to test correlations with traditional correlates of prosocial behavior (toward family: $\alpha = 0.91$; toward friends: $\alpha = 0.87$; toward strangers: $\alpha = 0.80$).

12.4. Planned analyses

12.4.1. Step one

Content validity of the prosocial measure produced in the literature review was on the Flourishing Families Project sample using the same CFA procedure detailed in Study 1.

12.4.2. Step two

Models were tested for invariance across gender using Little's (2013) Fixed Factor method of invariance.

12.4.3. Step three

Mean level differences between genders were tested using the same Beta difference method as Study 1.

12.4.4. Step four

Construct validity was determined by using SPSS 21 to correlate our multidimensional prosocial measure with an extant global measure of prosocial behavior, the Values in Action Inventory of Strengths (Peterson & Seligman, 2004). DeVellis (2012) wrote that an acceptable correlation level between variables to establish construct validity is any level above and beyond what shared variance can be attributed to chance. All five factors of our prosocial measure were correlated independently with the extant prosocial measure to determine if the new items were conceptually similar to established constructs of prosocial behavior.

12.4.5. Step five

Convergent validity was tested by correlating our multidimensional prosocial measure with cross-sectional measures of sympathy, self-regulation, authoritative parenting, and friendship quality.

12.4.6. Step six

Developmental gender differences were assessed by comparing the factor structure and mean differences for the models produced in the sample from Study 1 with the models produced in the sample from Study 2.

13. Results

13.1. Step one: CFAs

13.1.1. Family

The five factor model for family-oriented prosocial behavior revealed an acceptable fit to the data after adding inter-item correlations according to modification indexes suggested by Mplus: $\chi^2(342) = 595.415$, $p < 0.001$, CFI = 0.938, RMSEA (90% CI) = 0.059 [0.051, 0.067]. When hierarchical models were compared, the second-order model proved to have the best BIC (original model BIC = 17662.745, second-order BIC = 17623.059, and the bi-factor model failed to converge), and achieved adequate model fit ($\chi^2(356) = 630.162$, $p < 0.001$, CFI = 0.933, RMSEA = 0.060 (0.052, 0.068). The factor structure included a second-order latent variable model with prosocial behavior composed of the five factors with items that loaded above $\beta > 0.50$. Modification indices were examined and there was no evidence of cross-loading. This same procedure was followed for all subsequent CFA analyses.

13.1.2. Friend

The five factor model for family-oriented prosocial behavior revealed an acceptable fit to the data: $\chi^2(340) = 583.358$, $p < 0.001$, CFI = 0.934, RMSEA (90% CI) = 0.058 [0.050, 0.065]. When hierarchical models were compared, the second-order model proved to have the best BIC (original model BIC = 18102.811, second-order BIC = 18049.631, and bi-factor BIC = 18108.268), and achieved adequate model fit ($\chi^2(354) = 610.360$, $p < 0.001$, CFI = 0.931, RMSEA = 0.058 (0.050, 0.066).

13.1.3. Stranger

The five factor model for family-oriented prosocial behavior revealed an acceptable fit to the data: $\chi^2(340) = 630.151$, $p < 0.001$, CFI = 0.921, RMSEA (90% CI) = 0.063 [0.055, 0.070]. When hierarchical models were compared for stranger-oriented behavior, the hierarchical model had the best BIC (original model BIC = 20782.181, second-order BIC = 20782.181, and bi-factor model BIC = 20826.611) and acceptable model fit indexes ($\chi^2(354) = 684.261$, $p < 0.001$, CFI = 0.912, RMSEA = 0.066 (0.058, 0.073).

13.2. Step two: invariance testing across gender

Multiple group analyses were conducted to determine whether invariance could be established across gender. Using the Satorra-Bentler Chi-Square correction factor, all models tested were revealed to be strongly invariant by gender, thus enabling mean level comparisons across male and female models (see Table 3).

13.3. Step three: mean level differences between gender

13.3.1. Family

To determine mean level differences between men and women, Betas from the strong invariant male model were interpreted. Results indicated that women were only significantly higher than men on the emotional support factor ($\beta = -0.31$, $p < 0.001$). The other four factors indicated non-significant differences (defending behavior $\beta = -0.13$, $p = 0.183$; inclusion $\beta = -0.18$, $p = 0.060$; physical helping $\beta = -0.006$, $p = 0.957$; and sharing $\beta = -0.19$, $p = 0.067$).

13.3.2. Friend

To determine mean level differences between men and women, Betas from the strong invariant male model were interpreted. Results indicated that women were significantly higher than men for all five factors of prosocial behavior

Table 3
Model fit statistics for measurement invariance (Flourishing families sample).

Model tested	χ^2	df	p	CFI	CFIΔ	RMSEA (90% CI)	SRMR
1st Order Family:							
Configural	598.173	332	<0.001	0.935	–	0.061 (0.053, 0.069)	0.058
Factor loadings	611.196	347	<0.001	0.935	0.00	0.060 (0.052, 0.067)	0.066
Intercepts	595.414	342	<0.001	0.938	0.003	0.059 (0.051, 0.067)	0.060
2nd Order Family:							
Unconstrained	630.162	356	<0.001	0.933	–	0.060 (0.052, 0.068)	0.065
Constrained	634.210	360	<0.001	0.933	0.00	0.060 (0.052, 0.067)	0.072
1st Order Friend:							
Configural	592.021	330	<0.001	0.929	–	0.061 (0.053, 0.069)	0.057
Factor loadings	607.013	345	<0.001	0.929	0.00	0.059 (0.052, 0.067)	0.070
Intercepts	583.358	340	<0.001	0.934	0.005	0.058 (0.050, 0.065)	0.064
2nd Order Friend:							
Unconstrained	631.463	354	<0.001	0.925	–	0.060 (0.053, 0.068)	0.071
Constrained	635.765	358	<0.001	0.925	0.00	0.060 (0.052, 0.068)	0.076
1st Order Stranger:							
Configural	623.548	330	<0.001	0.922	–	0.064 (0.054, 0.072)	0.059
Factor loadings	651.684	345	<0.001	0.918	−0.004	0.064 (0.056, 0.071)	0.070
Intercepts	630.151	340	<0.001	0.923	0.004	0.063 (0.055, 0.070)	0.067
2nd Order Stranger:							
Unconstrained	684.261	354	<0.001	0.912	–	0.066 (0.058, 0.073)	0.072
Constrained	690.167	358	<0.001	0.912	0.00	0.065(0.058, 0.073)	0.080

(defending behavior $\beta = -0.32$, $p < 0.001$; emotional helping $\beta = -0.40$, $p < 0.001$; inclusion $\beta = -0.31$, $p < 0.001$; physical helping $\beta = -0.19$, $p < 0.05$; and sharing $\beta = -0.20$, $p < 0.05$).

13.3.3. Stranger

For the stranger-oriented model, results indicated that women were only significantly higher on emotional helping ($\beta = -0.18$, $p < 0.05$) and inclusion ($\beta = -0.41$, $p < 0.001$). There was a non-significant difference between men and women for defending behavior ($\beta = -0.13$, $p = 0.177$), physical helping ($\beta = 0.02$, $p = 0.827$), and sharing ($\beta = 0.01$, $p = 0.790$).

13.4. Step four: correlation testing with established prosocial measure

To determine construct validity, bivariate correlational analyses were conducted to analyze how the family, friend, and stranger measures were associated with previously established measures of prosocial behavior. Factors of our multidimensional measure were separately correlated with the extant measures of prosocial behavior with the corresponding targets. The analyses revealed that the five factors generally shared a large portion of variance with the extant measure of prosocial behavior. Specific factors that had lower than a 50% correlation included emotional support (strangers $r = 0.44$); physical help (family $r = 0.45$; strangers $r = 0.43$); sharing (family $r = 0.45$; friend $r = 0.40$; strangers $r = 0.34$) (see Table 4).

13.5. Step five: correlation testing with established correlates of prosocial behavior

To determine convergent validity, bivariate correlational analyses were conducted to examine the unique relations between each factor from the target models and expected correlates of those measures. A measure of sympathy was correlated with each measure of prosocial behavior. Additionally, the family-oriented prosocial measure was correlated with maternal and paternal closeness; the friend-oriented measure was correlated with friendship; and the stranger-oriented prosocial measure was correlated with self-regulation. Each of these correlations showed justifiable patterns; though paternal closeness and self-regulation had few significant correlations to the factors (see Table 4).

13.6. Step six: developmental gender differences

The CFA analysis conducted on the Flourishing Families sample revealed that the five factor models validated in the previous sample were replicated with good model fit for each model in this second sample. When mean levels were compared across samples, the family models showed the most marked difference with women scoring significantly higher on every factor in the older sample, compared to women only scoring higher on emotional support in the younger sample (see Table 5). This trend was reversed for the friend oriented models with the younger sample indicating more gender differences favoring women than the older sample. Women and men were generally equally likely to help strangers across most prosocial behaviors.

Table 4
Correlations between emergent prosocial factors and correlative measures.

	Family Model				Friend Model			Stranger Model		
	Extant Family Prosocial	Maternal Closeness	Paternal Closeness	Sympathy	Extant Friend Prosocial	Friendship	Sympathy	Extant Stranger Prosocial	Self-Reg.	Sympathy
F1	0.57*	0.12*	0.05	0.38*	0.53*	0.18*	0.38*	0.52*	0.13*	0.38*
F2	0.58*	0.15*	0.06	0.40*	0.56*	0.21*	0.43*	0.44*	0.08	0.43*
F3	0.59*	0.16*	0.12*	0.47*	0.52*	0.21*	0.47*	0.52*	0.15*	0.47*
F4	0.45*	0.16*	0.09	0.35*	0.51*	0.12*	0.36*	0.43*	0.06	0.36*
F5	0.45*	0.10*	0.07	0.28*	0.40*	0.11*	0.31*	0.34*	0.03	0.31*

Note: * indicates $p < 0.05$.

Table 5
Gender differences between samples.

	Family		Friend		Stranger	
	Amazon Turk (Older)	Flourishing Family (Younger)	Amazon Turk (Older)	Flourishing Family (Younger)	Amazon Turk (Older)	Flourishing Family (Younger)
F1: Defending	$\beta = -0.28$*	$\beta = -0.13$	$\beta = -0.22$	$\beta = -0.32$*	$\beta = -0.06$	$\beta = -0.13$
F2: Emotional	$\beta = -0.35$*	$\beta = -0.31$*	$\beta = -0.45$*	$\beta = -0.40$*	$\beta = -0.19$	$\beta = -0.18$*
F3: Inclusion	$\beta = -0.26$*	$\beta = -0.18$	$\beta = -0.29$*	$\beta = -0.31$*	$\beta = -0.24$*	$\beta = -0.41$*
F4: Physical	$\beta = -0.24$*	$\beta = -0.006$	$\beta = 0.04$	$\beta = -0.19$*	$\beta = 0.19$	$\beta = 0.02$
F5: Sharing	$\beta = -0.25$*	$\beta = 0.19$	$\beta = -0.29$*	$\beta = 0.20$*	$\beta = 0.04$	$\beta = 0.01$

Note: * indicates women's mean scores significantly higher than men's $p < 0.05$.

14. Study 2 discussion

The purpose of Study 2 was to validate our multidimensional measure on a different, younger sample, and to correlate it with an extant measure and cross-sectional correlates of prosocial behavior. In support of the relational perspective of prosocial behavior (Padilla-Walker & Christensen, 2011), patterns of behavior were different between men and women across different targets of prosocial behavior. One such difference is that men seemed less inclined to be prosocial toward their friends than did women. One explanation for this finding is the age of the participants in the sample; Hine and Leman (2013) found that adolescence is a particularly sensitive time for men to live up to emotionally stoic masculine stereotypes. During late adolescence both men and women are relentless promoters of hegemonic masculinity (Carrigan et al., 1985), and Pascoe (2007) found that pressure to conform to masculine ideals manifests constantly in high school relationship negotiations and communication rituals.

In this sample, every factor but emotional support had non-significant gender differences for prosocial behavior toward family. It has been posited that men might feel less pressure to "act like a man" when around their mothers. Fathers generally take the lead in things like norm compliance (Lamb, 2004) and emotional regulation (Roberts, 1999), but Roberts (1999) also found that paternal regulation can predict lower prosocial behavior. Whatever gendered pressures and socialization that are happening in the home, it is likely different from the socialization that men feel from their peers. This research sought to address whether stereotypes of masculine emotional stoicism and physical toughness extended to the field of prosocial behavior. Outcomes like this indicate that they do.

15. General discussion

As the study of prosocial research continues to diversify, more is being learned about helping behavior in various situations. Researchers are becoming increasingly concerned about the multidimensional factors that influence how and why individuals are prosocial. This measure seeks to continue this diversification of prosocial research by analyzing potential interactions that gender, age, and relationships might have on prosocial behavior. Prior prosocial measures have analyzed different types of prosocial behaviors and the contexts in which individuals help (Caprara, Steca, Zelli, & Capanna, 2005; Carlo & Randall, 2002). Other measures have honed in on how relationships might affect the frequency of helping behavior (Padilla & Christensen, 2011). This measure seeks to build on previous measures in an attempt to consider type and target, as well as specifically focusing on reducing gender differences that might be overestimated due to self-reported biases. A broad measure of potential prosocial behaviors was compiled, and factor analyses provided support for the items loading in five factors across different targets and samples. All twenty prosocial items achieved internal reliability and validity in the models, and held together in hypothesized factors across different targets and samples. One focus of this measure is to address potential gender differences in prosocial behavior by including helping behaviors that might be more in in line with how men are socialized to help others. The present analyses indicated that while differences exist in the ways that men and women help others, like gender differences in other areas, men and women are more alike than they are different (Zell, Krizan, & Teeter, 2015).

15.1. Less gender differences with age

With the exception of the family measure, the trend was for gender differences to become less pronounced in the older sample. Although the gender intensification theory specifies that pressures increase as physical bodies mature (with no projected end), Hine and Leman (2013) showed especially strong gendered pressure regarding prosocial values in *early* adolescence. Other research indicates that gender role proscriptions remain in force throughout emerging adulthood (Kimmel, 2008). Our data indicates that there may be a lessening of pressure for men to avoid certain types of helping as they age, though emotional support and inclusion generally remained low. Women seem to have a corner on the market regarding friend-oriented prosocial behavior, but in the older sample there was no difference in levels of defending behavior and physical help toward friends. This may be an indicator that the pressure Hine and Leman (2013) documented may be particularly strong within the adolescent friend-network. This tallies with accounts teens themselves give about the decreased emotional connectivity with their male friends they feel allowed to express throughout adolescence (Way, 2011). It also may indicate that there is a relaxing of pressure regarding certain kinds of helping behaviors within the friend network as men move through emerging adulthood. Caprara et al. (2005) found interesting developmental trends in prosocial behaviors with increased levels for adolescents and older adults. They found that people in these age groups did not have more prosocial inclinations than other age groups, but they received more personal satisfaction for each prosocial act. This finding, along with those in this research indicate the importance of treating prosocial behavior as a developmental construct in future research.

Another interesting age difference was family-oriented behavior: the older sample had consistent gender differences, with women higher across the board, whereas the younger sample had no differences except on emotional support. This might be explained by more of the younger sample living at home (49% of women, 37% of men) compared to the older sample (21% of women and 27% of men). It has been posited (Carlo et al., 2007; Padilla-Walker et al., 2015) that prosocial behavior toward family might be heavily associated with compliance to parental requests and standing arrangements about helping with chores, etc. Perhaps men still living at home participate in obligatory household chores, but are less willing than women (or have fewer opportunities) to help after moving out on their own.

15.2. Advantages of a multidimensional measure

The value of this measure extends beyond clarifying the gender differences of prosocial behavior. This measure capitalizes on the multidimensional directions currently being pursued by prosocial researchers including a developmental lens of the specific pressures and practices of late adolescent/college-aged samples, different types of prosocial behavior, how different relationships affect prosocial behavior, and the interaction between types and targets of prosocial behavior. Extant prosocial measures have honed in on a specific aspect of prosocial behavior such as types or targets of behavior, but lack the ability to capture these multiple dimensions at once. The Prosocial Tendencies Measure (Carlo & Randall, 2002) and Caprara and colleague's scale (2005) focus on prosocial behavior in different helping contexts and the broad description of helping items generally avoids an emotional bias, but there is no discussion on how relationships might affect one's behavior or motivations. Padilla-Walker and Christensen's (2011) adaptation of Peterson and Seligman's (2004) Inventory of Strengths adds a relational focus, but is limited to very emotion-centered items. Similar to the increasing specificity of research conducted on various types and aspects of aggression, prosocial research is progressing toward multidimensionality and requires measures capable of assessing differences such as types of behavior, gender differences and relational targets to keep moving the field forward.

15.3. Limitations and future directions

The present validation study was limited to participants from the United States. As described above, the development of the multidimensional prosocial measure was based on research from several countries, but additional validation of the measure from other nations would be a next step. It would also be fruitful to extend the present study to adolescents and adult populations because, as Greener and Crick (1999) suggested, different age groups might think more broadly about prosocial research than traditional research. Although this study attempted to include various forms of prosocial behavior, more research should continue to investigate the validity of humor, peacemaking, and inclusion as well as other potentially prosocial behaviors.

Moving forward, the similarities and target differences in masculine and feminine prosocial behaviors are arguably the most important finding of this work. It may be that the using a multidimensional scale does not have statistical significant gender differences, but it suggests practical differences in the way prosocial behavior should be taught to men versus women. For instance, practitioners interested in coaching prosocial behavior for men should pay attention to the pressures that men likely feel not to appear emotionally sensitive and overtly empathic. While we obviously wish to encourage emotional literacy in men, it is likely worthwhile to approach the topic with language couched in terms they are familiar with, similar to the way items in the new measure were rephrased.

As described at the outset, the purpose of this multidimensional measure was to create a measure of prosocial behavior that more adequately captured masculine prosocial behaviors. Given the present validation, we would suggest that this multidimensional measure may now be deployed to extend both theory and research in moral psychology. The use of this measure can address the increasing concern over the potential gender differences (Eisenberg et al., 2015; Zarbatany et al.,

1985) or biases (Chapman et al., 2006; Hastings et al., 2007; Skoe et al., 2002) in prosocial research that have been raised over the last 30 years. The data in these two samples indicates that tailoring measures to male-specific prosocial behaviors may yield more accurate views of the ways men help others.

Appendix A

Defending

1 If I see someone being given a hard time, I stand up for that person
2 If someone is being made fun of, I stick up for that person
3 When people are fighting with each other, I try to help them get along
4 I break up fights

Emotional Helping

5 If someone is sad, I try to make that person laugh
6 If someone is upset, I help that person let off steam
7 If someone is upset, I listen to that person
8 If someone is upset, I do something fun with that person

Inclusion

9 I am nice to others, even if I don't like them
10 If someone is new to a group, I make an effort to include that person
11 I accept others for who they are, even if they are different
12 If I notice someone who is lonely, I try to include that person

Physical Helping

13 I do physical acts of service for others [lifting heavy things, yard work, cleaning]

14 If I see someone hurt themselves, I help that person

15 I help people in an emergency

16 I would enter a risky situation in order to help someone

Sharing

17 I share with people [food, clothes, car]

18 I pick up the tab for people

19 I lend money to people

20 I share my personal belongings with people

Note: These items can be used to assess prosocial behavior toward different targets by prompting participants to think of their behavior toward a specific type of people (i.e. family, friend, or stranger) in repeated administrations of the measure. For example, one section might say: "the following questions will ask you about your behavior with strangers", and the second one would being with: "the following questions will ask you about your behavior with family".

References

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*, 411–423.

Barber, B. K., & Olsen, J. A. (1997). Socialization in context connection, regulation, and autonomy in the family, school, and neighborhood, and with peers. *Journal of Adolescent Research, 12*(2), 287–315. http://dx.doi.org/10.1177/0743554897122008.

Bergin, C., Talley, S., & Hamer, L. (2003). Prosocial behaviours of young adolescents: A focus group study. *Journal of Adolescence, 26*(1), 13–32. http://dx.doi.org/10.1016/S0140-1971(02)00112-4.

Caprara, G. V., Steca, P., Zelli, A., & Capanna, C. (2005). A new scale for measuring adults' prosocialness. *European Journal of Psychological Assessment, 21*(2), 77–89. http://dx.doi.org/10.1027/1015-5759.21.2.77.

Carlo, G., McGinley, M., Hayes, R., Batenhorst, C., & Wilkinson, J. (2007). Parenting styles or practices? Parenting, sympathy, and prosocial behaviors among adolescents. *The Journal of Genetic Psychology, 168*(2), 147–176. http://dx.doi.org/10.3200/GNTP.168.2.147-176.

Carlo, G., & Randall, B. A. (2002). The development of a measure of prosocial behaviors for late adolescents. *Journal of Youth and Adolescence, 31*(1), 31–44. http://dx.doi.org/10.1023/A:1014033032440.

Carrigan, T., Connell, B., & Lee, J. (1985). Toward a new sociology of masculinity. *Theory and Society, 14*(5), 551–604. http://dx.doi.org/10.1007/BF00160017.

Chapman, E., Baron-Cohen, S., Auyeung, B., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy: Evidence from the empathy quotient (EQ) and the "reading the mind in the eyes" test. *Social Neuroscience, 1*(2), 135–148. http://dx.doi.org/10.1080/17470910600992239.

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling, 12*(3), 471–492.

Crouter, A. C., Whiteman, S. D., McHale, S. M., & Osgood, D. W. (2007). Development of gender attitude traditionality across middle childhood and adolescence. *Child Development, 78*(3), 911–926. http://dx.doi.org/10.1111/j.1467-8624.2007.01040.x.

Davis, M. H. (1983). The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of Personality, 51*(2), 167–184. http://dx.doi.org/10.1111/j.1467-6494.1983.tb00860.x.

DeVellis, R. F. (2012). *Scale development: Theory and applications* (Vol. 26). Sage publications.

Eagly, A. H. (2009). The his and hers of prosocial behavior: An examination of the social psychology of gender. *American Psychologist, 64*(8), 644–658. http://dx.doi.org/10.1037/0003-066X.64.8.644.

Eberly, M. B., & Montemayor, R. (1999). Adolescent affection and helpfulness toward parents: A 2-year follow-up. *The Journal of Early Adolescence, 19*(2), 226–248. http://dx.doi.org/10.1177/0272431699019002005.

Eisenberg, N. (1983). Children's differentiations among potential recipients of aid. *Child Development, 54*(3), 594–602. http://dx.doi.org/10.2307/1130046.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg (Ed.), *Handbook of child psychology* (pp. 646–718). New York: Wiley. http://dx.doi.org/10.1002/9780470147658.chpsy0311.

Eisenberg, N., Spinrad, T. L., & Knafo-Noam, A. (2015). Prosocial development. In M. E. Lamb, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science, Vol. 3: Socioemotional processes* (7th ed., pp. 610–656). Hoboken, NJ, US: John Wiley & Sons.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of mechanical Turk samples. *Journal of Behavioral Decision Making, 26*(3), 213–224. http://dx.doi.org/10.1002/bdm.1753.

Greener, S., & Crick, N. R. (1999). Normative beliefs about prosocial behavior in middle childhood: What does it mean to be nice? *Social Development, 8*(3), 349–363.

Gretarsson, S. J., & Gelfand, D. M. (1988). Mothers' attributions regarding their children's social behavior and personality characteristics. *Developmental Psychology, 24*(2), 264–269. http://dx.doi.org/10.1037/0012-1649.24.2.264.

Hastings, P. D., McShane, K. E., Parker, R., & Ladha, F. (2007). Ready to make nice: Parental socialization of young sons' and daughters' prosocial behaviors with peers. *The Journal of Genetic Psychology, 168*(2), 177–200. http://dx.doi.org/10.3200/GNTP.168.2.177-200.

Hill, J. P., & Lynch, M. E. (1983). The intensification of gender-related role expectations during early adolescence. In J. Brooks-Gunn, & A. C. Petersen (Eds.), *Girls at puberty: Biological and psychosocial perspectives* (pp. 201–228). NY: Springer.

Hine, B., & Leman, P. J. (2013). The developing relationship between gender and pro-social behaviour. In P. J. Leman, & H. R. Tenenbaum (Eds.), *Gender and development* (pp. 78–108). New York, NY, US: Psychology Press.

Holden, C. J., Dennie, T., & Hicks, A. D. (2013). Assessing the reliability of the M5-120 on Amazon's mechanical Turk. *Computers in Human Behavior, 29*(4), 1749–1754. http://dx.doi.org/10.1016/j.chb.2013.02.020.

Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological Bulletin, 126*(5), 703–726. http://dx.doi.org/10.1037/0033-2909.126.5.703.

Kimmel, M. (2008). *Guyland: The perilous world where boys become men.* New York: Harper Collins.

Lamb, M. E. (2004). In *The role of the father in child development* (4th ed.). NY: Wiley.

Little, T. D. (2013). *Longitudinal structural equation modeling.* New York: Guilford Press.

Michalska, K. J., Kinzler, K. D., & Decety, J. (2013). Age-related sex differences in explicit measures of empathy do not predict brain responses across childhood and adolescence. *Developmental Cognitive Neuroscience, 3*, 22–32. http://dx.doi.org/10.1016/j.dcn.2012.08.001.

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus User's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Novak, S. P., & Clayton, R. R. (2001). The influence of school environment and self-regulation on transitions between stages of cigarette smoking: A multilevel analysis. *Health Psychology, 20*(3), 196. http://dx.doi.org/10.1037/0278-6133.20.3.196.

Padilla-Walker, L. M., & Carlo, G. (2014). The study of prosocial behavior: Past, present, and future. In L. M. Padilla-Walker, & G. Carlo (Eds.), *Prosocial development: A multidimensional approach* (pp. 3—16). New York: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199964772.003.0001.

Padilla-Walker, L. M., & Christensen, K. J. (2011). Empathy and self-regulation as mediators between parenting and adolescents' prosocial behavior toward strangers, friends, and family. *Journal of Research on Adolescence, 21*(3), 545—551. http://dx.doi.org/10.1111/j.1532-7795.2010.00695.x.

Padilla-Walker, L. M., Dyer, W. J., Yorgason, J. B., Fraser, A. M., & Coyne, S. M. (2015). Adolescents' prosocial behavior toward family, friends, and strangers: A person-centered approach. *Journal of Research on Adolescence, 25*(1), 135—150. http://dx.doi.org/10.1111/jora.12102.

Pascoe, C. J. (2007). *Dude, you're a fag: Masculinity and sexuality in high school.* Berkeley, CA, US: University of California Press.

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification.* Washington, DC: Oxford University Press.

Pleck, J. H. (1983). *The myth of masculinity.* Cambridge, MA: MIT Press.

Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research, 23*(1), 51—67. http://dx.doi.org/10.1207/s15327906mbr2301_3.

Roberts, W. L. (1999). The socialization of emotional expression: Relations with prosocial behaviour and competence in five samples. *Canadian Journal of Behavioural Science, 31*(2), 72—85. http://dx.doi.org/10.1037/h0087075.

Robinson, C. C., Mandleco, B., Olsen, S. F., & Hart, C. H. (2001). The parenting styles and dimensions questionnaire (PSQD). In B. F. Perlmutter, J. Touliatos, & G. W. Holden (Eds.), *Handbook of family measurement techniques: Vol. 3. Instruments & index* (pp. 319—321). Thousand Oaks: Sage.

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75,* 243—248.

Shigetomi, C. C., Hartmann, D. P., & Gelfand, D. M. (1981). Sex differences in children's altruistic behavior and reputations for helpfulness. *Developmental Psychology, 17*(4), 434—437. http://dx.doi.org/10.1037/0012-1649.17.4.434.

Skoe, E. A., Cumberland, A., Eisenberg, N., Hansen, K., & Perry, J. (2002). The influences of sex and gender-role identity on moral cognition and prosocial personality traits. *Sex Roles, 46*(9—10), 295—309. http://dx.doi.org/10.1023/A:1020224512888.

Söchting, I., Skoe, E. E., & Marcia, J. E. (1994). Care-oriented moral reasoning and prosocial behavior: A question of gender or sex role orientation. *Sex Roles, 31*(3—4), 131—147. http://dx.doi.org/10.1007/BF01547711.

Suri, S., & Watts, D. J. (2011). Cooperation and contagion in web-based, networked public goods experiments. *Plos ONE, 6*(3). http://dx.doi.org/10.1371/journal.pone.0016836.

Way, N. (2011). *Deep secrets.* Harvard University Press.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology, 28*(1), 126—136. http://dx.doi.org/10.1037/0012-1649.28.1.126.

Zarbatany, L., Hartmann, D. P., Gelfand, D. M., & Vinciguerra, P. (1985). Gender differences in altruistic reputation: Are they artifactual? *Developmental Psychology, 21*(1), 97—101. http://dx.doi.org/10.1037/0012-1649.21.1.97.

Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist, 70*(1), 10—20. http://dx.doi.org/10.1037/a0038208.