4-2018

Development and Validation of the Single Item Trait Empathy Scale (SITES)

Sara Konrath
Indiana University

Brian P. Meier
Gettysburg College

Brad J. Bushman
Ohio State University - Main Campus

Konrath, Sara, Brian P. Meier, and Brad J. Bushman. "Development and validation of the single item trait empathy scale (SITES)." Journal of Research in Personality 73 (2018): 111-122.

Development and Validation of the Single Item Trait Empathy Scale (SITES)

Abstract
Empathy involves feeling compassion for others and imagining how they feel. In this article, we develop and validate the Single Item Trait Empathy Scale (SITES), which contains only one item that takes seconds to complete. In seven studies (N = 5724), the SITES was found to be both reliable and valid. It correlated in expected ways with a wide variety of intrapersonal outcomes. For example, it is negatively correlated with narcissism, depression, anxiety, and alexithymia. In contrast, it is positively correlated with other measures of empathy, self-esteem, subjective well-being, and agreeableness. The SITES also correlates with a wide variety of interpersonal outcomes, especially compassion for others and helping others. The SITES is recommended in situations when time or question quantity is constrained.

Keywords
Empathy, Comparrison, Short measure, Single item scale, Scale validation, Prosocial behavior

Disciplines
Personality and Social Contexts | Psychology | Social Psychology

Comments
Original version available from the publisher at: https://www.sciencedirect.com/science/article/pii/S0092656617301137

Full Length Article

Development and validation of the Single Item Trait Empathy Scale (SITES)

Sara Konrath, Brian P. Meier, Brad J. Bushman

PII: S0092-6566(17)30113-7
DOI: https://doi.org/10.1016/j.jrp.2017.11.009
Reference: YJRPE 3684

To appear in: *Journal of Research in Personality*

Received Date: 14 July 2017
Revised Date: 6 November 2017
Accepted Date: 13 November 2017

Please cite this article as: Konrath, S., Meier, B.P., Bushman, B.J., Development and validation of the Single Item Trait Empathy Scale (SITES), *Journal of Research in Personality* (2017), doi: https://doi.org/10.1016/j.jrp.2017.11.009

Development and validation of the Single Item Trait Empathy Scale (SITES)

Sara Konrath [1,2,3]

Brian P. Meier [4]

Brad J. Bushman [5]

Author affiliations:

1. Lilly Family School of Philanthropy, Indiana University, Indianapolis, IN, USA

2. Institute for Social Research, University of Michigan, Ann Arbor, MI, USA

3. Department of Psychiatry, University of Rochester Medical School, Rochester, NY, USA

4. Department of Psychology, Gettysburg College, Gettysburg, PA, USA. Email: bmeier@gettysburg.edu

5. School of Communication & Department of Psychology, The Ohio State University, Columbus, OH, USA. Email: bushman.20@osu.edu

Correspondence should be addressed to Sara Konrath, Lilly Family School of Philanthropy, 301 University Boulevard, Suite 3000, Indianapolis, IN 46202. Email: skonrath@iu.edu. Phone: 317-278-8971. We thank the John Templeton Foundation (grants #47993 and 57942 to Konrath and Bushman), and the National Institutes for Health (1R21-HD-073549-01A1: NICHD and 1R01-CA-180015: NCI to Konrath) for support while writing this manuscript. All authors contributed to study design, data collection, and manuscript writing. Data analyses were primarily conducted by the first author. Note that we did not preregister any of these studies.

ABSTRACT

Empathy involves feeling compassion for others and imagining how they feel. In this article, we develop and validate the Single Item Trait Empathy Scale (SITES), which contains only one item that takes seconds to complete. In seven studies (N=5,724), the SITES was found to be both reliable and valid. It correlated in expected ways with a wide variety of intrapersonal outcomes. For example, it is negatively correlated with narcissism, depression, anxiety, and alexithymia. In contrast, it is positively correlated with other measures of empathy, self-esteem, subjective well-being, and agreeableness. The SITES also correlates with a wide variety of interpersonal outcomes, especially compassion for others and helping others. The SITES is recommended in situations when time or question quantity is constrained.

Abstract word count: 120

Key words: empathy; compassion; short measure; single item scale; scale validation; prosocial behavior

INTRODUCTION

"No one cares how much you know, until they know how much you care"

~Theodore Roosevelt, former US President

Empathy involves imagining others' perspectives and feeling care and concern for them (Davis, 1983). It is especially important when it comes to promoting prosocial behaviors, such as helping, cooperating, and sharing (Batson, 2011; Konrath & Grynberg, 2016). More empathic people also tend to be less self-focused, for example, they score lower on narcissism (Hepper, Hart, & Sedikides, 2014; Watson, Grisham, Trotter, & Biderman, 1984). This paper describes the development and validation of the Single Item Trait Empathy scale (SITES), which consists of this single item: *To what extent does the following statement describe you: "I am an empathetic person,"* rated using a scale that ranges from 1=*Not very true of me* to 5=*Very true of me*. Although caution should be taken when using short scales, this scale may be useful in situations when time or question quantity is constrained. After first reviewing current measures of empathy, we next give an overview of the SITES and its development.

Measuring empathy

Researchers are increasingly interested in collecting data from a lot of people in a short amount of time (i.e., data collection using crowdsourcing, mobile phones, or social media). In order to facilitate such data collection, and to better understand the trends mentioned above, scholars previously developed the Single Item Narcissism Scale (SINS) as a "quick and dirty" measure of narcissism (Konrath, Meier, & Bushman, 2014). The SINS was found to be a reliable and valid way to measure narcissism when a quick method is necessary (also see van der Linden

& Rosenthal, 2016 for a replication and extension of this work). In this article, we propose and validate a similar short measure of empathy called the Single Item Trait Empathy Scale (SITES).

Before describing our scale, it is useful to distinguish between trait and state empathy, in the same way researchers have distinguished between trait and state anger (Speilberger, 1983) and between trait and state anxiety (Spielberger, Sydeman, Owen, & Marsh, 1999). Personality traits are relatively stable over time and across situations. In contrast, emotional states are more fleeting and situationally determined. The SITES is a trait measure of empathy. In other words, it measures the extent to which people tend to vary in their empathic responses to others across situations. In this article, we show that the SITES is relatively stable over time and situations.

The SITES only takes seconds to complete. We offer it as an alternative to much longer trait measures of empathy. The most popular measure of trait empathy is the Interpersonal Reactivity Index (IRI), which consists of 28 items that are evenly divided into four subscales, with seven items per subscale (Davis, 1983). *Fantasy* assesses the extent to which individuals identify with fictional characters (e.g., "When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were happening to me"). *Perspective-taking* is a cognitive form of empathy, assessing the extent to which individuals spontaneously adopt another person's point of view (e.g., "Before criticizing somebody, I try to imagine how I would feel if I were in their place"). *Empathic concern* is an emotional form of empathy, assessing the extent to which individuals experience feelings of care and compassion for others (e.g., "When I see someone being taken advantage of, I feel kind of protective towards them"). *Personal distress* is a more self-focused emotional response to others, assessing the extent to which individuals experience discomfort and anxiety in response to others' suffering (e.g., "When I see someone who badly needs help in an emergency, I go to pieces"). For researchers who can afford

to use a longer scale, and who are interested in these different dimensions of empathy, the IRI works very well. However, the SITES might be useful when researchers cannot afford to use the 28-item IRI, such as when the number of items in a large data-collection session are limited, and when researchers are interested in a single-dimension measure of empathy.

Overview and Scale Development

Our approach to validating the SITES was to demonstrate its correlation with another widely used empathy scale (i.e., the IRI), examine its test-retest reliability, and then provide correlations with a number of theoretically relevant intrapersonal and interpersonal outcomes. Across seven studies, using several different participant populations (total N=5,724) and several different outcome measures, we present evidence for the SITES' discriminant validity, convergent validity, predictive validity, and test-retest reliability. We further divided the convergent and predictive validity outcomes into ones that are more intrapersonal (i.e. having implications for the self) versus interpersonal (i.e. having implications for others). This will help researchers to quickly determine whether this scale is relevant for their interests. (See Table 1 for scale descriptive statistics and Table 2 for an overview of the results.)

We chose the wording of the SITES carefully, aiming to create a face valid and easily understood measure of empathy that followed closely with scholars' development of the SINS (Konrath, Meier, & Bushman, 2014) and other single-item scales (e.g., the single-item self-esteem measure;(Robins, Hendin, & Trzesniewski, 2001).

We examined readability statistics of the item at the following website: http://www.readability-score.com. The SITES has Flesch-Kincaid grade level of 9.3 compared to

a grade level of 7.9 for the IRI.[1] Thus, although not all respondents will likely fully understand the rich theoretical connotations of the term "empathetic," the readability data and popular use of the term "empathy" among the general public suggests that most adult respondents will be able to understand the meaning of this term. Researchers could, however, modify the SITES by including a definition of empathic, such as "(Note: An empathetic person understands others' feelings, and experiences care and concern for them.) See Study 3.

METHOD

All studies were conducted after being approved by the Institutional Review Boards of the authors' three universities. We documented informed consent in writing for the in-person studies, however, the IRBs waived the requirement for written signatures for online studies, in which participants consented by clicking on a button. Data are available upon request to the first author. Researchers who wish to conduct secondary analyses on de-identified datasets will receive them upon presenting evidence of IRB approval and signing a data confidentiality agreement. (See Supplementary Tables for all means, standard, deviations, Cronbach's alphas, correlations, and regression output.) All participants were at least 18 years old.

Power Analysis

We used correlation analyses to examine the relevant effects of interest in each study. Subsequent analyses in some studies used ANOVAs or regression, but our main hypothesis in each study required a correlation analysis. We therefore based our power analyses on correlations. In order to calculate the sample size needed to achieve 80% power, we needed to find an estimated effect size or estimated *r*. We first examined the IRI scale development data (Davis, 1983). Davis examined convergent validity of the four subscales of the IRI with multiple

[1] The readability level can be moved down to grade 7.6 by making one small change. To what extent does *this* statement describe you: "I am an empathetic person." In the current paper, we used the original version of the scale with adults, but recommend that researchers use this revised version for adolescent populations.

variables including interpersonal functioning, self-esteem, emotionality, and sensitivity to others. He presented 120 *r* values that ranged from .00 to .59 with a mean of .19. Variables of this type are included in our studies and therefore we believe this estimate for *r* is valid. Yet, we also examined recent datasets that focused on trait empathy using the IRI and variables more specific to our studies such as agreeableness and prosocial behavior (Melchers et al., 2016; Mooradian, Davis, & Matzler, 2011; Sze, Gyurak, Goodkind, & Levenson, 2012). These additional datasets yielded 28 *r* values ranging from .00 to .63 with a mean of .23. These additional *r* values are similar to those provided by Davis (1983) and we therefore averaged all of the *r* values (148), which yielded an *r* of .20.

We conducted a power analysis using G Power (http://www.gpower.hhu.de/en.html) for correlations using an effect size estimate of $r = .20$ and a preferred power of 80%. These parameters resulted in a required sample size of 193 participants. The number of participants in each of our studies differs for various reasons (e.g., larger purpose of the study, size of the participant pool, resources available, time constraints, etc.). However, all of our studies but one included more than 193 participants. Due to financial and human resource constraints in running a complex 3-wave study, Study 5 included only 87 participants, which provides an achieved power of only 59%. Therefore, Study 5 was underpowered compared to the remaining studies.

STUDY 1

In Study 1 we aimed to examine the relationship between SITES and the longer, more established Interpersonal Reactivity Index (IRI, Davis, 1983). We expected the SITES to be positively correlated with both the emotional (i.e. empathic concern) and cognitive (i.e. perspective taking) IRI subscales.

Participants

Participants were 3,289 adults recruited online via MTURK for a small payment. They were 47.5% male, with a mean age of 46.1 (*SD*=16.7) years.

Procedures

Participants completed the SITES as part of a larger online study related to media and health behaviors. They also completed the Interpersonal Reactivity Index (IRI), which is a 28-item, widely used, multidimensional, measure of trait empathy with four 7-item subscales that were described above (Davis, 1983).

Results and Discussion

Participants' average SITES score was 4.02 (*SD*=1.09), and participants average IRI scores were as follows: empathic concern (*M*=3.80, *SD*=0.70, α=.81), perspective taking (*M*=3.51, *SD*=0.64, α=.74), fantasy (*M*=3.17, *SD*=0.79, α=.78), and personal distress (*M*=2.56, *SD*=0.53, α=.81). Empathic concern (emotional empathy) and perspective taking (cognitive empathy) are the most prototypical empathy subscales. As can be seen, participants' average scores on these two key subscales are above the midpoint on the 1-5 scale, just like the SITES.

Men had lower SITES scores (*M*=3.76, *SD*=1.14) than women (*M*=4.27, *SD*=0.97), $F(1,3204)=182.03$, $p<.001$, $d=0.48$. Gender differences for the prototypical IRI subscales were as follows: Men (*M*=3.57, *SD*=.68) scored lower than women (*M*=4.01, *SD*=0.68) in *empathic concern*, $F(1,3203)=363.47$, $p<.001$, $d=0.65$, and *perspective taking* (Men: *M*=3.40, *SD*=.64; Women: *M*=3.62, *SD*=0.62; $F(1,3203)=91.40$, $p<.001$, $d=0.35$). Overall, both the SITES and the IRI prototypical empathy subscales showed expected gender differences.

As expected, participants who scored higher on the SITES scored higher on the IRI empathic concern, $r(3282)=.50$, $p<.001$, and perspective taking, $r(3281)=.38$, $p<.001$ subscales. They also scored higher on the fantasy subscale, $r(3282)=.27$, $p<.001$, and lower on the personal

distress subscale, $r(3282)=-.08$, $p<.001$. Overall, the SITES was more highly correlated with empathic concern than perspective taking ($z=5.79$, $p<.001$;(Lee & Preacher, 2013), suggesting that the measure more strongly taps into emotional aspects of empathy, even though it also taps into cognitive aspects of empathy to some extent. In addition, participants who scored higher on the SITES scored slightly lower on the more self-oriented type of emotional response to others' distress. However, some caution is warranted considering the small effect size. Table 2 summarizes the relationships between the SITES and the measures collected in all studies.

We also tested the quadratic relationship between each of the IRI subscales and the SITES. Step 1 of each regression include the IRI subscale, and Step 2 included the same subscale squared. As can be seen from the Supplementary Tables, there was a significant change in the R^2 when adding the quadratic term in Step 2 for empathic concern, perspective taking, and fantasy, suggesting that the relationship between these subscales and the SITES was not strictly linear (the average change in R^2 ranged from 0.3% to 1.9%). There was no significant quadratic relationship for the personal distress subscale. See Supplementary Tables for IRI average values for each endpoint on the SITES.

STUDIES 2A & 2B

Studies 2A and 2B examined the relationship between the SITES and the Big Five personality traits, using two different measures of personality. Prior research has found that empathy is most strongly and consistently correlated with the Big Five factor of Agreeableness (Mooradian, Davis, & Matzler, 2011). Thus, we expected that participants who scored higher on the SITES to score higher on Agreeableness in both studies.

Participants

Participants in *Study 2A* were 250 Gettysburg College students, of whom 242 completed all measures. The final sample was 49.6% male, with a mean age of 19.3 (*SD*=0.9) years. Participants in *Study 2B* were a combined sample of 308 online adults and University of Michigan students, of which 307 completed all measures. The final sample was 44.2% male, with a mean age of 37.6 (*SD*=19.1) years.

Procedures

Participants completed the SITES and a measure of the Big Five personality traits as part of a questionnaire battery within larger studies. The personality scale used in *Study 2A* was the Mini International Personality Item Pool (IPIP), which is a 20-item scale shortened from the longer IPIP (Donnellan, Oswald, Baird, & Lucas, 2006). It includes 4 items from each of the five factors of Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (OCEAN). In *Study 2B* we used the Ten Item Personality Inventory (TIPI), which is a 10 item scale including 2 items from each of the five factors (Gosling, Rentfrow, & Swann Jr, 2003).

Results and Discussion

Participants' average SITES score was 3.78 (*SD*=0.96) in *Study 2A* and 4.17 (*SD*=0.90) in *Study 2B*. In *Study 2A* men scored lower on empathy (*M*=3.60, *SD*=1.01) than women did (*M*=4.01, *SD*=0.85), $F(1,240)=11.64$, $p=.001$ $d=0.38$. A similar pattern was found in *Study 2B* (men: *M*=4.05, *SD*=0.86; women: *M*=4.27, *SD*=0.93, $F(1,304)=4.25$, $p=.04$, $d=0.25$).

In *Study 2A*, participants who scored higher on the SITES scored higher on Agreeableness, $r(242)=.47$, $p<.001$ and Openness, $r(242)=.21$, $p=.001$. The SITES was unrelated to Conscientious, Extraversion, and Neuroticism, rs<.06, ps>.41.

In *Study 2B*, participants who scored higher on the SITES scored higher on Agreeableness, $r(307)=.12$, $p=.03$, and lower on Conscientiousness, $r(307)=-.15$, $p=.008$. The SITES was unrelated to openness, extraversion, and neuroticism, $rs<.08$, $ps>.17$.

As expected, both studies found that higher scores on the SITES were associated with higher Agreeableness scores. The inconsistencies with the other traits and relatively different degrees of association may be due to differences in the two personality inventories or participant populations. The important point is that both studies were consistent in showing that the SITES correlated positively with Agreeableness even when it was measured using two different scales.

STUDY 3

In Study 3, we examined the relationship between the SITES and self-reported environmental behaviors. Prior research has found that people who score higher on trait empathy, or who are induced to be in a more empathic state, have more pro-environmental attitudes and behaviors (Berenguer, 2007; Milfont & Sibley, 2016; Pfattheicher, Sassenrath, & Schindler, 2015; Schultz, 2002). Thus, we expected that people scoring high on the SITES would be more likely to report pro-environmental behaviors and would be more likely to identify as an environmentalist. Participants also completed a short narcissism measure to examine whether the SITES would be negatively correlated with narcissism, as in prior research (Hepper et al., 2014; Watson et al., 1984).

Participants

Participants were a nationally representative sample of 510 adults recruited from the market research company GfK for a small payment. Selection was made based on 2010 US Census data. Participants were randomly recruited through probability-based sampling, and households were provided with access to the Internet and hardware if needed. See recruitment and sampling

survey methodology details here: http://www.gfk.com/fileadmin/user_upload/dyna_content/US/documents/KnowledgePanel_Recruitment_Sample_Survey_Methodology.pdf. Three participants did not respond to the SITES, leaving a final sample of 507 (49.6% male; Mean age=50.1, *SD*=16.6).

Procedure

Participants completed the SITES[2] as part of a larger online study. Participants also completed the SINS (Konrath, Meier, & Bushman, 2014) described earlier (i.e., To what extent do you agree with this statement, "I am a narcissist"; 1=*not very true of me*; 7=*very true of me*). Pro-environmental behaviors were assessed by summing participants' responses to 10 questions about whether they had engaged in environmental behaviors in the past year (e.g. *Given money to an environmental organization; Recycled your newspapers or other papers; Taken steps to reduce your use of energy*; 1=yes, 0=no). Participants were also asked, *Would you describe yourself as an environmentalist?* (*Yes, definitely*; *Yes, somewhat*; or *No*). Participants also reported their political party identification, which we included because of its probable link to pro-environmental attitudes and behaviors (1=*Strong Republican*, 4=*Undecided/Independent*, 7=*Strong Democrat*).

Results and Discussion

We used GfK's post-stratification weight for the US general adult population in all analyses. Participants' average SITES score was 3.76 (0.98), and men scored lower on empathy (*M*=3.56, *SD*=0.94) than women did (*M*=3.98, *SD*=0.96), $F(1,505)=26.99$, $p<.001$, $d=0.44$. Participants who scored higher on the SITES scored lower on narcissism, $r(501)=-.13$, $p=.003$. In addition, as in prior research, higher SITES scores were associated with reporting a greater

[2] Study 3's version of the scale also included this definition, which we dropped for all other studies because we thought it was not needed: "*Note: An empathetic person understands others' feelings, and experiences care and concern for them.*"

number of pro-environmental behaviors, β=0.18, p<.001, even when controlling for political party identification β=0.17, p<.001.

We explicitly tested the predicted linear effect of environmental identification on the SITES using a one-way ANOVA, by coding the ordinal categories of "yes, definitely"; "yes, somewhat"; and "no" as 1, 0, and -1. As per a reviewer's suggestion, we also tested the quadratic effect in order to rule out this pattern of results (coding the categories 1, -2, and 1). Participants who described themselves as definitely an environmentalist scored higher on the SITES (M=4.23, SD=0.82, n=42) than those who described themselves as somewhat this way (M=3.80, SD=0.95, n=227), who in turn scored higher than those who did not self-identify as an environmentalist (M=3.65, SD=0.98, n=228), F(2,495)=6.63, p=.001. The hypothesized linear pattern was significant, t(493)=3.57, p<.001, and as expected, the quadratic effect was not significant, t(493)=1.34, p=.18. Taken together, empathy as assessed by the SITES was associated with more pro-environmental behaviors and more self-identification as an environmentalist. These effects remained even when controlling for political party affiliation, suggesting that they are independent of it.

STUDY 4

The main purpose of Study 4 was to examine the stability (i.e. test-retest reliability) of the SITES. In addition, we examined how the SITES was related to measures of narcissism, measures of psychological health and well-being, and rewarding behaviors. Consistent with prior research, we expected that participants who scored higher on the SITES would score lower on narcissism, as in Study 3 and in previous research (Hepper et al., 2014; Watson et al., 1984), higher on measures of psychological health and well-being (Cusi, MacQueen, Spreng, &

McKinnon, 2011; Davis, 1983; Suzuki & Kino, 2014; Wei, Liao, Ku, & Shaffer, 2011), and would score higher on social rewards motivation than on nonsocial rewards motivation.

Participants

Time 1 (baseline) participants were 834 adults recruited by Qualtrics Panels for a small payment who completed a battery of tests twice. They were 31.1% male, with a mean age of 44.7 (*SD*=15.3) years. Time 2 (follow up) participants (N=367) completed an identical survey approximately two weeks later.

Procedure

Participants were invited to join a 2-part study that involved completing two identical short surveys at two separate time points about two weeks apart.

Due to time constraints of the online survey, we only included one of the four subscales of the Interpersonal Reactivity Index (Davis, 1983), namely *empathic concern* (α=.82; see Study 1). This subscale showed the highest correlation with the SITES (See Study 1), and we wanted to replicate that finding. As in Study 3, participants also completed the Single Item Narcissism Scale (Konrath, Meier, et al., 2014).

Self-esteem was assessed using the 10-item Rosenberg Self-Esteem Scale (Rosenberg, 1965). A sample item is "*I feel that I have a number of good qualities*" (1=*Strongly Disagree*, 4=*Strongly Agree*; α=.89). *Life satisfaction* was measured with the 5-item Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin, 1985). A sample item is "*I am satisfied with my life*" (1=*Strongly Disagree,* 7=*Strongly Agree*; α=.92).

Frequency of *depressive symptoms* in the past week (1=*Rarely or none of the time*; 4=*Most or all of the time*; α=.93) such as sadness, loneliness, feelings of worthlessness were measured with the 20-item Center for Epidemiologic Studies Depression Scale (Radloff, 1977).

Anxiety was assessed using the 20-item State-Trait Anxiety Index, Form Y1 (Spielberger & Sydeman, 1994). A sample item is: "*In the past week, I have felt jittery*" (1=*almost never*, 4=*almost always*; α=.95).

A modified version of the Sensitivity to Reinforcement of Addictive and Other Primary Rewards scale was used for this study (Goldstein et al., 2010). We used a version that added preferences for self-esteem boosts and for helping others (Bushman, Moeller, Konrath, & Crocker, 2012). Participants were asked to think about *how pleasant* (i.e. liking) and *how much they wanted* (i.e. wanting) a number of *social* (i.e. seeing their best friend, helping others, their favorite sexual activity; α=.54) and *non-social* rewards (i.e. their favorite food, their favorite alcoholic beverage, a self-esteem boost, receiving their paycheck; α=.56; 1=*not at all*, 5=*extremely*), presented in random order. For each of the 7 rewards, we calculated the average of participants' liking and wanting scores.

Results and Discussion

Participants' average SITES score was 3.86 (*SD*=1.02) at Time 1 and 3.89 (*SD*=0.99) at Time 2. In comparison, participants' average IRI empathic concern score was 3.95 (*SD*=.70) at Time 1 and 3.88 (*SD*=0.72) at Time 2. These are again both above the midpoint on the 1-5 scales. Participants who scored higher on the SITES scored higher on empathic concern, $r(834)=.45$, $p<.001$. We also tested the quadratic relationship between empathic concern and the SITES, using the same procedure as in Study 1. As can be seen from the Supplementary Tables, for Wave 1 there was no significant change in the R^2 when adding the quadratic term in Step 2, but for Wave 2, there was significant change (of 5.7%), suggesting that the relationship between these subscales and the SITES was linear at Wave 1 but quadratic at Wave 2. See Supplementary

Tables for IRI average values for each endpoint on the SITES. Dropout status was unrelated to participants' SITES or empathic concern score.

Men scored lower on the SITES (M=3.50, SD=1.02) than women (M=4.01, SD=0.98), $F(1,791)=45.31$, $p<.001$, d=0.52 (based on Time 1 data). Men also scored lower on the IRI empathic concern subscale (M=3.62, SD=.65) than women (M=4.09, SD=0.67), $F(1,787)=85.10$, $p<.001$, d=0.71. Thus, there are expected gender differences, regardless of scale used.

Test-retest reliability of the SITES was assessed by computing correlations between Time 1 and Time 2 scores (approximately two weeks later, M=12.9 days, SD=2.6). The SITES test-retest correlation was $r(344)=.57$, $p<.001$, and the empathic concern one was $r(351)=.80$, $p<.001$. Although the IRI has better test-retest reliability, the SITES has reasonable consistency across time (see also Study 5).

We next examined the correlation between both measures of empathy and the other measures at Time 1. In terms of the SITES, those who scored higher on it also reported higher self-esteem, $r(834)=.13$, $p<.001$, and life satisfaction, $r(823)=.10$, $p=.003$, and lower narcissism, $r(825)=-.08$, $p=.02$, depressive symptoms, $r(824)=-.07$, $p=.04$, and anxiety, $r(819)=-.08$, $p=.03$. In comparison, those who scored higher on empathic concern also reported higher self-esteem, $r(841)=.25$, $p<.001$, and lower narcissism, $r(831)=-.33$, $p<.001$, depressive symptoms, $r(831)=-.10$, $p=.01$, and anxiety, $r(826)=-.07$, $p=.04$. However, there was no correlation between empathic concern and life satisfaction, $r(830)=.01$, $p=.71$. Taken together, with the exception of life satisfaction, both empathy assessments show similar directions of association with the other measures.

In terms of rewards that were most liked and wanted by those scoring higher on the SITES, these included (in order of magnitude): helping others, $r(833)=.34$, $p<.001$, seeing their

best friend, $r(831)=.26$, $p<.001$, self-esteem boosts, $r(832)=.23$, $p<.001$, eating their favorite food, $r(832)=.21$, $p<.001$, receiving their paycheck, $r(829)=.16$, $p<.001$, and their favorite sexual activity, $r(828)=.08$, $p=.02$. SITES scores were unrelated to being rewarded by their favorite alcoholic beverage, $r(829)=.04$, $p=.31$.

In comparison, the most liked and wanted rewards by those scoring higher on IRI empathic concern, included (in order of magnitude): helping others, $r(840)=.58$, $p<.001$, seeing their best friend, $r(838)=.44$, $p<.001$, self-esteem boosts, $r(839)=.32$, $p<.001$, receiving their paycheck, $r(836)=.26$, $p<.001$, eating their favorite food, $r(839)=.20$, $p<.001$. Empathic concern was unrelated to being rewarded by their favorite sexual activity, $r(835)=.05$, $p=.13$, and was negatively related to being rewarded by their favorite alcoholic beverage, $r(836)=-.08$, $p=.03$.

Although both the SITES and the IRI were both associated with other rewards, the most important motivators for those who scored higher on both measures of empathy were helping others and seeing their best friend.

STUDY 5

Study 4 revealed that the SITES has good test-retest reliability when the scale was completed twice with about two weeks between each measurement. In order to further examine test-retest reliability over a longer time period, we conducted Study 5. In addition, we examined how the SITES was related to self-report measures of empathy, the moral principle of care, adult attachment, psychological well-being, and narcissism. We expected that participants who scored higher on the SITES would be more likely to endorse the moral principle of care, have more positive attachment models of others, score higher on the psychological well-being facet of positive relations with others, and score lower on narcissism (Britton & Fuendeling, 2005; Hepper et al., 2014; Mikulincer et al., 2001; Watson et al., 1984; Wilhelm & Bekkers, 2010).

We also included three objective measures we expected to relate to empathy: (1) an emotion recognition task, (2) an observer rating of empathy, and (3) a helping task. We expected that individuals who scored higher on the SITES would be better at identifying emotions, be rated as more empathic by an observer, and feel more empathy for a distressed person, which should be associated with a greater likelihood of offering to help (Batson, 2011; Konrath, Corneille, Bushman, & Luminet, 2014).

Participants

Time 1 (baseline) participants were 87 college students recruited from the University of Michigan for a payment. They were 40.0% male, with a mean age of 21.0 (*SD*=4.3) years. Time 2 participants (*N*=76) completed the SITES about three weeks later (*M*=19.0 days, *SD*=5.4), and Time 3 participants (*N*=58) completed the SITES about 6 months later (*M*=6.2 months, *SD*=1.5).

Procedures

Study 5 was part of a 3-wave experimental study to examine the effect of empathy-building text messages (compared to control text messages) on participants' empathy and prosocial behavior (at Waves 2 and 3). The results of this experiment were reported elsewhere (Konrath et al., 2015).

In Study 5, we were primarily interested in the SITES's *test-retest reliability*, and as such, we examined correlations between the SITES at Time 1 and the SITES at Times 2 and 3. However, we also took advantage of the baseline survey measures to better understand the properties of the SITES. These measures were administered before participants were randomly assigned to experimental conditions.

Self-report measures included the Davis (1983) Interpersonal Reactivity Index and the Single Item Narcissism Scale described in Studies 3 and 4 (Konrath, Meier, et al., 2014). The

Principle of Care scale (Wilhelm & Bekkers, 2010, 2012) included 8 items assessing the moral importance of helping others. A sample item is: "*People should be willing to help others who are less fortunate*" (1 = *strongly disagree*, 5 = *strongly agree*; α=.67).

Adult attachment was assessed with the Relationship Questionnaire, which asks participants to rate to what extent each of four short paragraphs describing different attachment styles apply to them (1 = *not at all like me*, 5 = *very much like me*: Secure (positive views of self and others), Dismissing (positive self, negative others), Preoccupied (negative self, positive others), Fearful (negative self, negative others;(Bartholomew & Horowitz, 1991). *Positive models of self* were calculated by adding positive self models and subtracting negative self models: (Secure + Dismissing) – (Fearful + Preoccupied). *Positive models of others* were calculated with this formula: (Secure + Preoccupied) – (Fearful + Dismissing).

Psychological well-being was assessed with Ryff's 42-item scale (Ryff, 1989), which has 6 facets: Autonomy (e.g. "*I have confidence in my opinions even if they are contrary to the general consensus*"; α=.68), Mastery (e.g. "*I am quite good at managing the many responsibilities of my daily life*"; α=.49), Growth (e.g. "*I have the sense that I have developed a lot as a person over time*"; α=.75), Positive relations with others (e.g. "*Most people see me as loving and affectionate*"; α=.83), Purpose (e.g. "*I enjoy making plans for the future and working to make them a reality*"; α=.79), and Self-acceptance (e.g. "*In general, I feel confident and positive about myself*"; α=.90; 1 = *strongly disagree*, 6 = *strongly agree*).

We also examined three objective measures that should relate positively to trait empathy: (1) an emotion recognition task, (2) an observer rating of empathy, and (3) a helping task. To assess emotion recognition, participants completed the Facial Action Coding System (FACS)-verified University of California, Davis, Set of Emotion Expressions (UCDSEE; (Tracy, Robins,

& Schriber, 2009)), cropped to standard headshots. This consists of 10 emotional expressions (*anger, contempt, disgust, embarrassment, fear, happiness, pride, sadness, shame, surprise*) posed by two individuals: a White male and a White female (20 photos total; randomly presented). Participants selected the emotion that best corresponded to the face that they saw (scored +1 per correct response). We also asked research assistants to report how empathic participants were after interacting with them in the lab (1=*not very true of him/her*; 7=*very true of him/her*).

Finally, we assessed responses to a distressed person by using the Katie Banks task, which involves listening to a radio program about a woman who just lost her parents in a car crash (Batson et al., 1989; Batson et al., 1991). After listening to the program, participants were asked to what extent they felt distressed for themselves (4 items: e.g. "*directly distressed, as if I personally experienced something bad*"; α=.95; 1=*very slightly*, 5=*extremely*), versus for Katie Banks (4 items: "*distressed for the person being interviewed*"; α=.95). Participants were also asked to what extent they felt five compassionate emotions for Katie Banks (e.g. sympathetic, compassionate). Participants were then asked if they would be willing to volunteer to help Katie Banks, and if so, for how many hours.

Results and Discussion

Participants' average SITES score was 4.16 (*SD*=0.84; *N*=87) at Time 1, 4.14 (*SD*=0.74; *N*=76) at Time 2, and 4.03 (*SD*=0.99; *N*=58) at Time 3. Participants average IRI scores at Time 1 were as follows: empathic concern (*M*=3.92, *SD*=0.61, α=.81), perspective taking (*M*=3.59, *SD*=0.67, α=.81), fantasy (*M*=3.44, *SD*=0.85, α=.83), and personal distress (*M*=2.65, *SD*=0.67, α=.75). The following analyses compare the SITES to the IRI empathic concern (emotional empathy) and perspective taking (cognitive empathy), scales, since they measure prototypical

empathy factors. Dropout status at Time 2 or Time 3 was unrelated to empathy as measured by the SITES or the IRI empathic concern and perspective taking subscales.

There were no gender differences in the SITES (*M men*=4.19, *SD*=0.82; *M women*=4.13, *SD*=0.85), $F(1,87)=0.13$, $p=.72$, $d=0.07$. Similarly, there were no gender differences in empathic concern (*M men*=3.89, *SD*=0.59; *M women*=3.94, *SD*=0.63), $F(1,88)=0.15$, $p=.70$, $d=0.06$, or perspective taking (*M men*=3.67, *SD*=0.70; *M women*=3.53, *SD*=0.65), $F(1,88)=1.05$, $p=.31$, $d=0.10$.

Test-retest reliability was assessed by computing correlations between baseline and Time 2 and 3 scores. Baseline SITES scores were highly correlated with SITES scores at Time 2 (19.0 days later on average), $r(74)=.67$, $p<.001$, and Time 3 (6.2 months later on average), $r(56)=.62$, $p<.001$. Time 2 was also correlated with Time 3, $r(53)=.62$, $p<.001$.

Baseline IRI empathic concern scores were highly correlated with empathic concern scores at Time 2, $r(83)=.80$, $p<.001$, and Time 3, $r(60)=.76$, $p<.001$. Time 2 was also correlated with Time 3, $r(60)=.74$, $p<.001$.Baseline perspective taking scores were highly correlated with perspective taking scores at Time 2, $r(83)=.77$, $p<.001$, and Time 3, $r(60)=.59$, $p<.001$. Time 2 was also correlated with Time 3, $r(60)=.56$, $p<.001$. As in Study 4, the IRI has better test-retest reliability (especially empathic concern), however, the SITES also has reasonable consistency over time.

Self-report measures. All results reported from this point on are based on Time 1 (baseline) data because it had the largest sample size and hence the most reliable results.

Participants who scored higher on the SITES scored higher on empathic concern, $r(87)=.49$, $p<.001$, perspective taking, $r(87)=.37$, $p<.001$, and the principle of care, $r(87)=.30$, $p=.005$, and lower on narcissism, $r(87)=-.32$, $p=.002$. The SITES was unrelated to the IRI

subscales personal distress, $r(87)=.09$, $p=.43$, or fantasy, $r(87)=.03$, $p=.77$. In terms of adult attachment, the SITES was associated with having more positive models of others, $r(86)=.44$, $p<.001$, but was unrelated to positive models of self, $r(86)=.08$, $p=.46$. In terms of psychological well-being, the SITES was positively related to two facets: growth, $r(81)=.29$, $p=.01$, and positive relations with others, $r(81)=.28$, $p=.01$, but was unrelated to the others: autonomy, $r(81)=-.06$, $p=.61$; mastery, $r(81)=.11$, $p=.34$; purpose, $r(81)=.05$, $p=.66$; self-acceptance, $r(81)=-.03$, $p=.83$. We also tested the quadratic relationship between the IRI subscales and the SITES, using the same procedure as in Study 1 and 4. As can be seen from the Supplementary Tables, there was no significant change in the R^2 when adding the quadratic term in Step 2 for any of the subscales.

IRI empathic concern was similarly positively correlated with the principle of care, $r(89)=.49$, $p<.001$, and positive models of others, $r(87)=.36$, $p=.001$, but negatively correlated with narcissism, $r(89)=-.50$, $p<.001$. It was unrelated to positive models of self, $r(87)=.06$, $p=.56$. In terms of psychological well-being, empathic concern was positively related with growth, $r(84)=.41$, $p<.001$, positive relations with others, $r(84)=.39$, $p<.001$, purpose, $r(84)=.27$, $p=.01$, and self-acceptance, $r(84)=.28$, $p=.01$, but was unrelated to autonomy, $r(84)=.06$, $p=.61$ and mastery, $r(84)=.08$, $p=.45$.

IRI perspective taking was also associated with positive models of others, $r(87)=.24$, $p=.02$, but negatively correlated with narcissism, $r(89)=-.30$, $p=.005$. It was unrelated to the principle of care, $r(89)=.05$, $p=.64$, and positive models of self, $r(87)=.16$, $p=.15$. In terms of psychological well-being, perspective taking was positively related with growth, $r(84)=.38$, $p<.001$, but was unrelated to the other well-being facets: positive relations with others,

r(84)=.17, *p*=.12, purpose, *r(*84)=-.08, *p*=.48, self-acceptance, *r*(84)=.09, *p*=.40, autonomy, *r*(84)=.14, *p*=.22, and mastery, *r*(84)=.08, *p*=.50.

Objective measures. Contrary to prior research on empathy and emotion recognition (Konrath, Corneille, et al., 2014), none of the empathy measures were related to the number of correctly identified emotions in this study (SITES: *r*(86)=-.10, *p*=.36; empathic concern: *r*(90)=-.08, *p*=.47; perspective taking: *r*(90)=-.16, *p*=.13). Nor were they related to the degree to which research assistants saw participants as empathic (SITES: *r*(86)=-.08, *p*=.45; empathic concern: *r*(89)=.18, *p*=.09; perspective taking: *r*(89)=-.13, *p*=.23). This might be because it might be difficult for research assistants to accurately assess empathy in participants after very minimal interactions (e.g. reading consent form, giving them questionnaires, etc.). Note that because only one research assistant rated the participant, and no inter-rater reliability was assessed, this measure may not be a reliable way to rate participants' empathy.

In terms of prosocial emotions and behaviors, participants scoring higher on the SITES felt more compassion, *r*(82)=.35, *p*=.001, and distress for Katie Banks, *r*(82)=.21, *p*=.055, and also personal distress, *r*(82)=.24, *p*=.03, after listening to the radio program. Although the SITES was associated with feeling more compassion for Katie Banks, it did not significantly predict a greater likelihood of helping her, *β*=.32, *p*=.27, *OR*=1.38 [.78, 2.43]. Because of known gender differences in both empathy and helping behavior (Eagly, 2009; O'Brien, Konrath, Grühn, & Hagen, 2013), we examined whether this effect was moderated by gender, and found that it was not, *β*=-.22, *p*=.71, *OR*=.81 [.25, 2.56].

Participants scoring higher on IRI empathic concern felt more compassion, *r*(84)=.43, *p*<.001, and distress for Katie Banks, *r*(84)=.44, *p*<.001, and also personal distress, *r*(85)=.29, *p*=.007, after listening to the radio program. Unlike the SITES, empathic concern predicted a

greater likelihood of helping her, β=1.11, p=.01, OR=3.02 [1.30,7.03]. This effect was not moderated by gender, β=-1.36, p=.13, OR=.26 [.04, 1.51].

Participants scoring higher on perspective taking felt more compassion, $r(84)$=.23, p=.04, but not more distress for Katie Banks, $r(84)$=.11, p=.34, or personal distress, $r(85)$=.10, p=.35. Perspective taking did not significantly predict a greater likelihood of helping her, β=.40, p=.27, OR=1.50 [.74, 3.04], nor was this effect moderated by gender, β=.24, p=.75, OR=1.27 [.31, 5.24].

Summary. Overall, the SITES has good test-retest reliability, as now demonstrated in two studies. And although the SITES has predictable relationships with self-report measures, there are less consistent relationships with more objective measures, at least in these first examinations. Studies 6 and 7 will revisit other objective measures in different populations. Taken together, the SITES has more similar correlates with empathic concern than perspective taking.

STUDY 6

The main purpose of Study 6 was to examine the relationship between the SITES and emotion recognition in a larger and more diverse sample. We also examined its relationship to a self-reported measure of emotional intelligence (alexithymia). Finally, we examined its relationship to self-esteem and narcissism in order to replicate these important relationships.

Participants

Participants were 270 adults (46.2% male, mean age of 33.5, SD=11.6) years recruited from MTURK for a small payment.

Procedures

These measures were included in a questionnaire battery that was part of an unrelated study examining the relationship between self-construal and emotion recognition. Since this study manipulated self-construal, we ran both raw correlations, and correlations controlling for experimental condition. Self-esteem was measured using the Single Item Self-Esteem Scale (Robins et al., 2001), which asks participants to what extent they agree that they have high self-esteem (1=*not very true of me*; 5=*very true of me*). This measure of self-esteem is different than Rosenberg's (1965) Self-Esteem Scale used in Study 4. As in Studies 3, 4, and 5, participants also completed the Single Item Narcissism Scale (Konrath, Meier, et al., 2014).

Alexithymia (i.e. low emotional intelligence) was assessed using the 20-item Toronto Alexithymia Scale (Bagby, Parker, & Taylor, 1994; Bagby, Taylor, & Parker, 1994), which consists of a total score (α=.85) and 3 facets: (1) difficulty identifying feelings (e.g. "*I am often confused about what emotion I am feeling*"; α=.88), (2) difficulty describing feelings (e.g. "*It is difficult for me to find the right words for my feelings*"; α=.76), and (3) externally oriented thinking (e.g. "*I prefer talking to people about their daily activities rather than their feelings*"; α=.91; 1=*disagree strongly*; 5=*agree strongly*). Emotion recognition was again assessed using the UCDSEE (Tracy et al., 2009), with identical procedures as Study 5 (10 emotions, times 2 faces).

Results and Discussion

Participants' average SITES score was 3.81 (*SD*=0.96), and men scored lower on empathy (*M*=3.61, *SD*=0.94) than women did (*M*=4.01, *SD*=0.92), $F(1,266)=12.18$, $p=.001$, $d=0.43$.

Unlike in Study 4, which found a positive relationship between the SITES and self-esteem, Study 6 found that the SITES was unrelated to a different measure of self-esteem,

r(270)=.06, *p*=.32. However, as in Studies 3-5, those scoring higher in SITES also scored lower in narcissism, *r*(270)=-.26, *p*<.001.

In terms of self-reported emotional intelligence, those who scored higher on the SITES also scored lower on alexithymia, *r*(270)=-.25, *p*<.001, and its subscales. This indicates that those with higher empathy report having less difficulty in identifying, *r*(270)=-.19, *p*=.002, and describing their feelings, *r*(270)=-.22, *p*<.001, and have less externally oriented thinking, *r*(270)=-.25, *p*<.001. Although Study 5 found that the SITES was unrelated to performance on an emotion recognition task among college students, there was a positive association among this general sample of adults, *r*(270)=.18, *p*=.003. We return to this issue in the General Discussion. Note that when controlling for condition, the size of all correlations remained nearly identical.

STUDY 7

The main purpose of Study 7 was to again examine whether the SITES was associated with prosocial behavior (Batson, 2011). We also examined whether it was related to narcissism and aggressive cognitions.

Participants

Participants were 185 adults recruited from MTURK for a small payment. They were 37.8% male, with a mean age of 35.9 (*SD*=12.7).

Procedures

In Study 7, narcissism was assessed with the Single Item Narcissism Scale (Konrath, Meier, et al., 2014) as in our prior studies, but we also assessed it with the 40-item version of the Narcissistic Personality Inventory (Raskin & Terry, 1988). For each item, participants are asked to choose which of two statements best applied to them (e.g. "*I really like to be the center of attention*" versus "*It makes me uncomfortable to be the center of attention*"). The NPI-40 has a

total score (α=.86) and 7 subscales: vanity (α=.79), leadership/authority (α=.79), self-sufficiency (α=.52), superiority (α=.64), exhibitionism (α=.66), exploitativeness (α=.52), and entitlement (α=.62).

Aggressive cognitions were assessed using a short version of the aggressive word completion task (Anderson, Carnagey, & Eubanks, 2003). Participants were asked to complete 6 word stems, 4 of which had possible aggressive completions (e.g. MU _ _ ER), and 2 of which had only neutral or positive completions (e.g. L _ _ E). Participants' aggressive cognitions scores were calculated as the sum of their aggressive word completions (e.g. MURDER).

Prosocial behavior was assessed via two tasks. First, participants were told that this survey was being conducted by a student, and that we needed their confidential opinions of the student's work (1=*strongly disagree*; 5= *strongly agree*). We thus asked two questions about the survey itself (e.g. "*The student's online survey was user-friendly*") and two questions about whether the student deserved to be hired for a paid job in our research lab (e.g. "*If I were in charge of hiring people to design online surveys, I would hire this student*"). We then asked participants to recommend a grade for the student's survey, in half-grade increments ranging from F (coded as 1) to A+ (coded as 13). (Note that only 101 participants responded to this question.) For our second prosocial behavior measure, we asked participants if they would be willing to help us with our research by completing another survey for us for free as soon as they finished this one (1=*yes*, 0=*no*).

Results and Discussion

Participants' average SITES score was 4.07 (0.96), and men scored lower on empathy (*M*=3.87, *SD*=0.97) compared to women (*M*=4.18, *SD*=0.93), $F(1,184)=4.72$, $p=.03$, $d=0.33$.

As in Studies 3-6, those scoring higher in SITES also scored lower on the SINS, $r(184)=-.28$, $p<.001$. The SITES was also negatively correlated with overall narcissism on the NPI-40, $r(184)=-.16$, $p=.03$, and three of the NPI subscales: entitlement, $r(184)=-.29$, $p<.001$, superiority, $r(184)=-.16$, $p=.03$, and vanity, $r(184)=-.15$, $p=.048$. However, it was uncorrelated with the other NPI subscales, ps>.42. Previous research has shown that the entitlement subscale is the strongest predictor of aggression and violence (Bushman & Baumeister, 2002; Konrath, Bushman, & Campbell, 2006).

In terms of aggressive cognitions, self-reported empathy on the SITES was unrelated to the number of aggressive word stem completions, $r(184)=.04$, $p=.60$. This suggests that those scoring higher on the SITES do not necessarily have fewer aggressive cognitions as we might expect. However, it is possible that this relationship would only exist when in the presence of an aggressive situational cue. In addition, future research should examine whether the SITES is associated with less aggressive behavior, as per prior research (Eisenberg, Eggum, & Di Giunta, 2010).

As for prosocial behavior, participants who scored higher on the SITES rated the student's survey more positively, $r(184)=.17$, $p=.02$, and were more willing to recommend that we hire her in our research lab, $r(184)=.26$, $p<.001$. They also recommended a higher grade for her work, $r(101)=.24$, $p=.01$. On the second task, overall, empathy predicted a greater likelihood of helping the experimenters by completing a survey for free, $\beta=.48$, $p=.03$, OR=1.61 [1.05, 2.46]. However, the significant interaction by gender, $\beta=-.73$, $p=.03$, OR=.48 [.25, .94], found that empathy only predicted a greater likelihood of helping among women, $\beta=.48$, $p=.03$, OR=1.61 [1.05, 2.45], and not among men, $\beta=-.25$, $p=.34$, OR=1.61 [.78, 1.30].

Taken together, although Study 5 found that the SITES was unrelated to prosocial behavior among college students, it predicted prosocial behavior, especially among women, in a general sample of adults in Study 7. We return to this issue in the General Discussion.

General Discussion

Across seven studies using a wide range of participant samples (total *N*=5,724), we developed a single item measure of empathy that was found to be both reliable and valid. (See Table 2 for a summary of our results.) Although longer measures of empathy should be used when researchers are able to do so, our single item measure is recommended for use in exploratory research settings and in situations in which participant time and question quantity are under strain (e.g. online studies, mobile phone studies, pilot studies, etc). We offer a summary of our results as well as some limitations below.

Summary of Results

The SITES correlated positively with the most widely used measure of empathy, the Interpersonal Reactivity Index (Davis, 1983). However, not all subscales of the IRI were positively related to the SITES. The empathic concern and perspective taking subscales of the IRI were consistently positively related to the SITES, but the fantasy and personal distress subscales were not. Therefore, researchers who are interested in assessing the fantasy and personal distress components of empathy should use the IRI rather than the SITES. We also note that there was an inconsistent quadratic relationship in the studies that included the IRI (Studies 1, 4, and 5). Readers can examine the Supplementary Tables for more details.

The SITES was consistent over time: test-retest correlations from 2 weeks to 6 months were all greater than .55. In all but one of the studies, women had higher empathy scores than

men. As per prior research, empathy scores on the SITES also tended to increase with age (O'Brien et al., 2013), although the correlations were small.

Regarding *intrapersonal* outcomes, empathy was negatively related to narcissism in all six studies that measured both constructs. This makes good theoretical sense. Because narcissists are so focused on themselves, they do not necessarily consider how others might feel. The SITES was negatively related to depressive and anxiety symptoms (Study 4), although the correlations were small. The SITES was positively related to self-esteem (Studies 4 & 6), growth (Study 5), and positive relations with others (Study 5). As expected, the SITES was negatively related to alexithymia (Study 5). In terms of the Big 5 personality dimensions, the SITES was most strongly related to agreeableness (Studies 2A and 2B). Finally, people who score high on the SITES value both social rewards such as helping and seeing friends, and nonsocial rewards such as food and money (Study 4).

Regarding *interpersonal* outcomes, the SITES was positively related to adult attachment (Study 5), we also found that people high in empathy who are concerned about others, were also concerned about the planet (Study 3).

Not all findings were consistent. The SITES was unrelated to the number of correctly identified emotions in Study 5 (college students), although the expected relationship was obtained in Study 6 (general sample of adults online). There were also inconsistent results with respect to helping behaviors. In Study 5 (college students), participants scoring higher on the SITES felt more compassion for someone in distress, but were not more likely to help her like we might expect. Yet, in Study 7 (online adults), the SITES did predict increased prosocial behaviors directed toward two different targets (i.e. a student who purportedly designed the survey and a professor who needed help completing an additional survey).

We cannot fully explain these inconsistent results but they could have something to do with drawing on different participant samples. For prosocial behavior, it is possible that the nature of the task affected the results. Study 5 included a very distressed target whose parents had just died in a car crash, and participants scoring higher on the SITES also reported more feelings of personal distress, which could have interfered with their helping behavior (Batson, 2011). In Study 7, the targets were not distressed, and the prosocial behaviors were more along the lines of being polite and courteous to others. It is possible that the SITES is better at predicting these types of everyday helping behaviors rather than emergency or high distress helping behaviors. A multidimensional scale might be better at distinguishing these types of helping responses.

Likewise, the SITES was unrelated to aggressive cognitions in Study 6, perhaps due to the lack of aggressive cues in the paradigm. Future research should examine the relation between SITES scores and aggressive cognitions in the presence of aggressive cues.

Comparison with Interpersonal Reactivity Index

We compared the SITES and Davis (1983) Interpersonal Reactivity Index in terms of their correlations with key outcomes in the studies in which used both measures (see Table 2 for summary of these results). In general, the SITES and the IRI had correlations in the same direction, but the IRI often had larger effect sizes (especially empathic concern). This is interesting considering that we are comparing participants' direct self-perceptions on a construct ("I am an empathetic person") to more abstract concepts related to the construct (e.g. having feelings of compassion for others). Practically, this means that overall, the conclusions that researchers would make from using the SITES would generally be similar to the ones they would make if using the longer IRI subscales. However, with the smaller effect sizes, it would be more

difficult to find effects with the SITES compared to the IRI. One important exception was that the IRI empathic concern subscale predicted prosocial behavior in Study 5, while the SITES and the IRI perspective taking subscale did not. We recommend that researchers choose the scale that best suits their outcome of interest. Note that participants answered significantly above the midpoint for both the SITES and the IRI (see Table 1). This indicates that participants in our studies saw themselves as being somewhat empathic.

Limitations

The primary limitation of the SITES is that it does not assess the different dimensions or facets of empathy. We note that single-item measures typically suffer from this limitation as only one item is being used to measure a construct and therefore multiple facets cannot be measured. Although the SITES does predict theoretically relevant behaviors, it is not as comprehensive as longer measures because it consists of only a single item. Thus, when statistical power is low or effect-size estimates are expected to be small, a longer measure of empathy is recommended. Researchers who are interested in detecting fine differences in empathy levels should also use a longer measure. Thus, we believe that this single item measure should only be used when it would be difficult or impossible to include a longer empathy scale.

There may also be limits to the generalizability of our results. Future research should examine whether the results we reported generalize across different populations, especially in other cultures, since the meaning of empathy may depend upon one's culture. Future research should also test the validity of the SITES using different measures than we used in the current studies.

Conclusion

Although longer measures of empathy exist (e.g., IRI), there are situations when the SITES might be useful. For example, single-item scales can be useful for studies in which every single question counts in terms of time or participant attention levels (e.g. online studies, large nationally representative surveys, field studies in which a single page on a clipboard is an ideal survey length). In addition, this measure might be useful when using interactive electronic data collection techniques such as text messaging, EMA, or smartphone surveys, in which each number or response given takes effort for participants. Yet, in typical laboratory settings, we recommend the use of longer empathy scales. Future studies will help us better understand the predictive properties of the SITES, but for now, the SITES is one useful tool that can help to assess the complex aspects of empathy with one single item that can be completed in seconds. As Theodore Roosevelt noted, "No one cares how much you know, until they know how much you care." The SITES provides a short and valid measure of how empathic and caring people are.

References

Anderson, C. A., Carnagey, N. L., & Eubanks, J. (2003). Exposure to violent media: The effects of songs with violent lyrics on aggressive thoughts and feelings. *Journal of Personality and Social Psychology, 84*, 960-971.

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*(1), 23-32. doi: 10.1016/0022-3999(94)90005-1

Bagby, R. M., Taylor, G. J., & Parker, J. D. A. (1994). The twenty-item Toronto Alexithymia scale—II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research, 38*(1), 33-40. doi: 10.1016/0022-3999(94)90006-x

Bartholomew, K., & Horowitz, L. M. (1991). Attachment Styles Among Young Adults: A Test of a Four-Category Model. *Journal of Personality and Social Psychology, 61*(2), 226-244.

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

Batson, C. D., Batson, J. G., Griffitt, C. A., Barrientos, S., Brandt, J. R., Sprengelmeyer, P., & Bayly, M. J. (1989). Negative-state relief and the empathy—altruism hypothesis. *Journal of Personality and Social Psychology, 56*(6), 922.

Batson, C. D., Batson, J. G., Slingsby, J. K., Harrell, K. L., Peekna, H. M., & Todd, R. M. (1991). Empathic joy and the empathy-altruism hypothesis. *Journal of Personality and Social Psychology, 61*(3), 413.

Berenguer, J. (2007). The effect of empathy in proenvironmental attitudes and behaviors. *Environment and behavior, 39*(2), 269-283.

Britton, P. C., & Fuendeling, J. M. (2005). The relations among varieties of adult attachment and the components of empathy. *The Journal of Social Psychology, 145*(5), 519-530.

Bushman, B. J., & Baumeister, R. F. (2002). Does self-love or self-hate lead to violence? *Journal of Research in Personality, 36*(6), 543-545.

Bushman, B. J., Moeller, S. J., Konrath, S., & Crocker, J. (2012). Investigating the Link Between Liking Versus Wanting Self-Esteem and Depression in a Nationally Representative Sample of American Adults. *Journal of Personality, 80*(5), 1453-1469.

Cusi, A. M., MacQueen, G. M., Spreng, R. N., & McKinnon, M. C. (2011). Altered empathic responding in major depressive disorder: relation to symptom severity, illness burden, and psychosocial outcome. *Psychiatry Research, 188*(2), 231-236.

Davis, M. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113-126. doi: 10.1037/0022-3514.44.1.113

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71-75.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*(2), 192.

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.

Eisenberg, N., Eggum, N. D., & Di Giunta, L. (2010). Empathy-Related Responding: Associations with Prosocial Behavior, Aggression, and Intergroup Relations. *Social issues and policy review, 4*(1), 143-180.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504-528.

Hepper, E. G., Hart, C. M., & Sedikides, C. (2014). Moving Narcissus: Can Narcissists Be Empathic? *Personality and Social Psychology Bulletin, 40*(9), 1079-1091. doi: 10.1177/0146167214535812

Konrath, S., Bushman, B., & Campbell, W. K. (2006). Attenuating the link between threatened egotism and aggression. *Psychological Science, 17*(11), 995-1001. doi: PSCI1818 [pii]10.1111/j.1467-9280.2006.01818.x

Konrath, S., Corneille, O., Bushman, B. J., & Luminet, O. (2014). The Relationship Between Narcissistic Exploitativeness, Dispositional Empathy, and Emotion Recognition Abilities. *Journal of Nonverbal Behavior, 38*(1), 129-143. doi: 10.1007/s10919-013-0164-y

Konrath, S., Falk, E., Fuhrel-Forbis, A., Ho, S., Liu, M., Swain, J. E., . . . Walton, M. (2015). Can Text Messages Increase Empathy and Prosocial Behavior? The Development and Initial Validation of Text to Connect. *PLoS ONE, 10*(9), e0137585.

Konrath, S., & Grynberg, D. (2016). The positive (and negative) psychology of empathy. In D. Watt & J. Panksepp (Eds.), *The neurobiology and psychology of empathy*: Nova Science Publishers, Inc.

Konrath, S., Meier, B., & Bushman, B. (2014). Development and validation of the Single Item Narcissism Scale (SINS). *PLoS ONE, 9*(8), e103469.

Konrath, S., O'Brien, E., & Hsing, C. (2011). Changes in dispositional empathy in American college students over time: a meta-analysis. *Personality and Social Psychology Review, 15*(2), 180-198. doi: 1088868310377395 [pii]

Lee, I. A., & Preacher, K. J. (2013). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. . *Available from http://quantpsy.org*.

Melchers, M. C., Li, M., Haas, B. W., Reuter, M., Bischoff, L., & Montag, C. (2016). Similar personatliy patterns are associated with empathy in four difference countries. *Frontiers in Psychology, 7*, 290.

Mikulincer, M., Gillath, O., Halevy, V., Avihou, N., Avidan, S., & Eshkoli, N. (2001). Attachment theory and rections to others' needs: Evidence that activiation of the sense of attachment security promotes empathic responses. *Journal of Personality and Social Psychology, 81*(6), 1205.

Milfont, T. L., & Sibley, C. G. (2016). Empathic and social dominance orientations help explain gender differences in environmentalism: A one-year Bayesian mediation analysis. *Personality and Individual Differences, 90*, 85-88.

Mooradian, T. A., Davis, M., & Matzler, K. (2011). Dispositional empathy and the hierarchical structure of personality. *Am J Psychol, 124*(1), 99-109.

O'Brien, E., Konrath, S. H., Grühn, D., & Hagen, A. L. (2013). Empathic concern and perspective taking: Linear and quadratic effects of age across the adult life span. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 68*(2), 168-175.

Pfattheicher, S., Sassenrath, C., & Schindler, S. (2015). Feelings for the suffering of others and the environment compassion fosters proenvironmental tendencies. *Environment and behavior*, 0013916515574549.

Radloff, L. S. (1977). The CES-D scale a self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401.

Raskin, R., & Terry, H. (1988). A Principal-Components Analysis of the Narcissistic Personality Inventory and Further Evidence of Its Construct Validity. *Journal of Personality and Social Psychology, 54*(5), 890-902.

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring Global Self-Esteem: Construct Validation of a Single-Item Measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27*(2), 151-161. doi: 10.1177/0146167201272002

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Ryff, C. (1989). Happiness Is Everything, or Is It? Explorations on the Meaning of Psychological Well-Being. *Journal of Personality and Social Psychology, 57*(6), 1069-1081.

Schultz, W. (2002). Empathizing With Nature: The Effects of Perspective Taking on Concern for Environmental Issues-Statis. *Journal of Social Issues, 56*(3), 391-406.

Speilberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press. Palo Alto, CA.

Spielberger, C. D., & Sydeman, S. J. (1994). State-Trait Anxiety Inventory and State-Trait Anger Expression Inventory. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 292-321): Lawrence Erlbaum Associates, Inc.

Spielberger, C. D., Sydeman, S. J., Owen, A. E., & Marsh, B. J. (1999). Measuring anxiety and anger with the State-Trait Anxiety Inventory (STAI) and the State-Trait Anger Expression Inventory (STAXI).

Suzuki, Y., & Kino, K. (2014). Undergraduates' well-being from the perspective of social skills and empathic orientation. *Japanese Journal of Experimental Social Psychology*, *54*(2), 125-133. doi: 10.2130/jjesp.1312

Sze, J. A., Gyurak, A., Goodking, M. S., & Levenson, R. S. (2012). Greater emotional empathy and prosocial behavior in later life. *Emotion, 12*, 1129-1140.

Tracy, J. L., Robins, R. W., & Schriber, R. A. (2009). Development of a FACS-Verified Set of Basic and Self-Conscious Emotion Expressions. *Emotion, 9*(4), 554-559. doi: 10.1037/a0015766

Twenge, J., Konrath, S., Foster, J., Campbell, W. K., & Bushman, B. (2008). Egos inflating over time: a cross-temporal meta-analysis of the Narcissistic Personality Inventory. *Journal of Personality, 76*(4), 875-902; discussion 903-918. doi: JOPY507 [pii] 10.1111/j.1467-6494.2008.00507.x

van der Linden, S., & Rosenthal, S. A. (2016). Measuring narcissism with a single question? A replication and extension of the Single-Item Narcissism Scale (SINS). *Personality and Individual Differences, 90*, 238-241.

Watson, P., Grisham, S., Trotter, M., & Biderman, M. (1984). Narcissism and empathy: validity evidence for the Narcissistic Personality Inventory. *Journal of Personality Assessment, 48*(3), 301-305.

Wei, M., Liao, K. Y. H., Ku, T. Y., & Shaffer, P. A. (2011). Attachment, self-compassion, empathy, and subjective well-being among college students and community adults. *Journal of Personality, 79*(1), 191-221.

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

Wilhelm, M. O., & Bekkers, R. (2012). Principle of Care Scale. *Unpublished scale*.

Table 1: *Descriptive statistics for all studies*

Study	Mean (SD) & Difference from Midpoint	Sample Type	Skewness	Kurtosis
1	SITES: M=4.02, SD=1.09*** IRI EC: M=3.80, SD=0.70*** IRI PT: M=3.51, SD=0.64***	Online	-1.04 -.34 -.16	.47 -.05 -.08
2A	SITES: M=3.81, SD=0.95***	College students	-.59	-.06
2B	SITES: M=4.17, SD=0.90***	Online and college students	-.97	.57
3	SITES: M=3.80, SD=0.98***	Online	-.69	.31
4 Time 1	SITES: M=3.86, SD=1.02*** IRI EC: M=3.95, SD=0.70***	Online	-.76 -.42	-.80 -.21
4 Time 2	SITES: M=3.89, SD=0.99*** IRI EC: M=3.88, SD=0.72***	Online	.29 -.12	.38 -.60
5 Time 1	SITES: M=4.16, SD=0.84*** IRI EC: M=3.92, SD=0.61*** IRI PT: M=3.59, SD=0.67***	College students	-1.15 -.52 -.66	1.90 .01 .62
5 Time 2	SITES: M=4.14, SD=0.74*** IRI EC: M=3.79, SD=0.64*** IRI PT: M=3.65, SD=0.71***	College students	-.84 -.03 -.41	1.06 -.45 .17
5 Time 3	SITES: M=4.03, SD=0.99*** IRI EC: M=3.76, SD=0.70*** IRI PT: M=3.56, SD=0.67***	College students	-.95 -.01 -.72	.24 -.77 1.81
6	SITES: M=3.81, SD=0.96***	Online	-.47	-.47
7	SITES: M=4.07, SD=0.96***		-1.04	.95

Note: We ran one sample t-tests to determine if the mean of the SITES (and IRI, where applicable) was significantly different from the midpoint (3) of the scale. ~p<.10, *p<.05, **p<.01, ***p<.001. EC=empathic concern; PT=perspective taking.

Table 2: *Summary of SITES results*

Measure	Correlation with SITES	Correlation with IRI	Study #
General properties			
Empathy-related measures	Fantasy: *r*=.27***	--	Study 1
	Perspective taking: *r*=.38***	--	Study 1
	Empathic concern: *r*=.50***	--	Study 1
	Personal distress: *r*=-.08***	--	Study 1
	Empathic concern: r=.45***	--	Study 4
	Fantasy: r=.03	--	Study 5
	Perspective taking: r=.37***	--	Study 5
	Empathic concern: r=.49***	--	Study 5
	Personal distress: r=.09	--	Study 5
	Principle of care: r=.30**	EC: *r*=.49*** / PT: *r*=.05	Study 5
Test-retest reliability	2 weeks: *r*=.57***	EC: *r*=.80***	Study 4
	3 weeks: *r*=.67***	EC: *r*=.80*** / PT: *r*=.77***	Study 5
	6 months; *r*=.62***	EC: *r*=.76*** / PT: *r*=.59***	Study 5
Demographic variables			
Gender	Males < Females	Males < Females	Studies 1-4, 6, & 7
	Males = Females	EC & PT: Males = Females	Study 5
Age	Not applicable – college student samples	Not applicable – college student samples	Studies 2A & 5
	r=.03~	EC: *r*=.11*** / PT: *r*=-.01	Study 1
	r=.13*	N/A	Study 2B
	r=.02	N/A	Study 3
	r=.03	EC: *r*=.13***	Study 4
	r=.20**	N/A	Study 6
	r=.16*	N/A	Study 7
Intrapersonal outcomes			
Narcissism	SINS: r=-.11*	N/A	Study 3
	SINS: r=-.08*	EC: *r*=-.33***	Study 4
	SINS: r=-.32**	EC: *r*=-.50*** / PT: *r*=-.30**	Study 5
	SINS: r=-.26***	N/A	Study 6
	SINS: r=-.28***	N/A	Study 7
	NPI-40: r=-.16*	N/A	Study 7
Identification as environmentalist	Self-identified environmentalists scored higher on SITES than those who did not.	N/A	Study 3
Life satisfaction	*r*=.10**	EC: *r*=.01	Study 4
Depressive symptoms	*r*=-.07*	EC: *r*=-.10*	Study 4
Anxiety symptoms	*r*=-.08*	EC: *r*=-.07*	Study 4
Psychological well-being	Growth: *r*=.29**	EC: *r*=.41*** / PT: *r*=.38***	Study 5
	Positive relations with others: *r*=.28**	EC: *r*=.39*** / PT: *r*=.17	Study 5 Study 5
	Autonomy: *r*=-.06	EC: *r*=.06 / PT: *r*=.14	Study 5
	Mastery: *r*=.11	EC: *r*=.08 / PT: *r*=.08	Study 5
	Purpose: *r*=.05	EC: *r*=.27* / PT: *r*=-.08	Study 5
	Self-acceptance: *r*=-.03	EC: *r*=.28* / PT: *r*=.09	
Self-esteem	Rosenberg: r=.13***	EC: *r*=.25***	Study 4

	SISE: r=.06	N/A	Study 6
Alexithymia	Total alexithymia: *r*=-.25*** Difficulty identifying feelings: *r*=-.19** Difficulty describing feelings: *r*=.22*** Externally oriented thinking: *r*=-.25***	N/A	Study 6 Study 6 Study 6 Study 6
Openness	*r*=.21*** *r*=.02	N/A	Study 2A Study 2B
Conscientiousness	*r*=.05 *r*=-.15**	N/A	Study 2A Study 2B
Extraversion	*r*=.07 *r*=.08	N/A	Study 2A Study 2B
Agreeableness	*r*=.47*** *r*=.12*	N/A	Study 2A Study 2B
Neuroticism	*r*=.04 *r*=-.07	N/A	Study 2A Study 2B
Reward preferences	*Social rewards:* Helping others: *r*=.34*** Seeing best friend: *r*=.26*** Favorite sexual activity: *r*=.08* *Non-social rewards:* Self-esteem boosts: *r*=.23*** Eating favorite food: *r*=.21*** Receiving paycheck: *r*=.16*** Favorite alcoholic beverage: *r*=.04	 EC: *r*=.58*** EC: *r*=.44*** EC: *r*=.05 EC: *r*=.32*** EC: *r*=.20*** EC: *r*=.26*** EC: *r*=-.08*	 Study 4 Study 4 Study 4 Study 4 Study 4 Study 4 Study 4
Interpersonal outcomes			
Adult attachment	Positive models of the self: *r*=.44*** Positive models of others: *r*=.08	EC: *r*=.36** / PT: *r*=.24* EC: *r*=.06 / PT: *r*=.16	Study 5 Study 5
Pro-environmental behaviors	*β*=.17***	N/A	Study 3
Recognizing others' facial expressions of emotion	College students: *r*=-.10 Online adults: *r*=.18**	EC: *r*=-.08 / PT: *r*=-.16 N/A	Study 5 Study 6
Aggressive cognitions	*r*=.04	N/A	Study 7
Research assistant rating of participant empathy	*r*=-.08	EC: *r*=.18~ / PT: *r*=-.13	Study 5
Emotional responses to others in distress	Compassion: *r*=.35*** Distress for Katie: *r*=.21~ Personal distress: *r*=.24*	EC: *r*=.43*** / PT: *r*=.23* EC: *r*=.44*** / PT: *r*=.11 EC: *r*=.29** / PT: *r*=.10	Study 5 Study 5 Study 5
Helping behavior (1=yes, 0=no)	Helping someone in distress: *β*=.31 (not moderated by gender) Positive ratings of student: *β*=.17*, *β*=.26***, *β*=.24** Helping professor complete unpaid survey: *β*=.48* (moderated by gender; stronger in females)	EC: *β*=1.11* / PT: *β*=.40 N/A N/A	Study 5 Study 7 Study 7

Note: ~p<.10, *p<.05, **p<.01, ***p<.001. EC=IRI empathic concern; PT=IRI perspective taking.

Highlights

- Empathy involves feeling compassion for others and imagining how they feel.
- In this paper, we validate a single item measure of empathy.
- We find that it correlates in expected ways with intrapersonal and interpersonal outcomes.
- This scale may be useful in situations when time or question quantity is constrained.