University of Miami
Scholarly Repository

2019-04-03

Can Studies of Trait Altruism be Trusted?

William H.B. McAuliffe
University of Miami, w.mcauliffe@umiami.edu

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

UNIVERSITY OF MIAMI

CAN STUDIES OF TRAIT ALTRUISM BE TRUSTED?

By

William Harold Blake McAuliffe

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

May 2019

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

CAN STUDIES OF TRAIT ALTRUISM BE TRUSTED?

William Harold Blake McAuliffe

Approved:

_____
Michael McCullough, Ph.D.
Professor of Psychology

_____
Debra Lieberman, Ph.D.
Associate Professor of Psychology

_____
Elizabeth Simpson, Ph.D.
Assistant Professor of Psychology

_____
Cengiz Zopluoglu, Ph.D.
Assistant Professor of Education

_____
Edward Erwin, Ph.D.
Professor of Philosophy

_____
Guillermo Prado, Ph.D.
Dean of the Graduate School

MCAULIFFE, WILLIAM HAROLD BLAKE   (Ph.D., Psychology)

<u>Can Studies of Trait Altruism be Trusted?</u>   (May 2019)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Michael McCullough.
No. of pages in text. (50)

Studies using the self-report altruism scale (SRA) to measure trait altruism have found that altruism is uncorrelated with antisocial motives and positively correlated with being female and number of sex partners. However, it is unknown whether the SRA scale meets basic psychometric standards, such as unidimensionality. The act-frequency format of the SRA likely undermines its ability to isolate any one motive for behaving prosocially. In a pilot study ($N =$ 276), factor analysis indicated that the SRA contains dimensions related to the performance of favors (Factor 1), charitable giving (Factor 2), and politeness among strangers (Factor 3). Factor 2 predicted laboratory charitable giving, possibly due to having items related to charity. The present study ($N = 814$) replicated the three-factor structure of the SRA, with the majority of items loading on Factor 1. Factor 1 was *negatively* associated with criterion measures of altruism, and was the only factor that positively correlated with antisocial motives and number of sex partners. Factors 2 and 3, as well as other self-report measures of altruism, generally had the exact opposite correlates of Factor 1. Overall, treating the SRA as unidimensional has obscured trait altruism's relationships to other constructs.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

The self-report altruism scale (SRA) was designed to measure trait altruism. It consists of twenty "act-frequency" items that query how many times participants have performed various acts that benefit other people (Rushton, Chrisjohn, & Fekken, 1981). The SRA has been used to establish many influential empirical claims, including that women are more altruistic than men (Rushton, Fulker, Neale, Nias, & Eysenck, 1986) and that altruism is uncorrelated with antisocial behavior (Krueger, Hicks, & McGue, 2001). More generally, the SRA is an extremely popular instrument: As of March 2019 it has been cited 870 times on Google Scholar.

Construct Validity?

Do the psychometric properties of the SRA justify its wide use and key role in influential studies? To understand why the answer is probably negative, it is necessary to first specify what altruism is, and what would make a measure of trait altruism valid. The word "altruism" is sometimes used in the *behavioral* sense to refer to any helping behavior; it is also sometimes used in the *biological* sense to refer to genotypes that were naturally selected because their average phenotypic effect improved the fitness of *other, genetically related* organisms in ancestral environments. However, personality psychologists use the term "trait" to refer to stable differences in a psychological construct (Funder, 1991); thus, I use "altruism" here to refer to *psychological* altruism, which denotes a desire to improve the welfare of a given beneficiary as an end in itself

(Batson, 2011). The distinction between biological and psychological altruism is akin to the distinction between the *ultimate* cause of a trait—that is, the evolutionary process that explains why the trait came to fixation in the population—and the *proximate* cause of a trait—that is, the psychological processes that undergird trait-relevant behavior (Scott-Phillips, Dickins, & West, 2011). Psychological altruism can be contrasted with *egoistic* motives to help others as a means to benefitting the self (e.g., to accrue a reputation as a generous person) and *moral* motives to help as a means to upholding moral beliefs (e.g., to comply with a valued religious rule of donating a percentage of income).

The nature of altruistic motivation points to the criteria that a valid measure of trait altruism must meet. First, for altruism to count not only as a psychological construct, but also as a trait, there must be stable between-person differences in the strength or chronicity of altruistic motivation. Notably, Rushton's et al.'s (1981) explicit goal in creating the SRA was to provide evidence that trait altruism exists, counter to the then-prevailing wisdom that "it is undoubtedly futile to search for the altruistic personality" (Piliavin & Charng, 1990, p. 31). Second, a measure of trait altruism is only valid if there is a causal chain linking differences in scores on the measure to differences in trait altruism (Borsboom, Mellenbergh, & van Heerden, 2004). Thus, the SRA is valid only if the acts that participants have in mind when they respond to SRA items were caused by altruistic motives.

Rushton et al. (1981) were ill-poised to demonstrate that altruistic motives cause the behaviors reported on the SRA because their definition of trait altruism occurs at the level of behavior: "…there is a trait of altruism. That is, some people are consistently more generous, helping and kind than others" (p. 296). Thus, *any* prosocial behavior counts as altruism, and validating the SRA becomes merely a matter of showing that it correlates with other instances of prosocial behavior. Rushton et al.'s (1981) report that SRA scores were positively correlated with peer-reports of prosocial behavior, some laboratory prosocial behaviors, and self-reported prosocial orientation would thereby seem to prove conclusively in favor of the SRA's validity. Yet, it remains possible that these criteria measures themselves do not reflect altruistic motivation, but rather some other motive that can cause prosocial behavior. Consistent with this possibility, one study found that anonymous kidney donation, which likely *is* altruistically motivated, is unassociated with SRA scores (Brethel-Haurwitz et al., 2016). Perhaps, then, the SRA is correlated only with other instances of prosocial behavior that are *not* caused by altruism.

Unidimensionality?

Researchers have also yet to test whether the SRA is unidimensional—that is, whether its items covary due to a single common cause. Establishing unidimensionality is an integral component of a validity argument (Clark & Watson, 1995): If the SRA is multidimensional, then its scores do not reflect the strength of any one motive in particular, but rather an admixture of multiple

motives. Extant efforts to test the SRA's dimensionality have been inadequate. Rushton et al. (1981) reported that the SRA is highly reliable (alpha = .88). Although researchers often assume that high internal consistency implies unidimensionality, coefficient alpha is only a meaningful statistic if the scale in question is unidimensional. Indeed, coefficient alpha can yield spuriously high reliability estimates in the presence of multidimensionality (Cortina, 1993).

Landis et al. (2009) conducted a principal components analysis (PCA) on SRA scores, which yielded two components. However, this result is not conclusive because PCA is a data reduction technique, not a method for determining dimensionality (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Arnocky, Piché, Albert, Ouellette, and Barclay (2017) claimed that an exploratory factor analysis (EFA) indicated that the SRA is unidimensional. However, the authors provided no model fit statistics. Instead, the authors made a qualitative judgment based on a scree plot, which is unreliable (Streiner, 1998).

The behaviors about which the SRA queries are unlikely to be explained wholly by altruism, or any other one motive, because the act-frequency format deliberately omits the context in which the behaviors took place (Block, 1989). For instance, one SRA item asks participants how many times they have donated money to charity. Without more information it is not possible to know whether people who report that they donate to charity "very often" do so because they frequently experience egoistic motives (e.g., because they regularly acquiesced to solicitations at grocery stores to avoid negative social evaluation), altruistic

motives (e.g., because they experience compassion whenever they hear about a natural disaster on the news), moral motives (e.g., they internalized a religious maxim that requires its followers to donate a certain percentage of annual income), or some combination of motives.

It may turn out that certain prosocial acts have prototypical motives, even holding the context of such acts constant. Preliminary evidence that SRA items have prototypical causes comes from Wilhelm and Bekkers (2010), who divided ten SRA items into those that describe spontaneous helping and those that describe planned helping on the grounds that moral principles explain planned giving, whereas empathic responses dictate spontaneous giving. The authors found in a representative U.S. adult sample that moral principles indeed predicted planned helping particularly strongly, whereas the effect of dispositional empathy was relatively stronger for unplanned helping.

Measurement Invariance and Gender Differences

Another issue with the SRA is that it has been as used as evidence that women are more altruistic than men (Rushton, Fulker, Neale, Nias, & Eysenck, 1986) without first establishing that it is measurement invariant across genders. Measurement invariance refers to an isomorphism in how different groups interpret and respond to items on a measure (Gregorich, 2006). If some SRA items cause higher scores in one gender for reasons unrelated to altruism, then gender differences in altruism cannot be meaningfully assessed using the measure. Whether it is reasonable to expect women to score higher on the SRA

depends on what motive(s) the SRA measures. For instance, men are more likely

to help others to achieve status, whereas women are more likely to help others for

altruistic reasons (Böhm & Regner, 2013; Costa, Terracciano, & McCrae, 2001;

Falk & Hermle, 2018). Thus, if the SRA is valid, then one would expect that, on

average, women score higher on it than men.

The Present Study

The goal of this dissertation is to determine whether the SRA possesses

adequate psychometric properties. To achieve this end, I conducted a pilot study

in which I used factor analysis to assess the SRA's (a) unidimensionality, (b)

measurement invariance across genders, and (c) construct validity. Based on the

results of the pilot study, I proposed a study to (a) replicate the findings of the

pilot study, and (b) scrutinize previous findings that have used the SRA.

Chapter 2: Pilot Study

Method

Recruitment

276 participants (190 women; M_{age} = 21.67, SD_{age} = 7.49) were recruited from the University of Miami (n = 221) and in several locations around Miami-Dade County (n = 55) to participate in a study that was designed to assess the construct validity of various economic games. The full details of the protocol are described elsewhere (McAuliffe, Forster, Pedersen, & McCullough, 2019); here I describe only the details of the protocol that are relevant to measuring altruism.

Procedure

Participants were paid a $10 base payment, plus whatever money they earned from the series of computer-based interaction tasks that they chose not to donate to charity (see the procedure below). Study sessions contained between four and fourteen participants. Each participant was brought to a private carrel and wore headphones. The protocol was administered entirely through a computer script on E-Prime (Schneider, Eschman, & Zuccolotto, 2002).

Participants simultaneously read and listened to instructions about how to play the Dictator Game (Forsythe, Horowitz, Savin, & Sefton, 1994), a popular economic game used to study prosocial behavior. The Dictator Game involves two participants—a "Dictator" and a "Recipient." The Dictator is given a monetary endowment, and decides whether to send none, some, or all of the endowment to the Recipient; the Recipient, in contrast, has no endowment and does nothing.

The Dictator Game primarily measures an egoistic desire to avoid social censure for failing to treat others equitably (Dana, Cain, & Dawes, 2006; Dana, Weber, & Kuang, 2007).

Participants played six rounds of the Dictator Game, each with a different partner (in sessions with fewer than seven participants, participants were told that some rounds may feature previous partners, but neither party would know when a repeated interaction was occurring). Consistent with standard practice, Dictators and Recipients were anonymous to each other. Each participant had an endowment of $0.50 for half of the rounds, and $1.00 during the other three rounds. Transfer options ranged from $0.00 to $0.50 (in $0.05 increments) in half of the rounds, and ranged from $0.00 to $1.00 (in $0.05 increments) in the other three rounds. The percentage of money that participants transferred across rounds was consistent (omega = .91; M = 36%, SD = 34%).

After that, participants played other economic games that are not related to altruism (these other games took approximately 15 minutes to complete). Next, participants completed the SRA (omega = .90; M = 2.93, SD = 0.62) along with several other personality scales not relevant to the present investigation. Participants completed the SRA by indicating how often they have engaged in various prosocial acts on a 5-point scale (1 = Never, 5 = Very Often). Participants were allowed to skip the item, "I have given a stranger who needed a ride a lift in my car" if they indicated that they did not drive.

After completing the questionnaires, participants saw the total dollar amount they had earned from all of the economic games. Participants were told

they would watch videos from three different charities (Oxfam, Feeding America, and International Rescue Committee), and that after each video they would have the opportunity to donate as much money from their game earnings as they wanted (in $0.05 increments) to the charity that made the video. Charities were presented in random order. Scores for each participant were computed by dividing the total amount of money sent by the total amount of money earned from the economic games ($M = 54\%$, $SD = 40\%$). After changing charitable giving percentage scores of 0 and 1, respectively, to .025 and .975, I used a logit-transformation increase the normality of the distribution. After participants responded to each of three charity videos, they were awarded the amount of earned money they chose to keep rounded up to the near dollar, plus the $10 base payment.

A positive zero- correlation between the SRA and anonymous charitable giving provide evidence of convergent validity. However, it would only be weak evidence as responding positively to a solicitation can also be caused by a desire to save face (Andreoni, Rao, & Trachtman, 2017). Thus, we controlled for dictator game giving to partial out variance in charitable giving related to acquiescing to a solicitation.

Pilot Study Results

All analyses were conducted using the psych (Revelle, 2017), lavaan (Rosseel, 2012), and car (Fox et al., 2012) packages in R version 3.4.3 (R Core Team, 2014). All tests were two-tailed and used an alpha of .05. I used full information maximum likelihood with robust standard errors to account for

violations of normality and missing data. In evaluating model fit I relied on conventional cut-offs for adequacy (Kenny, 2015): SRMR $\leq$ .08; RMSEA $\leq$ .08; TLI $\geq$ .90). Given the relatively large sample sizes in this project, I did not rely on the chi-square test of exact model fit to detect model misfit.

Unidimensionality

A confirmatory factor analysis (CFA) in which all items loaded on one factor evinced poor fit ($\chi^2(170) = 521.27$, $p < .001$; SRMR = .076; RMSEA = .088, 90% CI[.080, .097]; TLI = .739; AIC = 14877.49; BIC = 15092.05). Consequently, I used Velicer's minimum average partial (MAP) test, which has a slight under-extraction bias, and parallel analysis, which has a slight over-extraction bias, to determine the number of factors to extract (O'Connor, 2000). To conduct Velicer's MAP test, I observed the number of factors that, when partialed out from a correlation matrix of the SRA items, produced the lowest average squared off-diagonal partial correlation. Velicer's MAP test yielded two factors. To conduct the parallel analysis, I generated 1,000 datasets to create the sampling distribution of the randomly generated eigenvalues, and used the 95th percentile of the sampling distribution to determine if the observed eigenvalue accounted for significantly more variance than the corresponding randomly generated eigenvalue. Parallel analysis yielded three factors rather than two.

To adjudicate between these conflicting findings, I conducted a series of minimal residual exploratory factor analyses (EFA) with an oblimin rotation (see Table 1 for the factor loadings of each EFA). The EFA with two factors achieved mediocre fit ($\chi^2(151) = 299.26$, $p < .001$; SRMR = .06; RMSEA = .066, 90%

CI[.055, .074]; TLI = .857). The items that loaded highly and uniquely on Factor 1 described favors in response to the sudden need of others (e.g., "I have given money to a stranger who needed it (or asked me for it)"). The three items that loaded on Factor 2 were all related to charitable giving (e.g., "I have donated goods or clothes to a charity"). The two factors correlated at .44. The three-factor model had adequate fit ($\chi^2(133) = 143.70$, $p = .250$; SRMR = .040; RMSEA = .044, 90% CI[.029, .054]; TLI= .939) and split many of the items that loaded onto Factor 1 of the two-factor model onto Factor 1 and Factor 3. Factor 3 items mostly related to politeness among strangers (e.g., "I have offered my seat on a bus or train to a stranger who was standing"); items that remained on Factor 1 related to favors that likely elicit reciprocation. Factors 1 and 3 each accounted for 15% of item variance, while Factor 2 accounted for 10%.

The three factors were moderately correlated with each other ($r_{1,2} = .22$; $r_{1,3} = .49$; $r_{2,3} = .41$). This is consistent with different motives for prosocial behavior positively covarying in the population. However, several items also loaded on multiple factors, suggesting that different participants endorsed the same items for different reasons.

Gender Differences

Testing for latent mean differences across gender is valid when the measure is at least scalar invariant (Gregorich, 2006). Scalar invariance subsumes three nested types of invariance: (a) configural (i.e., the same factor structure has adequate fit in both groups), (b) metric (i.e., the model has configural invariance and the magnitude of the items' factor loadings is the same across groups), and (c)

scalar (i.e., the model has metric invariance and the items' intercept values are the same across groups). If configural invariance is violated, then factor scores are not comparable across groups because the structure of the construct differs across groups. If metric invariance is violated, then genders cannot be compared on how they relate to other variables because the groups interpret the items in different ways. If scalar invariance is violated, then group mean comparisons are not valid because one group is scoring higher than another even after controlling for each participant's position on the latent continuum.

To test for gender differences across all three SRA factors, I conducted a multi-group CFA with gender as a group factor ($\chi^2(334) = 525.55$, $p < .001$; SRMR = .071; RMSEA = .066, 90% CI[.055, .077]; TLI= .851). I assigned items to the factor they loaded highest on in the three-factor EFA. I constrained factor loadings and item intercepts equal across gender, which reduced model fit and indicated a violation of measurement invariance. After freeing the two item intercepts that violated scalar invariance, I set the factor intercepts to zero for men. Examining the factor intercepts for women, I found that women scored similarly to men on Factor 1 (M_{diff} = .00, SE = .11, Z = 0.04, p = .971, d = .01), higher on Factor 2 (M_{diff} = 0.53, SE = 0.11, Z = 4.89, $p < .001$, d = 1.00), and higher on Factor 3 (M_{diff} = 0.17, SE = 0.09, Z = 2.03, p = .043, d = .32). The fact that Factor 1 does not replicate well-established gender differences in altruism (Costa et al., 2001; Falk & Hermle, 2018) may point to its limited construct validity.

Construct Validity

I specified a structural equation model in which the three SRA latent factors and a latent variable representing the six Dictator Game decisions were entered as predictors of charitable giving. Thus, the regression coefficient for each SRA factor represented the variance that it does not share with the other two factors or the Dictator Game. (However, results were qualitatively identical when excluding the Dictator Game from the model.) Approximate model fit was good, $\chi 2(315) = 472.95$, p < .001; SRMR = .056; RMSEA = .043; TLI = .928. As expected, the Dictator Game positively predicted charitable giving, b = 8.50, SE = 1.69, $Z = 5.04$, $p <.001$, $r = .33$. Factor 1 was non-significantly related to charitable giving, b = -0.76, SE = 0.50, $Z = -1.53$, $p = .127$, $r = -.21$; Factor 2 positively predicted charitable giving, b = 0.94, SE = 0.32, $Z = 2.96$, $p = .003$, $r = .24$; and Factor 3 was not significantly associated with charitable giving, b = 0.88, SE = 0.69, $Z = 1.28$, $p = .201$, $r = .18$).

Pilot Discussion

The pilot study yielded three main findings. First, although researchers have treated the SRA as unidimensional, an EFA revealed that it has three moderately correlated dimensions. Factor 1 included items that related to providing favors that typically elicit reciprocation. Factor 2 contained items related only to charitable giving. Factor 3 contained items related mostly to anticipating the needs of others. Second, there was qualified support for Rushton et al.'s (1986) contention that women score higher on the SRA: Women scored higher on Factors 2 and 3, but not Factor 1. Third, only Factor 2 positively and

significantly predicted charitable giving. Given the fact that items on Factor 2 all relate to charity, the positive relationship between Factor 2 and charitable giving may be due to criterion contamination rather than trait altruism per se.

Relationship to previous findings

If the SRA is multidimensional, then studies that have treated the SRA as unidimensional may have yielded invalid findings. For example, Krueger et al. (2001) found no relationship between the SRA and antisocial behavior. Aggregating factors that relate to other variables in opposite directions could yield null effects if the opposing correlations are of similar magnitudes.

But what if researchers are interested in prosocial behavior, regardless of how it is caused? In this case is it legitimate to treat the SRA as a unidimensional measure of prosocial behavior and remain agnostic about its psychological underpinnings? A recent paper reporting that SRA scores predict mating success provides an illustrative test case (Arnocky et al., 2017). The authors deliberately avoided theorizing about the motives underlying prosocial behavior on the premise of caring about the evolutionary origins of costly prosocial acts rather than their proximate cause: "By 'altruism,' we mean acts that decrease some fitness-proxy in the actor (at least temporarily) and function to confer benefits on a recipient, regardless of the underlying psychological motivations" (Arnocky et al., 2017, p.1). The choice to regard any prosocial behavior as altruism on evolutionary grounds is unjustified because evolutionary explanations apply to the psychological mechanisms that regulate behavior, not behavior itself (Tooby & Cosmides, 1992). As Arnocky et al. (2017) acknowledge in their supplementary

materials, different motives underlying prosocial behavior may be differentially linked to mating success, and may each have different evolutionary origins. All Arnocky et al.'s (2017) results indicate is that at least one SRA factor is positively associated with mating success. The other factors may be uncorrelated or even negatively correlated with mating success, a possibility that would considerably qualify Arnocky et al.'s (2017) conclusion that altruism promotes mating success, especially if the factor that positively correlated with mating success does not reflect trait altruism.

Interpreting Gender Differences

Even after correcting for violations of measurement invariance, gender differences on the SRA may not represent true differences in trait altruism. Trait altruism is a "global" trait in that it should be relevant to any domain in which people have a non-instrumental desire to improve the welfare of others (Funder, 1991). Because the SRA only asks about twenty of the indefinitely large number of behaviors that could be motivated by altruism, it is possible that putative gender differences in trait altruism are caused by factors that are specific to content domains that are particularly well-represented on the SRA. For instance, men scored similarly to women on Factor 1, which included items referring to acts that require high physical strength (e.g., "I have helped a stranger whose car has broken down or helped pushed a stranger's car out of snow (or mud, water, etc.") or involve high risk for those low in physical strength. For example, even someone who wants to perform favors may not be able to endorse "I have given a stranger a lift in my car" because the perceived risk of assault or rape were too

high when opportunities to perform the act arose. Possibly, women would have scored higher than men on Factor 1 if items were a more representative sample of the behaviors that Factor 1 causes. Similarly, women's greater social involvement in religious activities (Einolf, 2011), which often involve charitable giving, may explain their higher average score on Factor 2.

Chapter 3. Main Study

In the main study, I sought to clarify the implications of the pilot study. I had three goals: (a) to observe whether the three-factor structure—and the gender differences on each factor—of the SRA would replicate in a larger sample recruited from a different population, (b) to test whether any of the SRA factors are strongly associated with non-questionnaire measures of altruistic motivation, and (c) to test whether previously published findings using a unidimensional SRA would hold when using a multidimensional SRA.

Method

A power analysis revealed that data from 779 participants yields 80% power to detect "small" effects ($r = .1$; Cohen, 1992). Consequently, I planned to recruit 800 participants from Amazon.com's Mechanical Turk (MTurk), an online crowdsourcing service. Participants were offered a $1.00 base payment to complete the study, plus a bonus payment of $2.00 for completing the study. 814 participants ended up completing the study.

Self-Report Altruism Measures

I first planned to test whether I would recover a three-factor solution for the SRA in this new sample. Results from factor analyses require cross-validation because they tend to overfit models to the idiosyncrasies of a particular sample.

To address whether the item content rather than the latent causes underlying the SRA factors is responsible for the gender differences observed in the pilot study, I had participants respond to three items about how physically formidable they perceive themselves to be relative to the rest of the U.S. population (using a sliding percentage scale; omega = .90; $M = 39.67$, $SD = 22.75$), as well as the 4-item interpersonal sub-scale of the Religious Commitment Inventory (RCI; Worthington et al., 2003; omega = .94; $M = 1.84$, $SD = 1.15$). If the role of physical formidability in Factor 1 items is obscuring a gender difference favoring women, then women should score significantly higher on Factor 1 after controlling for formidability. If the increased opportunity to donate to charity from increased religious attendance explains why women score higher on Factor 2 than men, then there should be no significant gender difference in Factor 2 scores after controlling for RCI scores

In the pilot study, Factor 2 was just-identified and Factor 3 was just barely over-identified. To avoid potential under-identification in the present study, I also included nine additional items (Factor 1: "I have used my social connections to help an acquaintance obtain a job," "I have offered to pay for an acquaintance or stranger's purchase," and "I have provided first aid or other medical attention"; Factor 2: "I have participated in a charity fundraising event (e.g., a 5k run)," "I have made a 'pledge' to make a regular contribution to a charitable cause," and "I have performed an administrative role for a charitable cause"; Factor 3: "I have left a tip when it was not mandatory (e.g., for a barista)," "I have picked up a dropped item for a stranger," and "I have yielded to another person when I had the

right of way (while driving, biking, etc.)") that I planned to add to the SRA if an EFA on the original twenty-item SRA recovered any locally under-identified factors. Since the EFA did not uncover any locally under-identified factors, I did not analyze these items further.

If none of the SRA factors evince convergent validity, then it cannot be a valid measure of trait altruism. Such a finding would leave researchers uncertain whether idiosyncratic features of the SRA are to blame for its invalidity, or rather trait altruism simply cannot be assessed with face valid self-report measures. To preempt this potential ambiguity, I included two questionnaire measures may possess more validity than the SRA insofar as their items explicitly refer to altruistic motives and do not refer to any particular content domains.

The first of these questionnaire measures NEO-altruism scale (Costa & McCrae, 1992). The NEO is a personality inventory that is designed to capture the five major personality dimensions of the so-called five-factor model of personality. The five-factor model has been the most influential personality theory in recent decades (Saucier & Srivastava, 2015). The NEO-altruism scale is a lower-order "facet" measure of agreeableness, one of the five major personality dimensions. Participants responded to the scale by indicating their level of agreement with ten statements indicating their level of altruistic motivation (e.g., "I love to help others") on a 5-point Likert-type scale (1 = strongly disagree, 5 = strongly agree; omega = .88; $M = 3.98$, $SD = .63$).

The second addition was the HEXACO Interstitial Altruism Scale (Lee & Ashton, 2018). The HEXACO is a personality inventory that is designed to

capture the six major personality dimensions of the so-called six-factor model of personality. The HEXACO-Altruism scale contains items from three of the six major personality dimensions: honesty-humility, emotionality, and agreeableness. An increasing number of personality psychologists regard the HEXACO model as superior to the five-factor model (Saucier & Srivastava, 2015). Participants responded to the HEXACO-altruism scale by indicating their level of agreement with four statements indicating their level of altruistic motivation (e.g., "People see me as a hard-hearted person" [reverse-coded]) on a 5-point Likert-type scale (1 = strongly disagree, 5 = strongly agree; omega = .67; $M = 3.99$, $SD = .70$).

Addressing Previous Findings

Krueger et al. (2001) found that the SRA was uncorrelated with an act-frequency scale of antisocial behavior. Because many different motives cause antisocial behavior (Lilienfeld, 2018; Paulhus, Curtis, & Jones, 2018), I did not attempt to directly replicate this finding using an act-frequency measure of antisocial tendencies. Instead, I used measures that more directly tap psychological constructs that cause antisocial behavior: The psychopathy sub-scale of the Short Dark Triad (SD3; Jones & Paulhus, 2014) and the Assessment of Sadistic Personality (ASP); Plouffe, Saklofske, & Smith, 2017).

The SD3 is a 27-item measure of the "Dark Triad," which consists of psychopathy (e.g., "It's true that I can be mean to people"), narcissism (e.g., "I know I am special because everyone keeps telling me so"), and Machiavellianism (e.g., "I like to use clever manipulation to get my way"). I used only the 9-item psychopathy sub-scale in an effort to shorten the protocol (omega = .82; $M = 2.28$,

SD = .52). Moreover, there is evidence that Machiavellianism and psychopathy may be indistinguishable, and that all three sub-scales tap a core construct—callousness (i.e., indifference to the welfare of others; Miller, Hyatt, Maples-Keller, Carter, & Lynam, 2017; Persson, Kajonius, & Garcia, 2017).

The ASP is a 9-item measure scored on a 5-point Likert-type scale (1 = strongly disagree, 5 = strongly agree) that asks participants whether they agree with statements concerning the possession of sadistic desires (e.g., "Hurting people would be exciting"; omega = .90; *M* = 1.85, *SD* = .62). I regarded sadism as a proactive desire to engage in antisocial behavior, making it distinct from psychopathy, which taps a ruthless pursuit of short-term, self-interested goals rather than an intrinsic interest in antisocial behavior per se (Buckels, Jones, & Paulhus, 2013).

Mating Success Indices. I asked participants about the total number of sex partners they have had, as well as the number of casual sex partners they have had. These were the two measures that were most robustly correlated with measures of altruism in Arnocky et al. (2017). Bivariate correlations between mating success indices and individual items on the SRA suggest that Factor 1 items are the most strongly correlated with number of sex partners (see Table S2 in Arnocky et al., 2017).

Validity Measures

My goal in assessing convergent validity was to include validated measures of altruistic motivation that (a) share minimal method variance with the SRA, and (b) share minimal method variance with each other. The first measure I

used was the Values Embedded in Narrative (VEiN) coding procedure (Frimer,

Walker, & Dunlop, 2009) as applied to the Personal Strivings List (PSL;

Emmons, 1999). To complete the PSL, participants completed the sentence "I

typically try to . . . ." with their characteristic strivings—that is, the goals that they

are trying to achieve in their everyday lives. Participants reported the ten strivings

that best characterize their prototypical motives. Two raters who were trained in

the VEiN coding procedure independently evaluated each motive for evidence of

benevolence (95% raw agreement; kappa = .80) and *universalism* (99% raw

agreement; kappa = .67), which are values that underpin prosocial behavior cross-

culturally (Schwartz, 1992). Thus, participants scored between 0 and 10 on

benevolence ($M = 2.38$, $SD = 1.61$) and universalism ($M = .23$, $SD = .56$).

Benevolence was coded for when a striving indicated "concern for the welfare of

others in everyday interaction" (Frimer et al., 2009, p. 11). I regarded

benevolence as a convergent validity measure of altruistic motivation.

Universalism was coded for when a striving reflected "understanding,

appreciation, tolerance, and protection for the welfare of people beyond the

primary reference group (e.g., all people, the community, disadvantaged people)"

(Frimer et al., 2009, p. 10). I regarded universalism as a discriminant validity

measure of a *moral* motivation to benefit others (Batson, 2011).

  Although the SRA and PSL share method variance insofar as they are both

self-report measures, they have different methodological properties in that the

SRA requires participants to *respond* to items about *behavior*, whereas the PSL

has participants *produce* responses regarding *motives*. Importantly, the PSL also

minimizes socially desirable responding: Asking participants to report their characteristic motives cues much weaker concerns about appearing prosocial than do standard self-report measures that ask participants to what extent they possess prosocial goals (Frimer, Schaefer, & Oakes, 2014). Most importantly, benevolence and universalism strivings have been found to distinguish people who have been recognized for committing their lives to benefitting other people from matched controls (Frimer, Walker, Dunlop, Lee, & Riches, 2011), suggesting that benevolence and universalism indeed reflect unselfish motives.

I also included an experimentally validated behavioral measure of altruistic motivation: Prosocial behavior is caused by empathic concern, an emotion that reflects altruistic motivation towards distressed others (Batson, 2011). If the SRA factors reflect trait altruism, then they should predict empathy-based helping. Empathy-based helping is different from the PSL not only in that it is *behavioral* rather than *reported*, but also in that it measures how people *react* to situations that elicit altruistic motivation, whereas the PSL measures *proactive* altruistic motives that cause people to create and approach situations in which they can improve the welfare of others. To derive the association between empathic concern and prosocial behavior, participants watched a 90-second charity video (https://www.youtube.com/watch?v=6b-jmcZJVEk). To encourage participants to pay attention to the entire video, they read that they would be asked to given their impressions of the video after it ended.

After finishing the video, participants indicated how strongly they were feeling 13 emotions on a 7-point Likert-type scale (1= not at all, 7 = extremely).

Mixed among eight distractors (e.g., "bored" and "inspired") were five adjectives that formed an empathic concern index: sympathetic, compassionate, empathic, softhearted, and tender (omega = .94; $M = 5.62$, $SD = 1.46$; Batson, Early, & Salvarini, 1997). Participants then learned that they have an opportunity to donate some or all of their earnings from the study to the charity who made the video they watched. Participants indicated how much of their study payment (in $0.05 increments) they would like to donate to charity ($M = \$0.55$, $SD = \$0.66$).

Protocol

All participants received the measures in the same order. After completing demographics, participants listed their personal strivings. This task preceded the other measures so that participants would not infer that I was interested in whether they would report prosocial or antisocial goals. Participants then completed the SRA, the NEO and HEXACO altruism scales, the psychopathy and sadism scales, the sexual history items, the RCI, the physical formidability items, and finally the charitable giving task.

<div align="center">

Results

</div>

All analyses were conducted using the psych (Revelle, 2017), lavaan (Rosseel, 2012), and car (Fox et al., 2012) packages in R version 3.4.3 (R Core Team, 2014). All tests were two-tailed and used an alpha of .05. Number of sex partners (both lifetime and casual) were logit-transformed to improve their normality; as in the pilot study, charitable giving was logit-transformed as well. I used double mean-centering (Lin, Wen, Marsh, & Lin, 2010) in creating all latent interactions, and full information maximum likelihood with robust standard errors

to account for violations of normality and partially missing data. There was very little missing data; thus, nearly all participants were included in nearly all analyses.

Unidimensionality

I moved to exploratory methods after a CFA in which all 20 SRA items loaded on one factor had inadequate fit ($\chi^2(170) = 989.78$, $p < .001$; SRMR = .057; RMSEA = .076, 90% CI[.071, .081]; TLI = .799). Velicer's MAP yielded 1 factor, while parallel analysis yielded 6 factors. To address these inconsistent results, I conducted a series of EFAs. Extracting two factors yielded mediocre fit, $\chi^2(151) = 609.57$, $p < .001$; RMSR = .040; RMSEA = .062, 90% CI[.056, .067]; TLI = .867. Extracting three factors yielded adequate fit, $\chi^2(133) = 343.66$, $p < .001$; RMSR = .030; RMSEA = .051, 90% CI[.045, .056]; TLI = .911. As in the pilot study, the factors were moderately correlated, $r_{1,2} = .36$, $r_{1,3} = .46$, $r_{2,3} = .41$. Factor 1 accounted for more variance (19%) than Factor 2 (8%) and Factor 3 (9%). Four items that loaded on Factor 3 in the pilot study loaded on Factor 1 in the present study (see Table 2). These differences in factor loadings across studies may reflect sampling variation or real differences in the motives that underlie helping behaviors in the Mechanical Turk and Miami populations.

Parallel analysis revealed that the NEO has three factors and that the HEXACO has two. Velicer's MAP and exploratory factor analysis uncovered two factors for both the NEO and HEXACO altruism scales. Examination of the loadings on each scale's two-factor solution showed that the negatively valenced items loaded on a separate factor from the positively valenced items. Both the

NEO factors (r = .58) and HEXACO factors (r = .45) were moderately correlated. Because multidimensionality was caused by item valence rather than content per se—and it is not possible to excise all sources of method variance (Geiser, Eid, & Nussbeck, 2008)—I decided to treat the NEO and HEXACO as unidimensional for confirmatory analyses. The SRA likely also possesses method variance attributable to asking participants to endorse helpful acts, but this variance did not form its own factor because the SRA does not also have items about whether participants have committed unhelpful acts.

Gender Differences

To test for gender differences in the SRA factors, I created a three-factor CFA with gender as the group factor, $\chi 2(334) = 802.96$, p < .001; SRMR = .055; RMSEA = .059, 90% CI [.054, .064]; TLI = .862. Constraining the factor loadings and item intercepts equal across gender degraded model fit, indicating measurement noninvariance. I achieved scalar invariance by freeing two factor loadings and five item intercepts that modification indices indicated were noninvariant. Women scored *lower* on Factor 1, b = -0.15, SE = 0.05, Z = -3.01, p = .003, d = -0.25, but higher on Factor 2, b = 0.32, SE = 0.09, Z = 3.56, p <.001, d = 0.46, and Factor 3, b = 0.11, SE = 0.05, Z = 2.33, p =.020, d = 0.18. Regressing the factors on the RCI and formidability index did not eliminate these gender differences, consistent with the thesis that these gender differences generalize beyond the acts described on the SRA.

I compared the results from the multi-group CFA of the three-factor SRA to other indices of altruism. Beginning the NEO-altruism scale, I found that one

item violated scalar invariance. After freeing it across genders, I constrained the factor loadings and intercepts of the remaining items equal across gender, $\chi^2(87) = 537.49$, $p < .001$; SRMR = .087; RMSEA = .113, 90% CI [.106, .120]; TLI = .754. Females scored higher on the NEO, b = .27, $SE = .05$, $Z = 5.64$, $p < .001$, $d = .43$.

One item intercept was violated for the HEXACO. I estimated the factor-level gender difference after freely estimating this item across genders, $\chi^2(9) = 56.00$, $p < .001$; SRMR = .068; RMSEA = .113, 90% CI [.089, .139]; TLI = .792. Women scored higher on the HEXACO, b = .39, $SE = .06$, $Z = 6.89$, $p < .001$, $d = .54$.

Antisocial Motives

I created a series of structural equation models in which different altruism measures predicted psychopathy and sadism (see Table 3 for model fit and regression estimates). I first ran a model in which all twenty SRA items were treated as indicators of a single factor. Replicating Krueger et al. (2001), results indicated that the SRA does not significantly correlate with psychopathy ($p = .240$, $r = -.05$) or sadism ($p = .062$, $r = -.07$). Entering each of the three SRA factors as simultaneous predictors revealed that these null correlations are caused by positive and negative correlations between specific SRA factors and antisocial traits cancelling each other out. Specifically, Factor 1 positively predicted psychopathy ($p < .001$, $r = .56$) and sadism ($p < .001$, $r = .61$), while Factor 2 (psychopathy: $p = .001$, $r = -.31$; sadism: $p = .010$, $r = -.24$) and Factor 3 (psychopathy: $p < .001$, $r = -.49$; sadism: $p < .001$, $r = -.60$) evinced negative

associations. The results for Factors 2 and 3 converged with those of the NEO and HEXACO altruism scales. The NEO had a strongly negative association with sadism ($p < .001$, $r = -.67$) and psychopathy ($p < .001$, $r = -.71$), as did the HEXACO (sadism: $p < .001$, $r = -.72$; psychopathy: $p < .001$, $r = -.77$).

Sex Partners

I created a series of structural equation models in which different altruism measures predicted lifetime and casual sexual history (see Table 4 for model fit and regression estimates of the main effects models). Arnocky et al. (2017) found that the relationship between the SRA and sexual history was stronger for men, so I tested the interaction between altruism and gender in follow-up analyses.

I first ran a model in which all twenty SRA items were treated as indicators of a single factor. Results indicate that the SRA significantly predicts number of lifetime ($p <.001$, $r = .17$) and casual ($p < .001$, $r = .22$) partners. Next, I added gender and the product of gender and the SRA as an interaction term, $\chi^2(855) = 2506.48$, $p < .001$; SRMR = .055; RMSEA = .049, 90% CI [.047, .051]; TLI = .803. Women reported a significantly smaller history of lifetime (b = -0.17, $SE = 0.08$, $Z = -2.18$, $p = .029$, $d = -.17$) and casual sex (b = -0.31, $SE = 0.08$, $Z = -3.80$, $p < .001$, $d = -.31$). The interaction between gender and casual sex was significant (b = 0.38, $SE = 0.16$, $Z = 2.42$, $p = .016$, $r = .10$). Against Arnocky et al.'s (2017) finding that altruism has a stronger association with sexual history for men, the slope relating the SRA to casual sex partners was significant for *women* ($b = .20$, $SE = .04$, $\chi^2(1) = 4.65$, $p < .001$), but not men ($b = .08$, $SE = .06$, $\chi^2(1) = 1.40$, $p = .160$). The interaction between gender and lifetime sex was also

significant (b = 0.31, *SE* = 0.15, *Z* = 2.10, *p* = .036, *r* = .09). The slope relating the SRA to lifetime sex partners was large for women (*b* = .24, *SE* = .04, $\chi^2(1)$ = 5.97, *p* < .001), but still significant for men (*b* = .15, *SE* = .06, $\chi^2(1)$ = 2.66, *p* = .007).

Entering each of the three SRA factors as simultaneous predictors showed that only Factor 1 has a positive relationship with lifetime (*p* < .001, *r* = .33) and casual (*p* < .001, *r* = .37) sex partners. Factor 2 had a negative relationship with sexual history (lifetime: *p* = .091, *r* = -.16; casual: *p* = .007, *r* = -.23), while Factor 3 had no relationship with sexual history (lifetime: *p* = .874, *r* = .02; casual: *p* = .792, *r* = -.03). In three separate models I added gender and the product of gender and a particular SRA factor as predictors. The model with the product of Factor 1 and gender ($\chi^2(615)$ = 1761.61, *p* < .001; SRMR = .048; RMSEA = .048, 90% CI [.045, .051]; TLI = .858) yielded a significant interaction in predicting number of casual sex partners (*p* = .019) and lifetime sex partners (*p* = .038). Factor 1 scores positively predicted a history of casual sex for women (*b* = .57, *SE* = .13, $\chi^2(1)$ = 4.23, *p* < .001) but not men (*b* = .24, *SE* = .21, $\chi^2(1)$ = 1.13, *p* = .257). Similarly, Factor 1 scores predicted number of lifetime sex partners for women (*b* = .50, *SE* = .13, $\chi^2(1)$ = 4.03, *p* < .001) but not men (*b* = .23, *SE* = .19, $\chi^2(1)$ = 1.18, *p* = .236). There were no significant interactions between Factors 2 or 3 and gender in predicting the number of lifetime or casual sex partners.

To examine how other measures of trait altruism predicted sexual history, I first created a structural equation model in which NEO altruism predicted number of lifetime and casual sex partners (see Table 4 for full model information). The NEO was not significantly associated with lifetime sex (*p* =

.295, r = - .04), but was negatively associated with casual sex (p = .009, r = -.11).

Next, I added gender and its interaction with the NEO (χ^2(615) = 1292.01, p <

.001; SRMR = .070; RMSEA = .076, 90% CI [.073, .080]; TLI = .750). The NEO

did not significantly interact with gender in predicting number of lifetime (b = -

0.01, SE = 0.15, Z = -0.07, p = .946, r = -.00) or casual sex partners (b = 0.06, SE

= 0.17, Z = 0.35, p = .729, r = .02).

Finally, I examined whether HEXACO altruism predicts number of

casual and lifetime sex partners in a structural equation model (see Table 4 for full

model information). The HEXACO was nonsignificantly associated with lifetime

sex (p = .608, r = - .02) and marginally negatively associated with casual sex (p =

.052, r = -.09). A model with gender and its interaction with HEXACO altruism

added as predictors (χ^2(39) = 196.47, p < .001; SRMR = .062; RMSEA = .070,

90% CI [.062, .079]; TLI = .835) revealed that the HEXACO did not significantly

interact with gender in predicting number of lifetime (b = 0.05, SE = .14, Z = .03,

p = .716, r = .02) or casual (b = 0.04, SE = .16, Z = .27, p = .790, r = .01) sex

partners.

Empathy-Induced Helping

As a behavioral test of the SRA's construct validity, I examined each of its

factors' direct effects on charitable giving and indirect effects via empathic

concern (χ^2(290) = 877.93, p < .001; SRMR = .046; RMSEA = .050, 90% CI

[.046, .053]; TLI = .903; see Figure 1 for full model). Empathic concern

positively predicted charitable giving (p < .001, r = .20). Indirect effects were

significant for each SRA factor: Factor 1 reduced empathy-based helping (b = -

0.26, $SE = .09$, $Z = -2.92$, $p = .003$, $r = -.06$); Factor 2 increased empathy-based

helping ($b = 0.25$, $SE = .08$, $Z = 3.09$, $p = .002$, $r = .06$; and Factor 3 increased

empathy-based helping ($b = 0.46$, $SE = .16$, $Z = 2.89$, $p = .004$, $r = .08$). There was

also a direct effect that was significant for Factor 2 ($p = .020$, $r = .22$), perhaps

due to criterion contamination. In contrast, the direct effects of Factor 1 ($p = .117$,

$r = .11$) and Factor 3 ($p = .101$, $r = -.17$) were nonsignificant.

Next, I assessed the direct effects of the NEO altruism scale on charitable

giving and its indirect effect via empathic concern ($\chi^2(102) = 659.38$, $p < .001$;

SRMR $= .064$; RMSEA $= .082$, 90% CI [.077, .087]; TLI $= .851$). There was no

direct effect of the NEO on charitable giving, $b = 0.00$, $SE = 0.25$, $Z = 0.01$, $p =$

$.995$, $r = .00$. Instead, the effect of NEO on charitable giving was mediated by

empathic concern, $b = 0.76$, $SE = .15$, $Z = 5.02$, $p < .001$, $r = .16$.

Finally, I reran the same model using the HEXACO altruism scale instead

of the NEO altruism scale ($\chi^2(33) = 169.21$, $p < .001$; SRMR $= .033$; RMSEA $=$

$.071$, 90% CI [.063, .080]; TLI $= .932$). There was not a significant direct effect of

the HEXACO altruism scale, $b = 0.35$, $SE = 0.35$, $Z = 0.98$, $p = .326$, $r = .08$.

However, there was a significant indirect effect via empathic concern, $b = 0.60$,

$SE = 0.24$, $Z = 2.49$, $p = .013$, $r = .14$.

Benevolence Strivings

As a second test of the SRA's construct validity, I first examined whether

its factors predicted benevolence strivings, even after controlling for universalism

strivings ($\chi^2(204) = 716.75$, $p < .001$; SRMR $= .054$; RMSEA $= .056$, 90% CI

[.051, .060]; TLI $= .854$). Factor 1 was marginally negatively associated with

benevolence (b = -0.35, SE = .18, Z = -1.94, p = .052, r = -.14); Factor 2 was

positively associated with benevolence (b = 0.75, SE = 0.21, Z = 3.63, p <.001, r

= .31); Factor 3 was not significantly associated with benevolence (b = 0.16, SE =

.29, Z = 0.56, p = .574, r = .05). Universalism predicted benevolence, b = 0.42, SE

= .11, Z = 3.65, p = <.001, r = .15. Eliminating universalism from the model did

not qualitatively affect the relationships between the SRA factors and

benevolence. The fact that Factor 2 predicts benevolence strivings—most of

which are unrelated to charitable giving—suggests that its relationship with

charitable giving behavior in the present study and pilot study was not solely due

to criterion contamination.

   As before, I sequentially examined how the NEO and HEXACO relate to

benevolence strivings, controlling for universalism strivings. NEO altruism

($\chi^2(53)$ = 563.81, p < .001; SRMR = .071; RMSEA = .109, 90% CI [.103, .116];

TLI = .724) was positively associated with benevolence, b = 0.85, SE = 0.10, Z =

8.24, p <.001, r = .32. HEXACO altruism ($\chi^2(9)$ = 92.41, p < .001; SRMR = .064;

RMSEA = .107, 90% CI [.088, .127]; TLI = .716) was also positively associated

with benevolence, b = 0.68, SE = 0.11, Z = 5.99, p <.001, r = .29.

Chapter 4. General Discussion

The SRA has been used in hundreds of studies to measure trait altruism. However, the basic psychometric properties of the SRA, such as its dimensionality, has not yet been thoroughly vetted. Furthermore, Rushton et al. (1981) used a behavioral definition of altruism in creating SRA items, which is reflected in the items' lack of context or references to motivational states. Because helping behaviors have many different causes, the SRA is valid only to the extent that it measures behaviors that are the product of altruistic *motivation*—an intrinsic desire to improve the welfare of another person or group (Batson, 2011)—rather than other motives for helping. If the SRA is not a valid measure of altruistic motivation, then the conclusions that previous researchers have drawn about the relationship between trait altruism and other constructs may be spurious.

The pilot study and main study presented here suggest that the SRA in fact contains three factors. The factors were moderately correlated, likely because the behaviors described in SRA items are usually caused by one motive, but are also on occasion caused by others. Factors 2 and 3, like the NEO and HEXACO altruism measures, showed evidence of convergent validity with non-questionnaire measures of altruism, although findings for Factor 3 were less consistent. Factor 1, however, almost certainly does not measure altruistic motivation. The items that loaded on Factor 1 in both studies largely referred to favors that people could perform either out of altruistic motivation or a desire to elicit reciprocation. For example, a man could give a stranger a lift in his car in order to later solicit sex from her.

33

In line with this possibility, men scored higher on Factor 1 in the main study, despite the fact that women score higher on both behavioral and questionnaire measures of altruism (Costa et al., 2001; Falk & Hermle, 2018). Men, in contrast, score more highly on measures of malevolent motives (Muris, Merckelbach, Otgaar, & Meijer, 2017), which suggests that Factor 1 could reflect self-interested rather than altruistic motives for helping others. Consistent with this possibility, I found that variance unique to Factor 1 was *positively* associated to sadism and psychopathy, and generally negatively related to other measures of altruistic motivation.

The fact that Factor 1 likely represents a malevolent motive strongly suggests that past studies that treated the SRA as unidimensional have drawn incorrect conclusions about the relationship between altruism and other constructs. For instance, Krueger et al. (2001) observed no relationship between the SRA and questionnaire measures of antisocial traits. I replicated this finding (using different antisocial trait questionnaires) when treating the SRA as unidimensional. Separating the SRA into three factors revealed why a one-factor SRA is uncorrelated with antisocial traits: The negative associations that antisocial traits have with Factors 2 and 3 are cancelled out by their positive relationship with Factor 1. The negative relationships that the NEO and HEXACO altruism scales have with sadism and psychopathy suggest that altruism and antisocial motives really are negatively related, contra to Krueger et al. (2001).

The relationship between the SRA and sexual history found by Arnocky et al. (2017) —which I also replicated when treating the SRA as unidimensional—

also proved misleading. Although Factor 1 has a positive relationship with number of sex partners, this relationship is stronger for women than men, the opposite direction of the interaction that Arnocky et al. (2017) reported. Moreover, the other two factors and other self-report altruism measures were generally negatively correlated with number of casual sex partners and uncorrelated with number of lifetime sex partners.

The most direct evidence that Factor 1 does not measure trait altruism came from its negative relationship with benevolence strivings and empathy-mediated charitable giving. In contrast, Factor 2 (though not Factor 3) predicted benevolence strivings, and both Factor 2 and 3 predicted empathy-mediated charitable giving. The fact that the NEO and HEXACO altruism scales also positively predicted benevolence strivings and empathy-mediated charitable giving suggests that it is Factor 1, not Factors 2 or 3, that lacks construct validity.

Because Factor 2 was more consistent than Factor 3 in predicting charitable giving (in the pilot study) and benevolence strivings (in the present study), I tentatively conclude that it is the best measure of trait altruism within the SRA. The fact that Factor 2 only includes items about charitable giving, however, suggests that it has limited content validity, as evidenced by its direct effect on charitable giving. Furthermore, note that I reported Factor 2's *partial* correlations with other constructs, not its zero-order associations. Given Factor 2's high correlations with Factor 1 and Factor 3, it is likely that using Factor 2 in the absence of the other two factors would result in measuring not only altruistic

desires to donate to charity, but also giving to charity in order to be polite or to manipulate others.

Although I am arguing that the SRA contains three latent variables that vary somewhat independently across persons, this claim does not imply that examining the joint influence of the myriad causes that the three factors represent *within* a person would not be fruitful. Indeed, the decision to help another person could be viewed as a function of all the various reasons one might have for or against it in a particular context, weighted by the importance of each reason (Delton & Robertson, 2016). The weighting, in turn, may be due to either stable between-person variation or situational influences. For instance, a helping opportunity that would enable a person to solicit a later favor (perhaps engaging Factor 1), improve the welfare of a needy person (perhaps engaging Factor 2), and behave politely (perhaps engaging Factor 3) should very yield high rates of helping behavior, even though different people value each of these incentives to different extents. Similarly, if the magnitude of one incentive is much larger than that of another, then the motive engaged by the former incentive should be much more powerful within that situation than the motive engaged by the latter incentive. For instance, a helping act that is polite but does little to meaningfully improve the welfare of the beneficiary is most likely caused by a desire to be polite, even among people who value improving others' welfare more than they value being polite.

However, to study how people integrate different incentives to help others, researchers must study how people make decisions *within* particular situations

(Borsboom, Mellenbergh, & van Heerden, 2003). In particular, within-subject studies that vary the strength of various incentives for helping across a range of situations would reveal universalities and individual differences in how people integrate disparate pieces of information that are all relevant to deciding whether to help another person. Because the SRA is an individual differences measure, it cannot shed light on how the various latent variables that underlie it come together within a single person in a given situation. It is for this reason that a one-factor SRA cannot be meaningful: It combines different motives for helping others *across* people and *across* contexts. Personality measures are designed to predict individual differences in behavior, and thus have predictive power only when they tease apart each systematic source of variance.

The present findings have practical implications for the measurement of altruism. First, researchers should stop using the SRA. Substituting the SRA for a multi-method approach is likely optimal, but the NEO altruism scale is reliable and could serve well as a convenience measure. Second, if positive and negative regard for others' welfare really are opposite poles of a single dimension, then researchers should employ measures at both ends of the continuum with indifference at the mid-point. For instance, Vachon and Lynam's (2016) Affective and Cognitive Empathy scale measures both empathic concern and schadenfreude, thereby attenuating range restriction and differentiating people who are malevolent rather than merely indifferent towards others.

To sum up, the SRA is multidimensional and contains a factor that more likely represents antisocial motives than altruistic motives. The widespread use of

the SRA has perpetuated false conclusions about trait altruism, such as that it

promotes casual sex and has no relationship to antisocial motives. The study of

prosocial behavior would be better off without the SRA, especially since there are

other valid questionnaire measures of altruism available.

References

Andreoni, J., Rao, J. M., & Trachtman, H. (2017). Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy*, *125*(3), 625-653.

Arnocky, S., Piché, T., Albert, G., Ouellette, D., & Barclay, P. (2017). Altruism predicts mating success in humans. *British Journal of Psychology*, *108*(2), 416-435.

Batson, C. D. (2011). *Altruism in humans*. Oxford University Press, USA.

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imaging how you would feel. *Personality and Social Psychology Bulletin*, *23*(7), 751-758.

Block, J. (1989). Critique of the act frequency approach to personality. *Journal of Personality and Social Psychology*, *56*(2), 234-245.

Böhm, R., & Regner, T. (2013). Charitable giving among females and males: an empirical test of the competitive altruism hypothesis. *Journal of Bioeconomics*, *15*(3), 251-267.

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203-219.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*(4), 1061-1071.

Brethel-Haurwitz, K. M., Stoycos, S. A., Cardinale, E. M., Huebner, B., & Marsh, A. A. (2016). Is costly punishment altruistic? Exploring rejection of unfair offers in the Ultimatum Game in real-world altruists. *Scientific Reports*, *6,* 18974*.*

Buckels, E. E., Jones, D. N., & Paulhus, D. L. (2013). Behavioral confirmation of everyday sadism. *Psychological Science*, *24*(11), 2201-2209.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309-319.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98-104.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (Neo-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *Journal of Personality and Social Psychology*, *81*(2), 322-331.

Dana, J., Cain, D. M., & Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, *100*(2), 193-201.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), 67-80.

Delton, A. W., & Robertson, T. E. (2016). How the mind makes welfare tradeoffs: Evolution, computation, and emotion. *Current Opinion in Psychology*, *7*, 12-16.

Einolf, C. J. (2011). Gender differences in the correlates of volunteering and charitable giving. *Nonprofit and Voluntary Sector Quarterly*, *40*(6), 1092-1112.

Emmons, R. A. (1999). *The psychology of ultimate concerns: Motivation and spirituality in personality*. New York, NY: Guilford Press.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272-299.

Falk, A., & Hermle, J. (2018). Relationship of gender differences in preferences to economic development and gender equality. *Science*, *362*(6412), eaas9899.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic behavior*, *6*(3), 347-369.

Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., ... & Heiberger, R. (2012). Package 'car'. *Vienna: R Foundation for Statistical Computing*.

Frimer, J. A., Schaefer, N. K., & Oakes, H. (2014). Moral actor, selfish agent. *Journal of Personality and Social Psychology*, *106*(5), 790-802.

Frimer, J. A., Walker, L. J., & Dunlop, W. L. (2009). Values embedded in narrative(VEiN) coding manual. Unpublished manuscript, Department of Psychology, University of British Columbia, Vancouver, Canada.

Frimer, J. A., Walker, L. J., Dunlop, W. L., Lee, B. H., & Riches, A. (2011). The integration of agency and communion in moral personality: Evidence of enlightened self-interest. *Journal of Personality and Social Psychology*, *101*(1), 149-163.

Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science*, *2*(1), 31-39.

Geiser, C., Eid, M., & Nussbeck, F. W. (2008). On the meaning of the latent variables in the CT-C (M-1) model: A comment on Maydeu-Olivares and Coffman (2006). *Psychological Methods*, *13*(1), 49-57.

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, *44*(11), S78.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment*, *21*(1), 28-41.

Kenny, D. A. (2015). Measuring model fit. Retrieved from http://davidakenny.net/cm/fit.htm

Krueger, R. F., Hicks, B. M., & McGue, M. (2001). Altruism and antisocial behavior: Independent tendencies, unique personality correlates, distinct etiologies. *Psychological Science*, *12*(5), 397-402.

Landis, S. K., Sherman, M. F., Piedmont, R. L., Kirkhart, M. W., Rapp, E. M., & Bike, D. H. (2009). The relation between elevation and self-reported prosocial behavior: Incremental validity over the five-factor model of personality. *The Journal of Positive Psychology*, *4*(1), 71-84.

Lee, K., & Ashton, M. C. (2018). Psychometric properties of the HEXACO-100. *Assessment*, *25*(5), 543-556.

Lilienfeld, S. O. (2018). The multidimensional nature of psychopathy: Five recommendations for research. *Journal of Psychopathology and Behavioral Assessment*, *40*(1), 79-85.

Lin, G. C., Wen, Z., Marsh, H. W., & Lin, H. S. (2010). Structural equation models of latent interactions: Clarification of orthogonalizing and double-mean-centering strategies. *Structural Equation Modeling*, *17*(3), 374-391.

McAuliffe, W.H.B., Forster, D.E., Pedersen, E.J., & McCullough, M.E. (2019). Does cooperation in the laboratory reflect the operation of a broad trait? *European Journal of Personality*, *33*, 89-103.

Miller, J. D., Hyatt, C. S., Maples-Keller, J. L., Carter, N. T., & Lynam, D. R. (2017). Psychopathy and Machiavellianism: A distinction without a difference? *Journal of Personality*, *85*(4), 439-453.

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, Machiavellianism, and psychopathy). *Perspectives on Psychological Science*, *12*(2), 183-204.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods*, *32*(3), 396-402.

Paulhus, D. L., Curtis, S. R., & Jones, D. N. (2018). Aggression as a trait: the Dark Tetrad alternative. *Current Opinion in Psychology*, *19*, 88-92.

Persson, B. N., Kajonius, P. J., & Garcia, D. (2017). Testing construct independence in the Short Dark Triad using Item Response Theory. *Personality and Individual Differences*, *117*, 74-80.

Piliavin, J.A., & Charng, H.-W. (1990). Altruism: A review of recent theory and research. *Annual Review of Sociology*, *16*, 27–65.

Plouffe, R. A., Saklofske, D. H., & Smith, M. M. (2017). The assessment of sadistic personality: Preliminary psychometric evidence for a new measure. *Personality and Individual Differences*, *104*, 166-171.

R Core Team (2014). R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. Available online at: http. *www. R-project. org*.

Revelle, W. (2017). psych: Procedures for Personality and Psychological Research. https://cran.r-project.org/web/packages=psych. R package version 1.7.8.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1-36.

Rushton, J. P., Chrisjohn, R. D., & Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences*, *2*(4), 293-302.

Rushton, J. P., Fulker, D. W., Neale, M. C., Nias, D. K., & Eysenck, H. J. (1986). Altruism and aggression: The heritability of individual differences. *Journal of Personality and Social Psychology*, *50*(6), 1192-1198.

Saucier, G., & Srivastava, S. (2015). What makes a good structural model of personality? Evaluating the Big Five and alternatives. In Cooper, M. L., Larsen, R. (Eds.), *Handbook of Personality and Social Psychology: Vol. 3. Personality processes and individual differences*. Washington, DC: APA Books.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*. Psychology Software Incorporated.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology, 25,* 1-65.

Scott-Phillips, T. C., Dickins, T. E., & West, S. A. (2011). Evolutionary theory and the ultimate–proximate distinction in the human behavioral sciences. *Perspectives on Psychological Science*, *6*(1), 38-47.

Streiner, D. L. (1998). Factors affecting reliability of interpretations of scree plots. *Psychological Reports*, *83*(2), 687-694.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. *The adapted mind: Evolutionary psychology and the generation of culture*, 19-136.

Vachon, D. D., & Lynam, D. R. (2016). Fixing the problem with empathy: Development and validation of the affective and cognitive measure of empathy. *Assessment*, *23*(2), 135-149.

Wilhelm, M. O., & Bekkers, R. (2010). Helping behavior, dispositional empathic concern, and the principle of care. *Social Psychology Quarterly*, *73*(1), 11-32.

Worthington Jr, E. L., Wade, N. G., Hight, T. L., Ripley, J. S., McCullough, M. E., Berry, J. W., ... & O'Connor, L. (2003). The Religious Commitment Inventory--10: Development, refinement, and validation of a brief scale for research and counseling. *Journal of Counseling Psychology*, *50*(1), 84-96.

Table 1. Factor loadings from exploratory factor analysis of pilot study

Numbers of factors extracted	1	2		3		
Factor	1	1	2	1	2	3
1. I have helped a stranger whose car has broken down.	.52	.65	-.12	**.65**	.01	.01
2. I have given directions to a stranger.	.58	.40	.27	.01	.06	**.61**
3. I have made change for a stranger.	.58	.52	.11	.11	-.09	**.62**
4. I have given money to a charity.	.56	.11	.68	-.10	**.49**	.42
5. I have given money to a stranger who needed it.	.58	.56	.06	.18	-.10	**.57**
6. I have donated goods or clothes to a charity.	.51	.00	.76	.03	**.74**	.08
7. I have done volunteer work for a charity.	.42	-.07	.74	.06	**.84**	-.10
8. I have donated blood.	.35	.30	.12	**.38**	.22	-.08
9. I have helped carry a stranger's belongings.	.64	.60	.11	**.41**	.11	.29
10. I have delayed an elevator and held the door open for a stranger.	.43	.12	.44	-.20	.22	**.52**
11. I have allowed someone to go ahead of me in a lineup (e.g. in the supermarket)	.51	.42	.14	.07	-.03	**.53**
12. I have helped an acquaintance to move households.	.53	.62	-.07	**.57**	.04	.07
13. I have pointed out a clerk's error (e.g., in a bank, at the supermarket) in undercharging me for an item.	.49	.39	.16	**.27**	.14	.22

14. I have let a neighbor whom I didn't know too well borrow an item of some value to me.	.48	.56	-.07	**.58**	.07	-.02
15. I have bought 'charity" greeting cards deliberately because I knew it was a good cause.	.59	.43	.26	**.40**	.30	.11
16. I have helped a classmate who I did not know that well with a homework assignment when my knowledge was greater than his or hers.	.50	.24	.39	-.01	.26	**.42**
17. I have before being asked, voluntarily looked after a neighbor's pets or children without being paid for it.	.54	.52	.06	**.37**	.07	.23
18. I have offered to help a handicapped or elderly stranger across a street.	.63	.65	.03	**.41**	.00	.35
19. I have offered my seat on a bus or train to a stranger who was standing.	.63	.59	.10	.28	.00	**.45**
20. I have given a stranger a lift in my car.	.39	.59	-.23	**.61**	-.08	-.04

Table 2. Factor loadings from exploratory factor analysis of main study

Numbers of factors extracted	1	2		3		
Factor	1	1	2	1	2	3
1. I have helped a stranger whose car has broken down.	.61	.63	.04	**.63**	-.02	.08
2. I have given directions to a stranger.	.55	.11	.56	.17	.08	**.55**
3. I have made change for a stranger.	.55	.50	.11	**.53**	-.07	.18
4. I have given money to a charity.	.48	.07	.52	-.05	**.65**	.11
5. I have given money to a stranger who needed it.	.58	.44	.21	**.43**	.14	.14
6. I have donated goods or clothes to a charity.	.51	.01	.63	-.04	**.55**	.27
7. I have done volunteer work for a charity.	.49	.30	.27	.20	**.54**	-.11
8. I have donated blood.	.37	.34	.07	**.30**	.17	-.04
9. I have helped carry a stranger's belongings.	.58	.49	.15	**.55**	-.12	.26
10. I have delayed an elevator and held the door open for a stranger.	.48	-.07	.69	.00	.12	**.66**
11. I have allowed someone to go ahead of me in a lineup (e.g. in the supermarket)	.53	.16	.48	.20	.13	**.41**
12. I have helped an acquaintance to move households.	.57	.50	.13	**.49**	.09	.09
13. I have pointed out a clerk's error (e.g., in a bank, at the supermarket) in undercharging me for an item.	.53	.41	.19	**.38**	.19	.07

14. I have let a neighbor whom I didn't know too well borrow an item of some value to me.	.55	.70	-.12	**.66**	.05	-.13
15. I have bought 'charity" greeting cards deliberately because I knew it was a good cause.	.52	.47	.10	**.40**	.33	-.13
16. I have helped a classmate who I did not know that well with a homework assignment when my knowledge was greater than his or hers.	.48	.30	.25	**.31**	.12	.18
17. I have before being asked, voluntarily looked after a neighbor's pets or children without being paid for it.	.51	.42	.14	**.37**	.31	-.07
18. I have offered to help a handicapped or elderly stranger across a street.	.64	.54	.17	**.54**	.08	.14
19. I have offered my seat on a bus or train to a stranger who was standing.	.57	.31	.35	**.34**	.07	.32
20. I have given a stranger a lift in my car.	.42	.67	-.23	**.63**	.05	-.16

Table 3. Regressions of Malevolence on Altruism

	b	SE	Z	p	r
Psychopathy					
SRA	-.05	.04	-1.17	.240	-.05
Sadism					
SRA	-.07	.04	-1.87	.062	-.07

Model Fit: $\chi^2(662) = 2095.08$, $p < .001$; SRMR = .068; RMSEA = .052, 90% CI [.049, .054]; TLI = .840

	b	SE	Z	p	r
Psychopathy					
SRA1	.67	.12	5.82	<.001	.60
SRA2	-.33	.10	-3.25	.001	-.31
SRA3	-.66	.17	-3.89	<.001	-.49
Sadism					
SRA1	.71	.11	6.24	<.001	.61
SRA2	-.26	.10	-2.57	.010	-.24
SRA3	-.84	.18	-4.76	<.001	-.60

Model Fit: χ^2 (655) = 1775.27, p < .001; SRMR = .050; RMSEA = .046, 90% CI [.044, .048]; TLI = .874

	b	SE	Z	p	r
Psychopathy					
HEXACO	-1.02	.14	-7.26	<.001	-.77
Sadism					
HEXACO	-.70	.10	-7.31	<.001	-.72

Model Fit: χ^2 (206) = 746.59., $p < .001$; SRMR = .052; RMSEA = .057, 90% CI [.053, .060]; TLI = .878

	b	SE	Z	p	r
Psychopathy					
NEO	-.88	.11	-7.73	<.001	-.71
Sadism					
NEO	-.86	.11	-7.74	<.001	-.67

Model Fit: χ^2 (347) = 1410.52., $p < .001$; SRMR = .066; RMSEA = .061, 90% CI [.059, .064]; TLI = .834

Note: SRA = one-factor self-report altruism scale. SRA1 = Factor 1 of three-factor self-report altruism scale. SRA2 = Factor 2 of three-factor self-report altruism scale. SRA3 = Factor 3 of three-factor self-report altruism scale. HEXACO = altruism facet scale of the HEXACO inventory. NEO = altruism facet scale of the NEO inventory.

Table 4. Regressions of Number of Sex Partners on Altruism

	b	SE	Z	p	r
Lifetime					
SRA	.39	.07	5.35	<.001	.22
Casual					
SRA	.32	.08	4.02	<.001	.17
Model Fit: $\chi 2$ (208) = 936.81, *p* < .001; SRMR = .057; RMSEA = .066, 90% CI [.062, .070]; TLI = .829					
Lifetime					
SRA1	.57	.13	4.47	<.001	.33
SRA2	-.26	.16	-1.69	.091	-.16
SRA3	.04	.22	.16	.874	.02
Casual					
SRA1	.69	.15	4.67	<.001	.37
SRA2	-.41	.15	-2.71	.007	-.23
SRA3	-.06	.23	-.26	.792	-.03
Model Fit: $\chi 2$ (201) = 689.13, *p* < .001; SRMR = .050; RMSEA = .055, 90% CI [.050, .059]; TLI = .882					
Lifetime					
HEXACO	-.04	.07	-.51	.608	-.02
Casual					
HEXACO	-.15	.08	-1.94	.052	-.09
Model Fit: $\chi 2$ (8) = 60.77., *p* < .001; SRMR = .046; RMSEA = .090, 90% CI [.071, .110]; TLI = .885					
Lifetime					
NEO	-.08	.07	-1.04	.295	-.04
Casual					
NEO	-.22	.08	-2.61	.009	-.11
Model Fit: $\chi 2$ (53) = 549.06, *p* < .001; SRMR = .072; RMSEA = .107, 90% CI [.101, .114]; TLI = .767					

Note: SRA = one-factor self-report altruism scale. SRA1 = Factor 1 of three-factor self-report altruism scale. SRA2 = Factor 2 of three-factor self-report altruism scale. SRA3 = Factor 3 of three-factor self-report altruism scale. HEXACO = altruism facet scale of the HEXACO inventory. NEO = altruism facet scale of the NEO inventory.
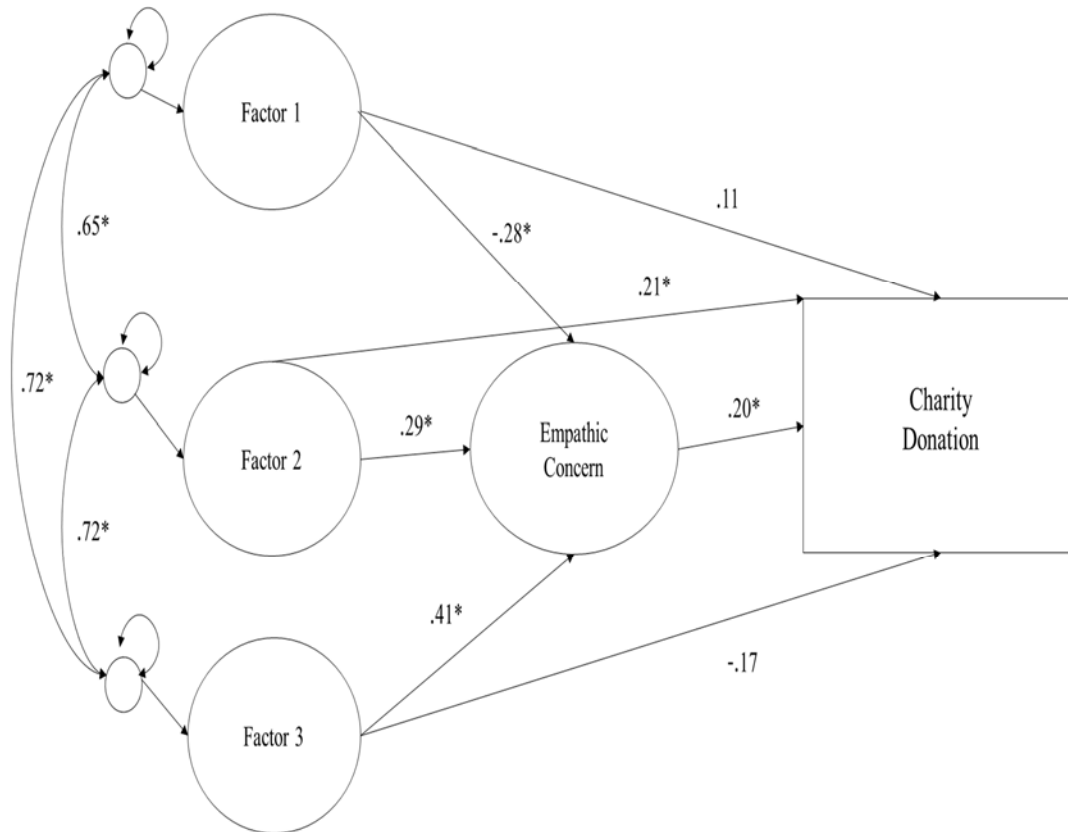
Figure 1. Direct and indirect effects via empathic concern of Self-Report Altruism scale factors on charitable donations. Notes: Coefficients are standardized; * = p <.05. Charitable giving is logit-transformed.