

Empathy, Gender, and Prosocial Behavior

Linda Kamas
Department of Economics
Santa Clara University
500 El Camino Real
Santa Clara, CA
lkamas@scu.edu
(Corresponding Author)

Anne Preston
Department of Economics
Haverford College
370 Lancaster Ave.
Haverford, PA 19041
apreston@haverford.edu

July 2020

Abstract

This paper examines how empathy is associated with gender differences in prosocial behavior in some economic games. Examining dictator, charitable giving, public goods, and trust games, we find that accounting for empathy eliminates the significance of gender differences in prosocial behavior in all four games. When the data are pooled from all games, the reduction in the female coefficients are significant after including either total empathy or empathic concern. The strong effect of empathy on prosocial behavior in the four games persists with controls for college major and demographics. This pattern replicates itself only with empathic concern, not the other empathy subscales. We conclude that measured gender differences in prosocial behavior in the games are not due wholly to sex per se, but at least partially to personality traits such as empathy that both sexes share so that the most empathic men act as prosocially as do women.

JEL Codes: C91, D03, D63, D64, J1

Key Words: Empathy, Prosocial, Gender, Social Preferences, Games

Funding: This work was supported by Haverford College, Santa Clara University, and the Leavey School of Business, Santa Clara University

Declarations of interest: none

Informed consent was obtained from all participants

Empathy, Gender, and Prosocial Behavior

1. Introduction

Philosophers and psychologists have long stressed the importance of empathy in engendering prosocial values and actions. The moral sentimentalists, David Hume and Adam Smith, argued that “sympathy” is a key determinant of moral values. More recently, feminists have identified empathy as an important element in the ethics of care (Folbre, 2001; Held, 2006; Noddings, 1984). Psychologists have long studied empathy; for example, the “empathy-altruism” hypothesis of Batson (1991, 2011a, 2011b) posits that empathy leads to altruistic motivation that results in prosocial behaviors, and there is a substantial amount of empirical evidence supporting this hypothesis. Much of the literature also finds that women are more empathic and more prosocial than men.

If empathy promotes prosocial actions and if women are more empathic than men, a reasonable hypothesis is that greater empathy leads women to exhibit more prosocial behavior than men. Empathy has been suggested as a partial explanation for gender differences in other-regarding behavior, but empirical work has not adequately tested this hypothesis. In this paper, we utilize a set of economic games that allow for other-regarding decisions and investigate whether gender differences in empathy can help us better understand gender differences in prosocial behavior in the games. We use an instrument to measure empathy that is well-established in the psychology literature (Davis’s Interpersonal Reactivity Index, IRI) and focus on both the total empathy score and specific types of empathy (empathic concern, perspective-taking, fantasy, and personal distress) to examine the effects of empathy on behavior in dictator, public goods, and trust

games and to examine the role played by empathy in generating the gender differences that are found.

We find that in all of the games, once empathy is accounted for, gender differences in prosocial behavior become insignificant. That is, more empathic men act as prosocially as do women. Empathy has a large positive and significant effect on amount sent in all games with the exception of the trust game where empathy has a strong effect on the amount returned. When empathy is divided into its four subscales, only empathic concern has a consistent effect on prosocial behavior and consistently reduces the effect of gender on prosocial behavior.

In the following section we discuss the literature that provides the basis for our hypothesis. We provide evidence that supports the following “stylized facts:”

- Empathy leads to prosocial behavior.
- Women are more empathic than men.
- Women act more prosocially than men in economic games.

Based on this evidence, we generate and test our hypothesis:

- Greater empathy is an important personality characteristic that leads women to act more prosocially than do men in economic games.

2. Empathy, Prosocial Behavior, and Gender

2.1. Defining and Measuring Empathy

A difficulty in analyzing empathy and behavior results from the lack of agreement on how empathy is defined. The word itself is of somewhat recent vintage; the moral philosophers used the word “sympathy” to indicate a trait similar to what we today call “empathy”. Empathy can be cognitive or affective: the former is knowing what the other

thinks or feels (Theory of Mind), while the latter is sharing the feelings of the other.

Empathy may include an element of feeling for (not just feeling like) the other, sometimes this is called compassion, sympathy, or empathic concern (Batson, 2011a, 2011b; Singer & Tusche, 2014; Zaki & Ochsner, 2012).

Batson (2011b) provides a useful taxonomy by identifying eight concepts of empathy that have been used in the literature: knowing another person's internal state including his or her thoughts and feelings, adopting the posture or matching the neural responses of an observed other, coming to feel as another person feels, intuiting or projecting oneself into another's situation, imagining how another is thinking or feeling, imagining how one would think and feel in the other's place, feeling distress at witnessing another person's suffering, and feeling for another person who is suffering (empathic concern). He concludes that empathic concern is the most important type of empathy in motivating prosocial behavior.

The measurement of empathy is difficult because it is an internal state and has a "multidimensional" nature (Batson, 1991, 2011a, 2011b; Davis, 1980, 1983). In principle, it would be best to have an objective test, an outside observer, or a physiological measurement to assess an individual's empathy. In practice, however, empathy is usually measured by questionnaires or anecdotal situations where subjects indicate the extent to which they are characterized by or agree with certain statements or stories. Two such self-report instruments that have been widely utilized are the Interpersonal Reactivity Index, IRI (Davis, 1980, 1983, 1994) and the Empathy Quotient, EQ (Baron-Cohen and Wheelwright, 2004). Such measures of empathy are subject to the obvious shortcomings of all self-reported measures: people may not really know the truth about what they are

reporting, they may prefer not to reveal information, or they may be dishonest. And these biases may characterize different groups of people in different ways. For example, men may be less willing to admit to thoughts, feelings, or actions that they consider to be “feminine” and vice-versa for women. Since empathy and some behaviors associated with it might be considered “feminine” in U.S. culture, men may report answers that result in lower scores on self-reported measures of empathy than do women.

In spite of the shortcomings of self-report measures of empathy, to date there is no unequivocally superior way to measure this internal and multidimensional quality. One alternative approach is to evaluate how well subjects are able to read faces or the expressions in eyes (Baron-Cohen et al. 1997, 2001; Hall, 1978). While this is useful in that it is a test rather than self-report, it measures a very narrow aspect of cognitive empathy (the ability to read facial cues) but not affective empathy. Another approach (Ickes, 1993, 1997; Ickes et al. 2000), has subjects watch videos of people interacting, asking them to assess what those in the videos were thinking or feeling at various points in the conversation. This measure of “empathic accuracy” also tends to focus on cognitive rather than affective aspects of empathy. To measure empathy in children who may be less able to self-report empathy, Eisenberg and Fabes (1990, 1991) used heart rate, skin conductance, and facial responses to supplement self-reported empathy. The neuroscience literature using brain imaging (fMRIs) provides a substantial amount of evidence that there is overlap in affective brain circuits that respond to one’s own emotions (e.g. pain, disgust, or pleasure) and emotions experienced by others (Fan et al. 2011; Jabbi et al. 2007; Lamm et al. 2011; Singer et al. 2004, 2006; Singer & Fehr 2005; Singer & Tusche 2014; Zaki & Ochsner 2012). Brain imaging holds the promise of being able to provide objective, third-

person measures of the multidimensional aspects of empathy, however, at this point in time, self-reported empathy scales such as the IRI and EQ are used to validate the neurological measures, suggesting that the self-reported surveys are considered to be a valid way to measure empathy.

2.2. Empathy and Prosocial Behavior

The long-held philosophical view that empathy and sympathy motivate prosocial behavior has received a substantial amount of empirical support. Psychologists distinguish between situational or state empathy, which refers to empathy for a particular other experiencing a particular distress at a particular time, from dispositional empathy, which is a personality characteristic or more general way of being or acting. Much of the psychology literature focuses on state empathy. Subjects are induced to feel empathy (exposed to a story or video of a person in a painful situation or asked to imagine how someone in a difficult position feels), empathy is measured (self-report, facial expressions, heart-rate decrease, skin salience), and then helpful responses are observed (e.g. running errands for the person, doing a task to raise donations, taking on painful shocks in place of the other, etc.). Studies of dispositional empathy often use a self-reported measure such as the IRI or the EQ to measure empathy and connect this to prosocial behavior. There is strong evidence that greater empathy (both situational and dispositional) is connected to more prosocial behavior (Batson, 2011a; Batson et al. 2015; Davis, 2015; Eisenberg & Miller, 1987; Hein et al. 2010).

Economists have begun to consider the role of empathy in motivating behavior¹ and some of this work has used games utilized in experimental economics. Empathy is found to be an important indicator of prosocial behavior in the dictator game (Artinger et al. 2014; Edele et al. 2013), ultimatum game (Hoffman et al. 2000), prisoner's dilemma games (Batson & Ahmad, 2001; Batson & Moran, 1999), and early rounds of a social dilemma game (Rumble et al. 2010). There are a few studies that do not find a consistent link between empathy and cooperative or prosocial behavior, for example, in a repeated prisoner dilemma game (Sautter et al. 2007), solidarity game (Büchner et al. 2007; De Oliveira et al. 2014), ultimatum game (Artinger et al. 2014), and trust game (Pelligra, 2011). These divergent findings may be due to differences in game structures causing empathetic motivation to be outweighed by strategic or reciprocal considerations in some cases.

2.3. Gender and Empathy

The empirical evidence largely supports the view that women are more empathic than men as measured by: self-reported questionnaires (Baron-Cohen & Wheelwright, 2004; Davis, 1980, 1994), reading faces or motor cues of others (Hall, 1978; McClure, 2000), reading eyes (Baron-Cohen et al. 1997, 2001; Kirkland et al. 2013), and in some studies of empathic accuracy (Ickes et al. 2000). The neuroscience literature using fMRIs of brain activity also finds gender differences in empathy (Christov-Moore et al. 2014; Derntl et al., 2010; Schulte-Rüther et al. 2008; Singer et al. 2006,). Chakrabarti and

¹ See, for example, Andreoni et al. 2017; Andreoni & Rao, 2011; Grohn et al. 2014; Kirman and Teschl 2010; Molnar-Szakacs, 2011; Singer & Fehr, (2005). Earlier papers on “social distance” implicitly support the role of empathy in eliciting prosocial behavior (Bohnet & Frey, 1999; Charness & Gneezy, 2008; Hoffman et al. 1996).

Baron-Cohen (2006) attempt to explain why women might be more empathic than men.

While some authors question findings that women are more empathic than men (Eisenberg & Lennon, 1983, Ickes et al. 2000, Klein & Hodges, 2001), the bulk of the literature supports the conclusion of greater empathy in women than men.

2.4. Gender and Prosocial Behavior

The issue of whether men or women are more prosocial is a complex one. It is likely that it depends on context and social norms.² We focus here on the economic games that are analyzed in our experiments in this paper.

Eckel and Grossman's (2008) survey on gender differences in economic decisions, concludes that when there is no risk, such as in dictator games, women are more generous than men. While there are some papers that do not find that women send more in dictator games (see the survey in Kamas and Preston 2015), most studies tend to find that women give more than men or there is no significant difference; rarely is it found that men are significantly more generous than women in dictator games.³

Contributions in a public goods game are prosocial in that the amount sent is split equally among the group with no strategic benefit to oneself of sending more. Croson and Gneezy (2009) cite two studies that find men contribute more, one where women contribute more, and four where the difference is insignificant. Therefore, there is little or no evidence that women behave more prosocially in this game than men.

² Croson & Gneezy (2009), Eckel & Grossman (2008), and Kamas & Preston (2015) summarize gender differences in preferences and behavior in the experimental economics literature.

³ There is a substantial literature on gender differences in charitable giving, much of which finds that women give more to charity than men (Barclays Wealth, 2009; Mesch, 2010; Mesch et al. 2006; Piper and Schnepf, 2008).

Offers in trust games are a combination of other-regarding preferences to share payoffs and strategic self-interest in the expectation that the other player will return more than has been sent. Croson and Gneezy (2009) cite eleven studies where men send significantly more than women, seven where the difference is not significant, and two where women send more than men. While this might be interpreted as women have less trust in the other or lower other-regarding preferences, Kamas and Preston (2015) conclude that instead this difference occurs because women are more often inequality averters and an equal outcome can be ensured with an offer of \$2.50 (given a \$10 stake with the amount sent multiplied by 3, each player ends up with \$7.50) while men are more often efficiency maximizers who relish the increase in total payoffs resulting from sending larger amounts. The amount returned in trust games is a better measure of prosocial preferences because there is no strategic payoff for sending more nor is there an increase in total payoffs. Croson and Gneezy (2009) cite seven studies where women return significantly more than men, ten where the difference is not significant, and one where men return more than women. Therefore, women return more or similar amounts as men in the trust game.

2.4. Empathy, Gender, and Prosocial Behavior

There are several economic studies that show that after nudging to induce situational empathy, women make more prosocial decisions than men (Van Rijn, Quiñones, & Barham 2018; Czap et al. 2014; Khachaturyan & Czap, 2016). The paper most closely related to this one, Willer, Wimer, & Owens (2015), posits that empathy may help explain gender differences in giving to charitable organizations in a dictator game. They measure empathy as the degree to which participants agree with the statement, “I am often quite touched by things that I see happen.” They find that empathy is positively

related to participants' stated willingness to give to a (fictitious) poverty-relief organization and that including this empathy measure in regressions reduces the size of the gender effect and eliminates its significance. We extend this work by using a more comprehensive measure of empathy, the Davis Interpersonal Reactivity Index (IRI), which consists of 28 statements measuring both cognitive and affective empathy. The statements are categorized into four different types of empathy: empathic concern, perspective taking, fantasy and personal distress. This provides more nuanced measures of different aspects of empathy (see description below). While the empathic concern scale includes their statement as one of the seven, the scale is richer because it includes questions that focus specifically on fairness and on feelings towards other people. We evaluate the effects of empathy on prosocial behavior in several economics games that enable us to consider how empathy influences behavior in various types of economic decision-making, not just charitable giving. And our giving game offers two very different types of charities, one that is directed to aiding people in distress (Feeding America) and another that focuses on supporting people in creating income earning activities (Accion USA). These allow subjects to make donations that have very different outcomes. Finally, our experiments use real money payouts so the participants are actually giving up money that they could keep (incentivized decision-making). It is much easier to *say* you will give away money than to actually do it. For example, if one wants to make a good impression on others or to feel good about oneself, this objective is achieved at zero cost if real money is not involved.

3. Methodology and Experiments

In summary, the empirical evidence supports the stylized facts that empathy leads to prosocial behavior and that women are more empathic and act more prosocially than men. In our experiments, we test the hypothesis that differences in empathy are associated with gender differences in prosocial behavior in games. We represent this hypothesis with an empirical model of prosocial behavior by individual i :

$$\text{Prosocial Action}_i = \beta_0 + \beta_1 \text{Female}_i + \beta_2 \text{Empathy}_i + \delta \mathbf{X}_i + \varepsilon_i$$

where \mathbf{X}_i is a vector of controls for individual i . If female and empathy were included separately in the equation with no inclusion of controls, their coefficients would be positive. However, because of the expected correlation between empathy and sex, it is possible that one or both of these coefficients may lose significance with the inclusion of both variables in the equation. If greater empathy is responsible for the more prosocial behavior of women rather than sex per se, we would expect that β_2 would be positive and significant and β_1 would be insignificant. At the same time, we acknowledge that there may be other personality characteristics that are associated with empathy that could also contribute to prosocial behavior. However, our focus here is to determine whether it is sex per se that accounts for the gender differences in prosocial behavior; therefore, even if the empathy measure is picking up the effects of related characteristics, our findings that including empathy in regressions eliminates the significance of the female dummy substantiates the hypothesis that it is not sex that is responsible for differences in prosocial behavior but empathy and, possibly, other personal characteristics that both sexes share.

To measure empathy, we utilize the Interpersonal Reactivity Index (IRI) developed by Davis (1980, 1983), which includes the various components of empathy identified by Batson, including both cognitive and affective aspects of empathy. The IRI has been a widely used measure of empathy in the psychology literature and it has been validated in numerous studies (Carey, et al. 1988, Chrysikou and Thompson 2015, Davis 1980, 1983, Hawk et al. 2013, and Pulos et al. 2004). Recent evidence shows that the IRI correlates with neurological measures of empathy. Subjects with higher scores on the IRI empathy scale show stronger brain activity measured in fMRIs in areas related to empathy when observing others experiencing pain (Singer et al., 2004, 2006) and pleasure or disgust (Jabbi et al. 2007). Therefore, the IRI is considered a useful tool for representing heterogeneous, multidimensional empathic responses among subjects.

The IRI consists of a questionnaire in which subjects are asked to score, on a five point scale, how well 28 statements describe them. In addition to the total empathy score, the statements are divided into four sub-categories: empathic concern (experience feelings of sympathy and compassion for unfortunate others); perspective-taking (spontaneously adopt the psychological point of view of others in everyday life); fantasy (imaginatively transpose oneself into fictional situations); and personal distress (experience distress and discomfort in response to extreme distress in others). Appendix Table 1 provides the statements of the IRI divided according to the sub-categories. Some of the statements are reverse-scored so that the larger the empathy index, the greater the empathy. In our empirical results below, for each study participant we calculate the average IRI score for all questions and for the questions in each sub-category.

The experiments were run at Santa Clara University. Students were recruited by email and flyers and paid \$10 show-up payment. The experiments took about 45 minutes and average earnings (including the show-up payment) were \$21.38. 180 students participated in eight sessions, 90 men and 90 women. There were eight exercises, one of which was randomly chosen to be paid out, the IRI survey, and a participant information questionnaire. Subjects were told that they should consider each exercise on its own, ignoring what they decided in other exercises and treating each exercise as if it were the one that would be paid out. Students were not informed in advance how many exercises there would be, and they were assured of complete anonymity. Code numbers were utilized to make the payments, and the person putting money in the envelopes was different from the person who passed them out to the students so that the latter would not know how much money was in the envelope. The exercises of the experiment included:⁴ three-person allocation game (used to categorize social preferences); dictator game (two anonymous people in the study), four charitable giving dictator games (with zero, 25%, 50%, 100% matching subsidies), public goods game, trust game, Davis IRI survey, socio-economic questionnaire. The three-person allocation game was played first in all sessions so that the social preference categorizations would not be influenced by decisions made by oneself or expected of others in the other games. The other exercises were provided in four different orderings among sessions. The IRI was completed at the end of all the sessions so that the questionnaire could not influence behavior in the games by focusing attention on other-regarding considerations or making emphatic thoughts more salient.

⁴ Experiment instructions are provided in the supplementary materials.

4. Results

4.1. Empathy

The IRI scores for the full sample and by gender as well as the correlations between the total empathy score and the subscales and between the total and subscale empathy scores and gender are provided in Table 1. As is usually found with self-reported measures of empathy, women's scores are significantly higher than men's, with the average female total empathy score 11% higher than the male score; however, the standard deviation of the empathy score is similar across men and women. The differences are largest for empathic concern and personal distress. The statements in the IRI that measure personal distress tend to suggest personal qualities or behaviors that might be considered "weak" ("feel helpless", "scares me", "I go to pieces") so that even if true, men might be more loathe to admit such behaviors. In conversations with students, many reported that the statements on personal distress were "off-putting". These concerns are reflected in the much lower scores on this measure for both men and women. The difference in the empathic concern scale between men and women is quite large (14%) and highly significant.

In columns 4 and 5 we present the correlations between each empathy score and the dummy variable for female and the correlations between the total score and the subscales respectively. Column 4 reveals that of the four subscales, female is most correlated with empathic concern and least correlated with perspective-taking while according to column 5, the total score is most closely correlated with empathic concern and least correlated with personal distress. Because Batson (1991, 2011a, 2011b), Eisenberg and Fabes (1990, 1991), and others have argued that empathic concern, more than the other types of

**Table 1: Total IRI and Empathy Subscales Scores
Total Sample and by Gender**

	(1) Total (n=180)	(2) Men (n=90)	(3) Women (n=90)	(4) Correlation Between Total Empathy Score and Female	(5) Correlation Between Total Empathy Score and Subscales
Total Empathy Score (IRI)	2.978 (0.456)	2.825*** (0.418)	3.132 (0.443)	0.340 (0.000)	
Empathy Subscales:					
Empathic Concern	3.344 (0.698)	3.130*** (0.691)	3.559 (0.641)	0.308 (0.000)	0.766 (0.000)
Perspective-Taking	3.232 (0.659)	3.183 (0.620)	3.283 (0.696)	0.076 (0.311)	0.590 (0.000)
Fantasy	3.189 (0.795)	3.029*** (0.781)	3.349 (0.780)	0.202 (0.000)	0.706 (0.000)
Personal Distress	2.146 (0.706)	1.959*** (0.614)	2.333 (0.744)	0.266 (0.000)	0.496 (0.000)

Note: Total IRI Empathy is the average score of all the questions in the IRI survey. The questions that are included in the empathy subscales are provided in Appendix Table 1. Figures in parentheses in columns 1-3 are the standard deviations. Figures in parentheses in columns 4 and 5 are the p values related to the hypothesis that the correlation is significantly different from zero.

*** Average score of men is significantly different than the average score of women at the 0.01 level.

empathy leads to prosocial behavior, and because the empathic concern score is most highly correlated with the total score and most highly correlated with female, we will focus below on both the total empathy score and the empathic concern score as the facets of empathy that may be associated with gender differences in prosocial behavior in these games. However, all tables related to the games are replicated for the other subscales and presented in the appendix.

4.2. Dictator Game

In dictator games, Player A is given an amount of money (\$10 in this case) and asked to make a decision on how to divide the allocated amount between himself or herself and another anonymous person, Player B, who has no influence on the distribution of money. All participants were asked to make a decision on how they would choose to allocate the \$10, not knowing if they were Player A or B. If this game was randomly chosen to be paid out, actual decisions implemented were determined by a coin toss: if the coin toss was a head, people with even code numbers were identified to be Player A and their choices were paid out, while those with odd code numbers were identified as Player B and they received the amount decided on by a random other Player A person. If a tail came up, decisions by subjects with odd code numbers were implemented.⁵ The average amount sent in the dictator game was \$3.91, but men sent \$3.59 (s.d.=2.87) while women sent \$4.22 (s.d.=1.98), a difference significant at the 0.10 level.

Table 2 provides two-limit (0, 10) tobit regressions for the dollar amount sent in the dictator game. Column 1 reveals that women send \$0.93 more than men in the dictator games (significant at 0.10 level).⁶ Empathy, as measured by the total IRI score (column 2) has a large (significant at 0.05 level) effect on amount sent: a one standard deviation increase in the empathy score increases the amount sent by \$0.52, or about a quarter of a standard deviation in giving. In column (3), the effect of empathic concern is

⁵ Therefore, each person took on both roles and knew that there was a 50% chance they would be paid as a dictator and 50% chance they would be paid as recipient. Because there is no strategic benefit to giving in anonymous dictator games in terms of reciprocity, we believe this does not bias giving. It is possible that knowing that they could be paid as recipients may have made dictators more generous, however, there is no reason to believe that this would influence decisions of men and women differently.

⁶ This value is different than the difference in means since we are estimating two level tobits which assume a normal distribution of giving preferences that might extend beyond the end points (0, 10) determined by the game and thus to lead bunching at 0 and 10.

Table 2: Two Level Tobits on Amount Sent in Dictator Game

	(1)	(2)	(3)	(4)	(5)
Female	0.932* (0.488)			0.731 (0.507)	0.648 (0.529)
Empathy		1.131** (0.565)		0.859 (0.594)	
Empathic Concern			0.808** (0.354)		0.662* (0.385)
Constant	3.199*** (0.418)	0.336 (1.723)	0.970 (1.223)	0.777 (1.721)	1.128 (1.224)
Observations	180	179	180	179	180

Notes: Empathy is the average score of all the questions in the IRI survey. Empathic concern is the average score on the subset of questions in this subscale (see Appendix Table 1). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

similar to the effect of empathy, and a one standard deviation increase in the empathic concern score increases giving by \$0.56. When both empathy and female are included in the equation (column 4), the coefficients on female and empathy fall and become insignificant. When both empathic concern and female are included (column 5), both coefficients fall in magnitude but the coefficient on empathy remains significant at the 0.09 level. While the continued significance of the empathic concern coefficient in column 5 implies that empathic concern may be picking up some of the female effect, tests on the coefficients in column 5 reveal that we cannot reject the hypothesis that these coefficients are no different than those estimated in columns 1 and 3. Appendix Table 2 gives results from the same models but including perspective-taking, fantasy, or personal distress to represent empathy and in each case, female remains a positive significant determinant of dictator giving while none of the other components of empathy have a significant impact on dictator giving.

4.3. Charitable Giving

In this set of four exercises, the decision was similar to the dictator game, divide \$10 between self and other, but here the recipient was one of two charities and in three of the exercises the donation was matched or subsidized by the researchers. Students were asked to choose between two charities, Feeding America (a hunger relief organization that distributes food) and ACCION USA (a microfinance organization that gives small loans to micro-entrepreneurs). These two charities were picked because they appeal to different motivations for giving, aiding people suffering from hunger (Feeding America) versus supporting people in creating income earning activities (Accion USA). In addition, it is likely that the two charities would be attractive to people with different types of social preferences highlighted in the literature: those who are inequality averse may choose the former while efficiency maximizers may be more willing to give to the latter (See Charness & Rabin, 2002; Engelmann & Strobel, 2004; Fehr & Schmidt, 1999). In the first of these exercises, subjects were given \$10 which they could either keep or give some portion to their chosen charity. In the three subsequent versions, participants were allowed to keep or donate some or all of the \$10 to the same charities, but the researchers matched the amount donated at 25%, 50%, or 100% subsidies, implying prices of giving of 1 (for no subsidy), 0.80, 0.67, and 0.50, respectively.

The dollar amount sent to the charities at various subsidy rates for the full sample and by gender is provided in Table 3. Without a subsidy, women donate 24% more than men (\$5.78 compared to \$4.66). However, as the subsidy increases (price of giving falls), men increase giving while women decrease giving so that, although women continue to give more than men at the higher subsidy levels (50% and 100%), the gender difference is

**Table 3: Charitable Giving – Mean Dollar Amount Sent
Full Sample and by Sex at Various Subsidy Rates**

	Full Sample	Men	Women
Dollar Amount Sent:			
No Subsidy (price of giving = 1.0)	\$5.217 (3.709)	\$4.656** (3.846)	\$5.778 (3.499)
25% Subsidy (price of giving = 0.8)	\$5.183 (3.705)	\$4.600** (3.885)	\$5.767 (3.438)
50% Subsidy (price of giving = 0.67)	\$5.150 (3.675)	\$4.778 (3.815)	\$5.522 (3.510)
100% Subsidy (price of giving = 0.5)	\$5.061 (3.619)	\$4.844 (3.897)	\$5.278 (3.325)

Notes: Standard deviations are in parentheses; ** Dollar amount for men is significantly different from dollar amount for women at the 0.05 significance level using t-tests.

no longer significant.⁷

The influence of empathy is evaluated in tobits estimating amount donated and displayed in Table 4. In column (1) the coefficient on the female dummy is significant and large for the no subsidy and 25% subsidy treatments (rows 1 and 4), implying that women donate between \$1.77 and \$1.89 more than men. However, it declines in size and is insignificant at the higher subsidy rates (rows 7 and 10). In column 2, empathy has a large significant effect on donations at all subsidy rates, with a one standard deviation increase in empathy associated with between \$2.14 and 2.40 increase in donations. Similarly, in column (3), empathic concern has a large, significant impact on charitable donations and the effect of a one standard deviation change ranges from \$1.96 to \$2.36. In column 4

⁷ This result is similar to Andreoni and Vesterlund (2001) who find male giving increases more as the price of giving declines; however, they find that men give a larger amount than women at low prices of giving. This result may be related to findings that women are more likely to be inequity averters while men are more likely to be social surplus maximizers (Kamas and Preston, 2012). Subsidies increase the total size of payouts so, as subsidies increase, efficiency or surplus maximizers would likely increase giving while inequality averters may respond by reducing giving to keep payoffs more equal between donor and recipient. A similar result was found for charitable giving by Kamas and Preston (2010).

Table 4: Two Level Tobits on Amount Donated to Charity with Various Levels of Subsidy

	(1)	(2)	(3)	(4)	(5)
			No Subsidy		
1. Female	1.765* (0.942)			0.337 (0.976)	0.370 (0.992)
2. Empathy		5.045*** (1.071)		4.916*** (1.177)	
3. Empathic Concern			3.182*** (0.657)		3.094*** (0.732)
			25% Subsidy		
4. Female	1.887** (0.944)			0.293 (0.967)	0.406 (0.970)
5. Empathy		5.272*** (1.045)		5.163*** (1.137)	
6. Empathic Concern			3.375*** (0.672)		3.280*** (0.738)
			50% Subsidy		
7. Female	1.191 (0.898)			-0.313 (0.889)	-0.145 (0.912)
8. Empathy		4.885*** (0.974)		4.998*** (1.035)	
9. Empathic Concern			3.017*** (0.641)		3.050*** (0.693)
			100%Subsidy		
10. Female	0.715 (0.883)			-0.791 (0.857)	-0.550 (0.885)
11. Empathy		4.703*** (0.954)		4.988*** (1.002)	
12. Empathic Concern			2.802*** (0.626)		2.924*** (0.668)

Notes: Empathy is the average score of all the questions in the IRI survey. Empathic concern is the average score on the subset of questions in this subscale (see Appendix Table 1). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

when both empathy and female are included, and in column 5 when both empathic concern and female are included, the coefficients on the female dummy variables uniformly fall and become insignificant. In contrast, the coefficients on the two measures of empathy change very little and retain their significance. For each female coefficient estimated in Column 1,

we test the null hypothesis that it is equal to the female coefficient estimated in column 4 or column 5 with the alternative being that the column 4 or column 5 coefficient is lower than the column 1 coefficient. In all cases, we can reject the null hypothesis in favor of the alternative that with the inclusion of empathy, the coefficient on female has fallen in magnitude at better than the 0.08 significance level. In none of the cases does the effect of empathy fall when female is included. These results strongly imply that the gender differences in giving at zero or low subsidy rates are largely due to differences in empathy. Appendix Table 3 gives results from the same models but perspective-taking, fantasy, or personal distress are the empathy measures. Perspective-taking and fantasy are positive determinants of charitable giving while personal distress has no effect. However, the effect of female does not fall significantly in any model and remains significant in models with perspective-taking.

4.4 Public Goods Game

In the public goods game, participants were told they would be put into a group of five people and each person would be given \$10 to split between their own private account (keeping all the money put into this account) and the public account (where the sum of donations by all group members would be multiplied by two and split equally among the five people). Women give more of the \$10 to the public account than do men (\$5.46 compared to \$4.97), but the difference is not significant.⁸

Table 5 provides tobits on the amount invested in the public account. As expected women invest \$1.30 more than men (column 1) but the coefficient is not significant.

⁸ Because sending money increases the size of total payoffs, the insignificant gender difference in this game may be partially due to men being more likely to be efficiency or surplus maximizers.

Table 5: Two Level Tobits on Amount Donated to Public Good

Variables	(1)	(2)	(3)	(4)	(5)
Female	1.295 (1.043)			0.334 (1.035)	0.402 (1.057)
Empathy		3.653*** (1.156)		3.534*** (1.187)	
Empathic Concern			2.174*** (0.749)		2.087*** (0.778)
Constant	4.867*** (0.821)	-5.304 (3.451)	-1.757 (2.555)	-5.120 (3.432)	-1.671 (2.548)
Observations	180	179	180	179	180

Notes: Empathy is the average score of all the questions in the IRI survey. Empathic concern is the average score on the subset of questions in this subscale (see Appendix Table 1). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Empathy (column 2) and empathic concern (column 3) on the other hand have large, significant effects on the amount put into the public account: a one standard deviation increase in the total empathy score increases investment in the public good by \$1.67, while the effect of a one standard increase in empathic concern is slightly lower at \$1.52.

Adding both the female dummy and the empathy measures (columns 4 and 5) has little effect on the coefficients on empathy which remain large, positive, and significant, but the coefficient on female declines in both magnitude and significance. While a test that the female coefficients are equal between column 1 and columns 4 and 5 can only be rejected at the 20% significance level, these results indicate that empathy is an important determinant of the amount invested in the public good and may contribute to gender gaps in contributions. Appendix Table 4 gives results from the same models but including perspective-taking, fantasy, and personal distress, and, of the three, only fantasy is a positive significant determinant of contributions to the public good.

4.5. Trust Game

In the trust game, Player A was given \$10 and asked how much he or she would send to an anonymous participant in the study, Player B. All participants were told that the amount sent to Player B would be multiplied by three and Player B would then decide how much to return to Player A. Player A's payout was the amount he or she kept in the first round plus the amount returned by Player B. Player B's payout was three times the amount sent by Player A minus the amount he or she returned to Player A. Subjects made decisions as both Player A and as Player B and were told that whether they would actually be paid as Player A or Player B would be randomly determined at the end of the experiment.⁹ All subjects first made a decision as Player A on how much to send to Player B. Then, they were asked to make ten decisions on how much they would return back to Player A for each of the ten possible amounts they might receive if they were Player B (ranging from \$3 to \$30, using the strategy method). At the end of the experiment, if this exercise was chosen to be paid out, subjects were put into anonymous two-person groups in which one person had an even code number and the other an odd code number. A coin was tossed and if a head came up, subjects with even code numbers were paid as Player A and those with odd code numbers were paid as Player B (and vice-versa for tails).

Table 6 provides the mean dollar amounts sent and returned and the average proportion of the amount sent that was returned. Women sent less than men but the

⁹ Therefore, each person took on both roles in the trust game and knew that there was a 50% chance they would be paid as Player A and 50% chance they would be paid as Player B. Having people decide both as sender and receiver maximized the number of decisions we obtained. The meta-analysis of Johnson and Mislin (2011) shows that a large number of trust games in the literature have participants play both roles. The aim of the randomness of the payoff was to reduce biases in giving behavior, however, we acknowledge that knowing that they could be paid in either roles may affect subjects' decisions.

**Table 6: Trust Game – Amounts Sent and Returned
Full Sample and by Gender**

	Total	Men	Women
1. Dollar Amount Sent	4.678 (3.268)	\$4.878 (3.457)	\$4.478 (3.073)
2. Average Dollar Amount Returned	5.445 (3.075)	\$4.847*** (3.114)	\$6.049 (2.930)
3. Average Proportion Returned	0.293 (0.177)	0.263** (0.184)	0.324 (0.165)

Notes: Rows 2 and 3 represent averages for all ten possible amounts returned. Standard deviations are in parentheses; ***, ** Score of men is significantly different from score of women at the 0.01 and 0.05 significance levels respectively using t-tests.

difference is not significant. Women returned an average \$6.05 while men returned an average \$4.85 (a 25% difference, significant at 0.01 level). Similarly, women returned about 32% of the amount received while men returned 26% (difference significant at 0.05 level). The amount returned in the trust game is a better measure of prosocial behavior than the amount sent because there is no expectation of reward for sending money back. Sending money as the first mover may be motivated by concern for the other but it also may be a self-interested action based on the belief (trust) that the other will return more than was sent. Thus, the amount sent conflates prosocial and self-interested motivations. Because women do not send more than men in the trust game, we focus our analysis below on the effect of empathy on the amount returned in the trust game.¹⁰

Table 7 provides tobits for the average percent that is returned in the trust game over the ten possible amounts received. Column (1) shows that women return a

¹⁰ We estimated tobits of the dollar amount sent in the trust game and empathy does not have a significant effect on the amount sent; both the empathy variable and its interaction with female are small and insignificant. These results are available from the authors.

Table 7: Two level Tobits on Percent of Money Sent Back in Trust Game

Variables	(1)	(2)	(3)	(4)	(5)
Female	0.073** (0.030)			0.054* (0.031)	0.047 (0.032)
Empathy		0.095*** (0.033)		0.073** (0.036)	
Empathic Concern			0.070*** (0.021)		0.059** (0.024)
Constant	0.248*** (0.022)	0.009 (0.097)	0.052 (0.071)	0.043 (0.100)	0.064 (0.073)
Observations	179	178	179	178	179

Notes: Empathy is the average score of all the questions in the IRI survey. Empathic concern is the average score on the subset of questions in this subscale (see Appendix Table 1). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

significantly higher percentage than men (0.073). Empathy and empathic concern also have strong effects on how much the second mover returns to the first mover (columns 2 and 3). One standard deviation increases in the empathy scale and the empathic concern scale increase the percent returned by 4.3 and 4.9 percentage points respectively. Once the empathy and female variables are included together in the column 4 and 5 models, both the female and the empathy coefficients fall in magnitude. However, the empathy coefficients remain significant at the 0.05 significance level. In the model with the total empathy score (column 4) the female coefficient is significant at the 0.10 level but becomes insignificant in the model with empathic concern (column 5). Again, results from the trust game support the hypothesis that gender differences in prosocial behavior are linked to empathy. Appendix Table 5 gives results from the same models but including perspective-taking, fantasy, or personal distress to represent empathy. While perspective-taking and fantasy have positive significant effects on the amount sent back, in all models the coefficient on female remains positive and significant and of a similar magnitude to the model without a measure of empathy.

5. Pooled Sample of Responses

We acknowledge that the small number of observations reduces the power of any test we might conduct concerning the value and change in the coefficients. As a result, we pool the responses to the games to get a variable we call “prosocial” which represents the action in each game. All actions range from 0-10 except for the percentage returned in the trust game which varies from 0 to 1. Therefore, we multiply this variable by 10 before it gets added to the pooled responses. We then replicate the two level tobits of Tables 2, 4, 5, and 7, and estimate them with dummy variables for the different games and with errors clustered by id. The results are presented in Table 8.

Table 8: Two level Tobits on Pooled Responses in 7 Games

Variables	(1)	(2)	(3)	(4)	(5)
Female	1.190** (0.564)			0.220 (0.563)	0.277 (0.576)
Empathy		3.397*** (0.645)		3.316*** (0.695)	
Empathic Concern			2.148*** (0.411)		2.085*** (0.449)
Constant	2.88*** (0.445)	-6.587*** (1.918)	-3.692*** (1.382)	-6.457*** (1.965)	-3.623*** (1.409)
Observations	1259	1252	1259	1252	1259

Notes: Pooled responses are from the dictator game, the four versions of the giving to charity, the public goods game, and the return decisions in the trust game. The tobits include dummy variables for each game. Empathy is the average score of all the questions in the IRI survey. Empathic concern is the average score on the subset of questions in this subscale (see Appendix Table 1). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The large significant coefficient in column 1 reveals that women have a larger prosocial response than men, and the difference is approximately equal to one third of a standard deviation of the prosocial variable. As in earlier analyses the empathy variables in columns two and three also have large positive and significant effects on the prosocial response. The magnitude of the effects are quite similar: a one standard deviation increase in the empathy variable increases the prosocial

response by just over 40% of a standard deviation. In the column 4 and 5 models where both female and empathy measures are included, the female coefficient falls from 1.19 to 0.22 and 0.28 in the regressions which include total empathy and empathic concern respectively. The reductions in the female coefficient are significant at the 0.043 and the 0.056 levels respectively, implying that empathy is at least partially responsible for the gender differences in prosocial behavior in economics games.

6. Does Empathy Remain a Significant Contributor to Prosocial Behavior with Controls for Education and Demographics?

With empathy levels being higher for women than for men, one might anticipate that empathy is just signaling differences in experiences of men and women that teach different forms of prosocial behavior. For example, students of business or economics may learn to engage in more self-oriented behavior (profit maximization), and women are less likely to major in these subjects. Therefore, we re-estimate the two level tobits for each of the games including female, empathy and controls for education (dummies for economics, business, and engineering majors), race (dummy variables for Asian and other race, which includes Hispanic, Native American and African American), practicing a religion, income (five income categories), volunteering over the past year, and political views (dummies for conservative and moderate views).

The results are presented in Table 9 with each column presenting the estimates of the game identified at the top of the column. Rows 1 and 2 present the coefficients on female and empathy in regressions that included the total empathy measure and rows 3 and 4 present the same coefficients in regressions that include the empathic concern score. The results are particularly strong for empathic concern; it is a strong predictor of prosocial behavior in giving to charity at various subsidy levels (columns 2-5), the public goods

Table 9: Does the Effect of Empathy Persist with Controls for Education and Demographics?

VARIABLES	(1) Dictator game	(2) Charity no Subsidy	(3) Charity 25% Subsidy	(4) Charity 50% Subsidy	(5) Charity 100% Subsidy	(6) Public Good Game	(7) % Sent Back in Trust Game
1. Female	0.411 (0.534)	0.691 (1.002)	1.101 (0.955)	0.327 (0.876)	-0.231 (0.868)	0.551 (1.069)	0.058** (0.029)
2. Empathy	0.540 (0.594)	5.431*** (1.225)	5.811*** (1.166)	5.455*** (1.082)	5.289*** (1.049)	3.103** (1.231)	0.075* (0.039)
3. Female	0.296 (0.553)	0.574 (1.012)	1.085 (0.984)	0.348 (0.903)	-0.252 (0.913)	0.594 (1.106)	0.049 (0.0302)
4. Empathic Concern	0.603 (0.371)	3.468*** (0.730)	3.678*** (0.749)	3.383*** (0.696)	3.096*** (0.667)	1.625** (0.783)	0.062** (0.0253)

Notes: Dependent variable is given at the top of each column. Controls include dummies college major, income, race, political affiliation, volunteer activity, and practicing a religion. Empathy is the average score of all the questions in the IRI survey. Empathic concern is the average score on the subset of questions in this subscale (see Appendix Table 1). All regressions are two level tobits. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

game (column 6), and the trust game (column 7), even after including controls for education and demographics. While the magnitude of the empathic concern coefficient in the dictator game falls slightly, reducing the significance level, it is almost significant with a p value of 0.107. The effect of female (row 3) remains insignificant for all models. The results are similar for the total empathy score presented in rows 1 and 2; empathy is a significant predictor of prosocial behavior in all games except for the dictator game, and female is not a significant predictor in any of the games except the trust game. We also re-estimate each of these models excluding the empathy variable, and in all of the models

except the dictator game (where it falls), the female coefficient is similar in magnitude and significance to its levels in earlier tables. The results remain strong. Empathy is a stronger predictor of prosocial behavior in games than sex, even controlling for education and demographics.¹¹

7. Conclusion

This paper investigates the hypothesis that gender differences in prosocial behavior found in some economic games are associated with women's higher empathy. Davis's Interpersonal Reactivity Index describing total empathy and the four subscales, empathic concern, perspective-taking, fantasy, and personal distress are utilized to measure empathy and to examine the influence of empathy on decisions of subjects in the dictator game, charitable giving, public goods game, and trust game. Empathy is found to play an important role in all the games; more empathetic subjects act more prosocially. We find that levels of empathy are higher for women than for men, and women tend to act more prosocially than men. In all games, controlling for total empathy or empathic concern causes the female effect to become insignificantly different from zero, and in some games, the female coefficient becomes significantly smaller. When the data are pooled from all games, the reduction in the female coefficients are significant after including either total

¹¹ We also estimated these models with dummy variables for social preference categories determined by a set of three-person allocation decisions. These decisions can separate those who are self-interested, inequality averse, efficiency maximizing, or compassionate social surplus maximizing (the latter two being two different types of social surplus maximizers), categories used in the social preference literature (Charness and Rabin, 2002; Engelmann and Strobel, 2004; and Fehr and Schmidt, 1999). Except for inequality averse individuals showing more empathy than others, these groupings do not coincide well with empathy. In models where we include these categories, the coefficients on female and either total empathy or empathic concern are virtually identical to the results we present in Table 9.

empathy or empathic concern. These consistent results do not hold for the other empathy subscales. These results suggest that it is empathic concern that drives the gender differences rather than sex per se. Men and women with similar levels of empathic concern exhibit similar levels of prosocial behavior.

We address concerns that empathy is only signaling different experiences of men and women in which they learn different levels of prosocial behavior. Rerunning all the models with controls for college major and demographics does not impact the coefficients on empathy. Empathy has an independent effect on prosocial behavior in games. Further, it is only the inclusion of total empathy score or empathic concern and not any of the other measures of empathy that causes the gender effect on behavior in these games to shrink and become insignificant.

References

- Andreoni, J., Rao J.M., 2011. The power of asking: How communication affects selfishness, empathy, and altruism. *Journal of Public Economics*, 95, 513-20.
- Andreoni, J., Rao, J.M. and Trachtman, H., 2017. Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy*, 125, 625-653.
- Andreoni, J., Vesterlund, L., 2001. Which is the fair sex? Gender differences in altruism. *The Quarterly Journal of Economics* 116, 293-312.
- Artinger, F., Exadaktylos, F., Koppel, H. and Sääksvuori, L., 2014. In others' shoes: do individual differences in empathy and theory of mind shape social preferences?. *PloS one*, 9, p.e92844.
- Barclays Wealth, 2009. Tomorrow's Philanthropist.
http://www.barclayswealth.com/insights/assets/pdf/Barclays-Wealth-Tomorrows-Philanthropist_UK.pdf
- Baron-Cohen, S., Jolliffe, T., Mortimore, C., Robertson, M., 1997. Another advanced test of theory of mind: evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry*, 38, 813-822.
- Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34, 163-175.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I., 2001. The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42, 241-251.
- Batson, C.D., 1991. *The Altruism Question, Toward a Social-Psychological Answer*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Batson, C.D., 2011a. *Altruism in Humans*. Oxford: Oxford University Press.
- Batson, C.D., 2011b. These things called empathy: eight related but distinct phenomena. In: Decety, J., Ickes, W. (Eds). *The Social Neuroscience of Empathy*. Cambridge, MA: The MIT Press, 3-15.
- Batson, C.D. Ahmad, N., 2001. Empathy-induced altruism in a prisoner's dilemma II: what if the target of empathy has defected? *European Journal of Social Psychology*, 31, 25-36.

Batson, C.D., Ahmad, N. and Lishner, D.A., 2015. Empathy and altruism. In: Schroeder, D.A., Graziano, W.G. (Eds.). *The Oxford Handbook of Prosocial Behavior*, Oxford: Oxford University Press, 259-281.

Batson, C.D., Moran, T., 1999. Empathy-induced altruism in a prisoner's dilemma. *European Journal of Social Psychology*, 29, 909-924.

Bohnet, I. and Frey, B.S., 1999. Social distance and other-regarding behavior in dictator games: Comment. *The American Economic Review*, 89, 335-339.

Büchner, S., Coricelli, G. and Greiner, B., 2007. Self-centered and other-regarding behavior in the solidarity game. *Journal of Economic Behavior & Organization*, 62, 293-303.

Carey, J. C., Fox, E.A., Spraggins, E.F., 1988. Replication of structure findings regarding the interpersonal reactivity index. *Measurement and Evaluation in Counseling and Development*, 21, 102-105.

Chakrabarti, B. and S. Baron-Cohen, 2006. Empathizing: neurocognitive developmental mechanisms and individual differences. In: Anders, S., Ende, G., Junghofer, M., Kissler, J., Wildgruber, D. (Eds.). *Understanding Emotions; Progress in Brain Research 156*, Amsterdam: Elsevier, , 403-417.

Charness, G., Gneezy, U., 2008. What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior & Organization*, 68, 29-35.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. *The Quarterly Journal of Economics* 117, 817-869.

Christov-Moore, L., Simpson, E. A., Coudé G., Grigaityte, K., Iacoboni, M., Ferrari P.F., 2014. Empathy: Gender effects in brain and behavior. *Neuroscience and Biobehavioral Reviews*, 46, 604-627.

Chryssikou, E.G., Thompson, W.J., 2015. Assessing cognitive and affective empathy through the interpersonal reactivity index: an argument against a two-factor model. *Assessment*, 1-9.

Croson, R., Gneezy, U., 2009. Gender differences in preferences. *Journal of Economic Literature* 47, 448-74.

Czap, N.V., Czap, H.J., Burbach, M.E., Lynne, G.D., 2014. Gender in environmental context: the effect of property rights, fines, and empathy nudging. *International Journal of Economics and Finance*, 6, 11-23.

Davis, M.H., 1980. A Multidimensional Approach to Individual Differences in Empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.

Davis, M.H., 1983. Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach. *Journal of Personality and Social Psychology*, 44, 113-126

Davis, M.H., 1994. *Empathy: A Social Psychological Approach*. Madison, WI: WCB Brown & Benchmark Pub.

Davis, M.H., 2015. Empathy and prosocial behavior. In: Schroeder, D.A., Graziano, W.G. (Eds.). *The Oxford Handbook of Prosocial Behavior*, Oxford: Oxford University Press, 282-306.

De Oliveira, A.C., Eckel, C.C. Croson, R.T., 2014. Solidarity among the poor. *Economics Letters*, 123, 144-148.

Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D.I., Schneider, F., Habel, U., 2010. Multidimensional assessment of empathic abilities: neural correlates and gender differences. *Psychoneuroendocrinology*, 35, 67-82.

Eckel, C.C., Grossman, P.J., 2008. Differences in the economic decisions of men and women: experimental evidence. In: Plott, C. R., Smith, V.L. (Eds.). *Handbook of Experimental Economics Results*, Volume 1. Amsterdam: Elsevier B.V., 509-519.

Edele, A., Dziobek, I. and Keller, M., 2013. Explaining altruistic sharing in the dictator game: The role of affective empathy, cognitive empathy, and justice sensitivity. *Learning and Individual Differences*, 24, 96-102.

Eisenberg, N., Fabes, R. A., 1990. Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14, 131-149.

Eisenberg, N., Fabes, R. A., 1991. Prosocial behavior and empathy: A multimethod developmental perspective. In: Clark, M.S. (Ed.). *Prosocial Behavior*, *Review of Personality and Social Psychology*, Vol.12. Newbury Park, CA: Sage, 34-61.

Eisenberg, N., Lennon, R., 1983. Sex differences in empathy and related capacities. *Psychological Bulletin*, 94, 100-131.

Eisenberg, N., Miller, P.A., 1987. The Relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, 101, 91-119.

Engelmann, D., Strobel, M., 2004. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review* 94, 857-869.

Fan, Y., Duncan, N.W., de Greck, M. and Northoff, G., 2011. Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience & Biobehavioral Reviews*, 35, 903-911.

Fehr, E., Schmidt, K. M., 1999. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics* 114, 817-868.

- Folbre, N., 2001. *The Invisible Heart; Economics and Family Values*. New York: The New Press.
- Grohn, J., Huck, S. and Valasek, J.M., 2014. A note on empathy in games. *Journal of Economic Behavior & Organization*, 108, 383-388.
- Hall, J.A., 1978. Gender effects in decoding nonverbal cues. *Psychological Bulletin*, 85, 845-857.
- Hawk, S.T., Keijsers, L., Branje, S.J., Graaff, J.V.D., Wied, M.D., Meeus, W., 2013. Examining the interpersonal reactivity index (IRI) among early and late adolescents and their mothers. *Journal of Personality Assessment*, 95, 96-106.
- Hein, G., Silani, G., Preuschoff, K., Batson, C. D., Singer, T., 2010. Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, 68, 149-160.
- Held, Virginia, 2006. *The Ethics of Care, Personal Political, and Global*. Oxford: Oxford University Press.
- Hoffman, E., McCabe, K. and Smith, V. L., 1996. Social distance and other-regarding behavior in dictator games. *The American Economic Review*, 86, 653-660.
- Hoffman, E., McCabe, K. and Smith, V. L., 2000. The impact of exchange context on the activation of equity in ultimatum games. *Experimental Economics*, 3, 5-9.
- Ickes, W., 1993. Empathic accuracy. *Journal of Personality*, 61, 587-610.
- Ickes, W. (Ed.), 1997. *Empathic Accuracy*. New York: The Guilford Press.
- Ickes, W., Gesn, P.R. and Graham, T., 2000. Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships*, 7, 95-109.
- Jabbi, M., Swart, M. and Keysers, C., 2007. Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage*, 34, 1744-1753.
- Johnson, N.D., Mislin, A.A., 2011. Trust games: a meta-analysis. *Journal of Economic Psychology*, 32, 865-889.
- Kamas, L., Preston, A., 2010. What can social preferences tell us about charitable giving? Evidence on responses to price of giving, matching, and rebates. In: Isaac, R. M., Norton, D.A. (Eds.). *Charity with Choice, Research in Experimental Economics*, Vol. 13. United Kingdom: Emerald Group Publishing Limited, 165-199.
- Kamas, L., Preston A., 2012. Gender and social preferences in the US: an experimental study. *Feminist Economics*, 18, 135-160.

- Kamas, L., Preston, A., 2015. Can social preferences explain gender differences in economic behavior? *Journal of Economic Behavior & Organization*, 116, 525-539.
- Khachatryan, M., Czap, N.V., 2016. Different Strokes for Different Folks: Gender and emotions in an Environmental Game. *Sustainable Agricultural Research*, 5, 81-93.
- Kirkland, R.A., Peterson, E., Baker, C.A., Miller, S. and Pulos, S., 2013. Meta-analysis Reveals Adult Female Superiority in "Reading the Mind in the Eyes Test". *North American Journal of Psychology*, 15, 121-146.
- Kirman, A. and Teschl, M., 2010. Selfish or selfless? The role of empathy in economics. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365, 303-317.
- Klein, K.J. and Hodges, S.D., 2001. Gender differences, motivation, and empathic accuracy: When it pays to understand. *Personality and Social Psychology Bulletin*, 27, 720-730.
- Lamm, C., Decety, J. and Singer, T., 2011. Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*, 54, 2492-2502.
- McClure, E.B., 2000. A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, 126, 424-453.
- Mesch, D.J., 2010. *Women give 2010; New research about women and giving*. Women's Philanthropy Institute, Center on Philanthropy at Indiana University.
- Mesch, D.J., Rooney, P.M., Steinberg, K.S., Denton, B., 2006. The effects of race, gender, and marital status on giving and volunteering in Indiana. *Nonprofit and Voluntary Sector Quarterly* 35, 565-587.
- Molnar-Szakacs, I., 2011. From actions to empathy and morality—A neural perspective. *Journal of Economic Behavior & Organization*, 77, 76-85.
- Noddings, N., 1984. *Caring, A Feminine Approach to Ethics and Moral Education*. Berkeley: University of California Press.
- Pelligrà, V., 2011. Empathy, guilt-aversion, and patterns of reciprocity. *Journal of Neuroscience, Psychology, and Economics*, 4, 161-73.
- Piper, G., Schnepf, S.V. 2008. Gender differences in charitable giving in Great Britain. *Voluntas: International Journal of Voluntary & Nonprofit Organizations* 19, 103-124.

Pulos, S., Elison, J. and Lennon, R., 2004. The hierarchical structure of the Interpersonal Reactivity Index. *Social Behavior and Personality: an International Journal*, 32, 355-359.

Rumble, A.C., Van Lange, P.A. and Parks, C.D., 2010. The benefits of empathy: When empathy may sustain cooperation in social dilemmas. *European Journal of Social Psychology*, 40, 856-866.

Sautter, J.A., Littvay, L. and Bearnes, B., 2007. A dual-edged sword: Empathy and collective action in the prisoner's dilemma. *The ANNALS of the American Academy of Political and Social Science*, 614, 154-171.

Schulte-Rüther, M., Markowitsch, H.J., Shah, N.J., Fink, G.R. and Piefke, M., 2008. Gender differences in brain networks supporting empathy. *Neuroimage*, 42, 393-403.

Singer, T., Fehr, E., 2005. The neuroeconomics of mind reading and empathy. *American Economic Review*, 95, 340-45.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J. and Frith, C.D., 2004. Empathy for pain involves the affective but not sensory components of pain. *Science*, 303, 1157-1162.

Singer, T., Seymour, B., O'Doherty, J.P., Stephan, K.E., Dolan, R.J. and Frith, C.D., 2006. Empathic neural responses are modulated by the perceived fairness of others. *Nature*, 439, 466-469.

Singer, T., Tusche, A., 2014. Understanding others: brain mechanisms of theory of mind and empathy. In: Glimcher, P.W., Fehr, E. (Eds.). *Neuroeconomics; Decision Making and the Brain*, 2nd edition. London: Academic Press, 513-532.

van Rijn, J., Quiñones, E.J., Barham, B.L., 2018. Does Empathic Concern for Children Explain the Gender-Donations Gap? *Agricultural & Applied Economics Staff Papers Series*, University of Wisconsin-Madison.

Willer, R., Wimer, C. and Owens, L.A., 2015. What drives the gender gap in charitable giving? Lower empathy leads men to give less to poverty relief. *Social Science Research*, 52, 83-98.

Zaki, J., Ochsner K.N., 2012. The neuroscience of empathy: progress, pitfalls and promise. *Nature Neuroscience*, 15, 675-680.

Appendix Table 1: Interpersonal Reactivity Index (IRI)

(Statements describing empathic tendencies are grouped according to the sub-category of empathy. The number preceding each statement gives the order in which it is presented in the questionnaire.)

Empathic Concern

2. I often have tender, concerned feelings for people less fortunate than me.
4. Sometimes I don't feel very sorry for other people when they are having problems. (reverse scored)
9. When I see someone being taken advantage of, I feel kind of protective towards them
14. Other people's misfortunes do not usually disturb me a great deal. (reverse scored)
18. When I see someone being treated unfairly, I sometimes-don't feel very much pity for them. (reverse scored)
20. I am often quite touched by things that I see happen.
22. I would describe myself as a pretty soft-hearted person.

Perspective-Taking

3. I sometimes find it difficult to see things from the "other guy's" point of view. (reverse scored)
8. I try to look at everybody's side of a disagreement before I make a decision.
11. I sometimes try to understand my friends better by imagining how things look from their perspective.
15. If I'm sure I'm right about something, I don't waste much time listening to other people's arguments. (reverse scored)
21. I believe that there are two sides to every question and try to look at them both.
25. When I'm upset at someone, I usually try to "put myself in his shoes" for a while.
28. Before criticizing somebody, I try to imagine how I would feel if I were in their place.

Fantasy

1. I daydream and fantasize, with some regularity, about things that might happen to me.
5. I really get involved with the feelings of the characters in a novel.
7. I am usually objective when I watch a movie or play, and I don't often get completely caught up in it. (reverse scored)
12. Becoming extremely involved in a good book or movie is somewhat rare for me. (reverse scored)
16. After seeing a play or movie, I have felt as though I were one of the characters.
23. When I watch a good movie, I can very easily put myself in the place of a leading character.
26. 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.

Personal Distress

6. In emergency situations, I feel apprehensive and ill-at-ease.
10. I sometimes feel helpless when I am in the middle of a very emotional situation.
13. When I see someone get hurt, I tend to remain calm. (reverse scored)
17. Being in a tense emotional situation scares me.
19. I am usually pretty effective in dealing with emergencies. (reverse scored)
24. I tend to lose control during emergencies.
27. When I see someone who badly needs help in an emergency, I go to pieces.

Appendix Table 2: Two Level Tobits on Amount Sent in Dictator Game

Variables	(1)	(2)	(3)
Female	0.948** (0.477)	0.879* (0.497)	0.891* (0.503)
Perspective-Taking	0.474 (0.377)		
Fantasy		0.171 (0.317)	
Personal Distress			0.109 (0.343)
Constant	1.695 (1.309)	2.681** (1.039)	2.985*** (0.795)
Observations	179	180	180

Notes: The empathy measures Fantasy, Perspective-Taking, and Personal Distress are the average scores of the scales identified in Appendix Table 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 3: Two Level Tobits on Amount Sent in Charitable Giving Game

Variables	(1)	(2)	(3)
		No Subsidy	
Female	1.657* (0.904)	1.222 (0.945)	1.565 (0.994)
Perspective-Taking	2.357*** (0.683)		
Fantasy		1.706*** (0.598)	
Personal Distress			0.519 (0.717)
		25% Subsidy	
Female	1.678* (0.907)	1.263 (0.933)	1.829* (1.002)
Perspective-Taking	2.598*** (0.672)		
Fantasy		1.954*** (0.565)	
Personal Distress			0.150 (0.722)
		50% Subsidy	
Female	0.989 (0.855)	0.584 (0.873)	1.168 (0.943)
Perspective-Taking	2.615*** (0.594)		
Fantasy		1.978*** (0.521)	
Personal Distress			0.061 (0.674)
		100% Subsidy	
Female	0.466 (0.827)	0.135 (0.854)	0.736 (0.929)
Perspective-Taking	2.937*** (0.588)		
Fantasy		1.918*** (0.522)	
Personal Distress			-0.053 (0.631)

Notes: The empathy measures Fantasy, Perspective-Taking, and Personal Distress are the average scores of the scales identified in Appendix Table 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 4: Two Level Tobits on Amount Sent in Public Goods Game

Variables	(1)	(2)	(3)
Female	1.307 (1.027)	0.820 (1.023)	1.035 (1.080)
Perspective-Taking	1.049 (0.687)		
Fantasy		1.570** (0.642)	
Personal Distress			0.680 (0.752)
Constant	1.534 (2.318)	0.105 (2.110)	3.536** (1.690)
Observations	179	180	180

Notes: The empathy measures Fantasy, Perspective-Taking, and Personal Distress are the average scores of the scales identified in Appendix Table 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 5: Two Level Tobits on Amount Sent Back in Trust Game

Variables	(1)	(2)	(3)
Female	0.073** (0.029)	0.061** (0.029)	0.084** (0.032)
Perspective-Taking	0.050** (0.021)		
Fantasy		0.035* (0.019)	
Personal Distress			-0.028 (0.022)
Constant	0.091 (0.069)	0.144** (0.059)	0.303*** (0.047)
Observations	178	179	179

Notes: The empathy measures Fantasy, Perspective-Taking, and Personal Distress are the average scores of the scales identified in Appendix Table 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.