Fall 2024

Compassion Cartography: Mapping the Psychological Landscape of Altruistic Equity and Effectiveness

Kyle Fiore Law
University at Albany, State University of New York, kflaw@albany.edu

COMPASSION CARTOGRAPHY:

MAPPING THE PSYCHOLOGICAL LANDSCAPE OF ALTRUISTIC EQUITY AND EFFECTIVENESS

by

Kyle Fiore Law

A Dissertation

Submitted to the University at Albany, State University of New York

In Partial Fulfilment of

The Requirements for the Degree of

Doctor of Philosophy

Department of Psychology

August 2024

ABSTRACT

Psychologists and philosophers have often championed deliberative reasoning over empathy as a better approach to overcoming parochial biases and guiding altruistic equity and effectiveness. Advocates of the effective altruism (EA) philosophy even advise philanthropists to downregulate empathic responses to maximize the impact of their donations. However, recent research reveals that extraordinary altruists (XAs), such as those who donate organs to strangers, are driven primarily by empathy. This dissertation explores whether empathy necessarily impedes altruistic equity (impartial regard for others' welfare) and effectiveness (prioritizing impact), revealing a more nuanced reality that challenges this perspective. Across two phases of research involving three subject groups (N_{Total} = 360; N_{EA} = 119, N_{XA} = 65, $N_{control}$ = 174), I investigate the cognitive, affective, moral, and attitudinal underpinnings of equitable and effective altruistic behavior among self-identifying effective altruists (EAs), extraordinary altruists (XAs), and ordinary adults (controls). The findings suggest that empathy is elevated among XAs, while reasoning is elevated among EAs. However, both capacities predict the prioritization of equity and effectiveness across populations, at times only showing their maximal prosocial benefits when both are high. Furthermore, expansive moral concern, utilitarian beliefs in impartial beneficence, compassionate love, identification with all of humanity, and intriguingly, group loyalty, also play critical roles. Bringing EAs and XAs into the lab for the first time in a single, unified investigation, the current insights suggest ongoing debates over the altruistic utility of reasoning versus empathy may be based on a false dichotomy. Instead, the data suggest that there is no single recipe for exceptional altruism, and that both reasoning *and* empathy are critical ingredients that work best *together*.

Keywords: Effective Altruism, Extraordinary Altruism, Reasoning, Empathy, Morality, Attitudes

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Unlike any other time in history, modern-day humans have a more expansive capacity to help and even save the lives of distant strangers, thanks to advances in technology, economics, medicine, and agriculture (Anthony & Emami, 2018; Kang et al., 2023; Mwalupaso et al., 2019). But this capacity is often underused. Generally, people care deeply about and help socially close others such as family and friends, but are much less likely to extend care and help to distant and unrelated others (Abrams et al., 2018; Cikara et al., 2011, 2014; Law, Amormino, et al., 2024). This issue is compounded by the fact that the greatest impact through altruism can be achieved by helping distant strangers, especially from the standpoint of people who have the greatest amount to give, as their resources can save more lives and alleviate more suffering in less affluent regions (MacAskill, 2018; Singer, 2015). As such, overcoming persistent biases in care and altruism may be necessary if humanity is to overcome many of the grandest challenges it currently faces, including pandemic diseases, global poverty and the effects of climate change, which require individual and collective actions to help distant strangers in order to most impactfully reduce suffering and maximize collective welfare (Bloom, 2016; Gates & Gates, 2021; MacAskill, 2018, 2022; Roser & Ortiz-Ospina, 2019; Singer, 2016; Theron Pummer & William MacAskill, 2019). Addressing these challenges requires understanding what features of the human mind and sense of morality give rise to altruistic equity–an impartial sense of concern for and willingness to help others across social, spatial and temporal divides–and altruistic effectiveness–a proclivity towards helping behavior that maximizes gains in welfare for the greatest number.

Dominant psychological and biological models can account for, and emphasize, biases in moral concern and altruism. Biological models of cooperation demonstrate that favoring socially

close individuals fosters the formation and maintenance of cohesive small groups (e.g., kin selection, Hamilton, 1964; reciprocity, Dunbar & Shultz, 2007; Rand & Nowak, 2013; Trivers, 1971), and that favoring genetically close individuals–who are also typically socially close–maximizes inclusive fitness (Lehmann & Rousset, 2020). At a psychological level, preferential valuation of close others' welfare is promoted by prevailing tendencies to favor individuals who are personally familiar, more similar to, or otherwise categorized with the self (Amodio & Cikara, 2021; Cameron & Payne, 2011; B. A. Jones & Rachlin, 2009; B. Jones & Rachlin, 2006; Kogut & Ritov, 2005a; Syropoulos, Law, et al., 2023; Syropoulos, Law, & Young, 2024d). Consequently, people generally feel more empathy for suffering ingroup members compared to suffering outgroup members (Bloom, 2016; Cikara et al., 2014; Fowler et al., 2021) and are more willing to sacrifice limited resources to help socially close others relative to acquaintances or strangers (Rachlin & Jones, 2008; Syropoulos, Law, & Young, 2024d; K. Vekaria et al., 2017).

The persistence of parochial biases in care and altruism is notable in part because one of the predominant aims of psychological research over the last half century has been to better understand various forms of bias in an effort to reduce them (Boggio et al., 2023; Dovidio et al., 2010; Sherif et al., 1961; Tajfel, 1970, 1979; Turner & Reynolds, 2003). Furthermore, attenuating parochial bias may have substantial pragmatic consequences for global well-being. Considering the stark disparities in welfare and living conditions between affluent societies and developing countries, philanthropic contributions from the affluent can achieve maximal impact on human welfare gains when channeled towards beneficiaries facing greater hardships in less developed regions of the world (Bloom, 2016; Singer, 2015, 2016). This inherent divide between those most in need and those best equipped to help underscores the urgency to confront deep-rooted biases that favor close others over distant strangers and one identifiable victim over the many.

In recent efforts to counteract parochiality and promote effectiveness in altruism, philosophers (see the "effective altruism" philosophy; e.g., Singer, 1981, 2015, 2016), economists (e.g., Mandler, 2020; Wade-Benzoni, 2006, 2017) and psychologists (e.g., Caviola et al., 2021, 2022; Greene, 2014; Huang et al., 2020; Wilks et al., 2023) alike have proposed encouraging deliberative reasoning–specifically emphasizing utilitarian goals–as a means of reducing bias to maximize collective welfare[1]. Despite decades of research connecting empathy to heightened prosocial behavior (Batson et al., 1983; Batson & Shaw, 1991), emotion, and empathy in particular (Bloom, 2016; MacAskill, 2018; Prinz, 2011; Singer, 2015), have more recently been cautioned against as too inherently biased to be useful in promoting altruistic equity and impact. Echoing this sentiment, psychologist Paul Bloom (2016) asserts, "Empathy is biased, pushing us in the direction of parochialism and racism."

This claim above is not entirely unfounded. Empathic emotion *is* often felt in a biased manner. Indeed, many people empathize more strongly with those who are closer and more similar to themselves (Cikara et al., 2011; Dovidio et al., 2010; Fowler et al., 2021). Moreover, empathy does regularly guide to whom we direct our altruistic efforts (Batson et al., 1983; Batson & Shaw, 1991; Coke et al., 1978). Taken together, when causes that evoke a greater emotional response are not optimally effective, emotionality can hinder effective resource allocations and drive parochial bias in altruism. For instance, emotionally-appealing but

[1]To clarify, the recent ethical philosophy of effective altruism (EA) emphasizes utilitarian reasoning over other forms of ethical reasoning (e.g., virtue ethics; Kant, 1873) because it has two primary goals: (1) reducing parochiality (i.e., promoting equity in altruism) and (2) maximizing impact (i.e., promoting effectiveness in altruism). These goals are often intertwined in real-world contexts. That is, the most cost-effective causes–from the perspective of those most affluent who have the most to give–typically are those that benefit distant strangers (MacAskill, 2018; Singer, 2015). While utilitarianism and Kantian ethics *both* oppose parochiality, EA focuses on utilitarianism in particular because it aims to not only overcome parochiality, but also enhance the effectiveness of altruistic actions. By reducing parochial bias, EA seeks to ensure that the most cost-effective causes, which typically benefit distant strangers, receive support. This approach aligns with the utilitarian objective of maximizing global welfare. Thus, parochiality is seen as counterproductive to EA's aim of prioritizing effectiveness, and utilitarian reasoning is championed by EAs as a solution to promote both equity *and* effectiveness simultaneously.

ineffective causes are normatively prioritized over more effective alternatives which lack emotional appeal to prospective donors (Berman et al., 2018).

Even more significantly, the research by Berman and his colleagues (2018) emphasizes that emotions still strongly shape philanthropic choices even in the presence of information about effectiveness. Recognizing the persistent influence of emotions on prosocial decision-making underscores the complex dynamics between people's feelings and their capacity for reasoned analysis in guiding charitable giving. Thus, while emotionality and deliberative reasoning *can* pose countervailing influence on philanthropic decision-making (Bloom, 2016; J. D. Greene et al., 2004; Singer, 2015), the significant impact of emotions on these decisions suggests that harnessing both emotional and rational appeals could be crucial in bolstering equitable and effective altruistic engagement. That is, rather than suppressing emotions in philanthropic decisions, as has been suggested by some proponents of the effective altruism philosophy and social movement (Bloom, 2016; Singer, 2015), there may be value in aiming to nurture greater empathy for more distant beneficiaries and for causes with substantial potential for effective impact.

A recent pre-registered study confirms that expanding empathy can potentially enhance altruistic giving. These findings reveal that when it comes to encouraging prosocial behavior toward strangers in need, emotional appeals are just as effective as rational ones (Lindauer et al., 2020; see Small et al., 2007 for a related finding). The results of this study and others (e.g., Balliet et al., 2019; Batson et al., 1983; Olivola & Shafir, 2013) demonstrate that parochial bias in altruism *can* be overcome in everyday life by emotional *and* rational pathways, with many participants given the opportunity to donate to an anonymous suffering stranger voluntarily choosing to do so. Similarly, despite the general population tending to *experience* empathy more strongly for the suffering of closer others than for the suffering of more distant others, recent

research is beginning to demonstrate that people *judge* feeling equal empathy for both close others and distant strangers to be most morally praiseworthy (Fowler et al., 2021). In other words, people see the moral value of experiencing empathy towards others more equitably even though they sometimes fall short of doing so. In the real world, volunteerism, financial donations, and efforts to help strangers are not uncommon (Lasby & Barr, 2018), yet people overwhelmingly prefer helping close others with their time and resources. Consequently, a critical as yet unanswered question remains unaddressed: Why are some individuals compelled to alleviate the suffering of distant others, even when it demands significant personal sacrifice?

In this dissertation, I seek to determine whether care and altruism possess untapped potential within the general populace. Adopting a multidisciplinary method, I explore how emotion and deliberative reasoning may both play roles in cultivating an expansive concern for the well-being of others and a tendency to prioritize the impact of altruistic resource allocations, focusing on exceptionally caring groups like extraordinary and effective altruists, two cohorts that demonstrate profound equity and impartiality in *who* they choose to help and achieve a robust prosocial impact in *how* they help others. By studying these exceptionally altruistic populations, the primary aim is to gain insight into how equitable and effective prosociality may be nurtured and broadened in the wider community.

1.1. The Ethics of Expanding Care

The majority of the world's major religious and philosophical traditions emphasize the importance of extending care and altruism to socially distant others. The Hebrew Torah and Christian Bible include invocations to not only "honor thy father and thy mother" (Exodus 20:12), but also to "love thy neighbor" (Leviticus 19:17) and "rescue the weak and the needy" (Psalm 82:4). The Christian parable of the Good Samaritan makes explicit the high moral value of sacrificing to help socially distant strangers, as does Maimonides' hierarchy of charity, which

places anonymous helping above direct giving. Among the five pillars of the Islamic faith is charity for those who are less fortunate. Buddhist teachings go further in explicitly endorsing the value of benevolence and compassion toward all beings, with an emphasis on universality in valuing the welfare of others in need (Desbordes et al., 2015; Lama, Dalai, 2001). This emphasis is paralleled by some non-religious philosophical movements like effective altruism (Riedener et al., 2021), which has emerged to promote engagement in equitable altruism by encouraging adherents to minimize the extent to which their altruistic decisions are driven by intuitive emotional preferences for close others and instead rationally evaluate which altruistic decision would lead to the greatest overall gains in welfare (Burum et al., 2020; Caviola et al., 2021; GiveWell, 2023; Schubert & Caviola, 2021; Singer, 1981, 2016).

Although various religious and philosophical traditions overwhelmingly endorse care and altruism for distant others as morally good (Barasch et al., 2014; Bostyn et al., 2018; Bostyn & Roets, 2016; Carlson & Zaki, 2018; Pizarro et al., 2003), many individuals do not share the specific belief that care and altruism should disregard social preferences and relationships (Everett et al., 2018; Roff & Fairbairn, 2007), and some even perceive feeling care for and extending help to distant strangers as *immoral*. In emerging research, most respondents view donating resources to socially distant others *instead of* family members and friends as less moral than donating to family members and friends *instead of* distant others (Everett et al., 2018; Kahane et al., 2018; Law et al., 2022; McManus et al., 2020, 2021). The effect of parochial bias on moral views of altruism is so pronounced that even when the objective welfare gains associated with donating to a distant other explicitly and substantially outweigh the welfare gains associated with donating to a family member or friend (a life saved vs. a brief period of happiness), people still continue to rate helping distant others as less moral.

These pervasive norms render all the more remarkable the existence of exceptionally caring populations like effective altruists, who donate large portions of income to distant strangers suffering severe hardship and extraordinary altruists who take on extreme risks and sacrifices, such as donating a kidney or portion of their liver to benefit complete strangers. What psychological features of these exceptionally caring populations give rise to patterns of such equitable and effective prosocial tendencies and regard for all life, and can they help to elucidate how the normal boundaries of compassion and altruism might be expanded in the rest of us?

1.2. The Plurality of Avenues Towards Universal Care and Altruism

Decades of seminal research across cognitive science and social psychology has detailed extensively the relationships between emotion, cognition and altruism. Early research in psychological science provided strong support for the notion that empathy serves as a force for good, with vast potential to promote altruism across social, geographic and ideological divides. In this vein, numerous studies have substantiated a connection between experimentally induced empathy and subsequent acts of altruism (Batson et al., 1983; Batson & Shaw, 1991; Coke et al., 1978; Eisenberg & Miller, 1987). This link between empathy and altruism has been observed for both affective components of empathy, such as sharing in another's emotions (Batson & Powell, 2003; Dickert et al., 2011; Toi & Batson, 1982), and cognitive components, like perspective-taking, which can lead to empathic emotions and increased altruistic behavior (Batson et al., 1997; Coke et al., 1978).

Moreover, research examining empathy's impact on intergroup relations suggests that taking the perspective of or experiencing empathic concern for outgroup members can enhance attitudes and reduce prejudice toward these outgroups (Dovidio et al., 2010; Galinsky et al., 2005; Galinsky & Moskowitz, 2000; Vescio, Theresa K. et al., 2003), as well as other groups that are either socially remote or historically stigmatized (Lamm et al., 2007; Tarrant & Hadert,

2010). Similarly, investigations into individual differences in empathy have yielded consistent findings regarding its prosocial benefits (Fowler et al., 2021; Lamm et al., 2007, 2011; Weisz & Cikara, 2021), showing that individuals with greater empathic capacity tend to behave more prosaically towards others, even those who are socially distant or otherwise dissimilar. Collectively, this body of research suggests that empathy may occupy a significant place in our altruistic toolkit, enabling individuals and groups to cultivate heightened concern for others and fostering altruistic behavior, even across group boundaries.

Despite empirical evidence supporting empathy as a positive force for increasing engagement in altruistic behaviors, recent scholarship has raised concerns and resistance regarding the assertion that empathy should be actively promoted (Bloom, 2016; Carlson & Zaki, 2018; Prinz, 2011; Zaki, 2018). According to these accounts, empathy is intrinsically biased and parochial in nature, meaning it tends to be limited in scope and primarily felt more strongly toward a select group of individuals. This viewpoint finds support in research on the Identified Victim Effect, which illustrates that it is easier for individuals to experience empathy when faced with a person they can associate with a face and a name (Kogut & Ritov, 2005a, 2005b), and scope insensitivity, which illustrates that people often display markedly muted empathic signatures in response to mass suffering when compared to the empathy they feel when confronted with the suffering of a single individual (Kogut et al., 2015; Sharma & Morwitz, 2016; Västfjäll et al., 2014).

Similarly, empathy tends to be more readily extended to those who share similarities with oneself, potentially rendering emotionality problematic in intergroup contexts (Amodio & Cikara, 2021; Bruneau et al., 2017; Cikara et al., 2011, 2014; Kteily & Bruneau, 2017; Vollberg et al., 2019). For instance, a study exploring sports rivalries revealed that soccer fans exhibited similar neural activation in the anterior insula when they received a shock themselves and when

they observed fellow fans of their favored teams receiving a shock. However, this neural signature of empathy was significantly diminished when they watched fans of rival teams experiencing the same physical pain, which in turn predicted a strong in-group bias in costly helping behavior (Hein et al., 2010). This line of inquiry highlights that empathy is often unevenly directed towards individuals who are closer, more similar, or more easily identifiable, with this bias in empathic responses leading to in-group favoritism in the context of prosociality. Additionally, arguments have been raised that feeling empathy for a specific individual can lead to biased helping behavior, favoring that individual in particular over others who may have equal or greater needs (Batson et al., 1995; Batson & Ahmad, 2009), potentially posing a challenge to the collective action for the greater good (see Batson & Ahmad, 2009 for review). Thus, while empathy may indeed motivate prosocial behavior in some situations, it appears that it in others it is associated with a parochial favoritism that may at times run counter to the pursuit of equity and effectiveness in altruism.

Based on extensive research on empathy bias, numerous scholars in the fields of psychology and philosophy have suggested that when the aim of altruism is to maximize the preservation of lives (i.e., effectiveness), particularly those of individuals facing the most dire circumstances, without regard for their social proximity or relatedness (i.e., equitability), it is advisable to reduce the influence of empathy and instead prioritize rational, utilitarian reasoning to inform altruistic decisions (Caviola et al., 2021; J. Greene, 2014; J. D. Greene et al., 2001, 2004, 2008; Patil et al., 2018). Research in this vein shows that deliberative reasoning, rather than intuitive emotional responses, support altruistic and ethical decisions that benefit the greater good. Notably, this perspective pervades much of the philosophical discourse pertaining to effective altruism, which encourages combatting empathy with reasoning in order to achieve maximal gains in welfare through altruism (MacAskill, 2018; Singer, 2015, 2016). This

viewpoint contrasts starkly with prior research that emphasizes empathy's role in fostering altruism beyond parochial boundaries (Batson & Ahmad, 2009; Batson & Powell, 2003; Coke et al., 1978; Toi & Batson, 1982). Moreover, it presents a critical yet unanswered question: Does empathy *invariably* limit altruism to those who are closer and more similar, or can it be channeled as a force for promoting equitable and effective altruism?

To summarize, many psychologists and philosophers, including those associated with the effective altruism movement, have advocated for the promotion of equitable care and effective altruism through the enhancement of rational or utilitarian reasoning (Bloom, 2016; Caviola et al., 2022; J. Greene, 2014; J. D. Greene et al., 2004; Singer, 1981). Nonetheless, there is emerging evidence which corroborates seminal findings on the prosocial advantages of emotionality. This work suggests that empathy may also offer a viable path toward achieving the objectives of altruistic equity and effectiveness that are central to the effective altruism philosophy and social movement.

For instance, Buddhist meditation practices combine elements of both logical rationality (e.g., Chen & Jordan, 2020; Sodadee et al., 2018) and empathic emotion (e.g., Bayot et al., 2020; McRae, 2017) to promote universal and equitable care for all living beings, with some research having demonstrated that experts in Buddhist meditation practices show enhancements in empathic ability (Leung et al., 2013; Lutz et al., 2008). Furthermore, emerging research has shown that individuals who pledge to donate a portion of their lifetime income to charities–though not explicitly to charities which equitably and impartially maximize welfare–show heightened reasoning ability, but also heightened emotional abilities as well (Wilks et al., 2023). Notably, feeling balanced empathy for all, regardless of social proximity, is deemed morally superior to feeling biased empathy favoring close over distant others (Fowler et al., 2021). Overall, this research implies that the commonly observed parochial bias in empathic

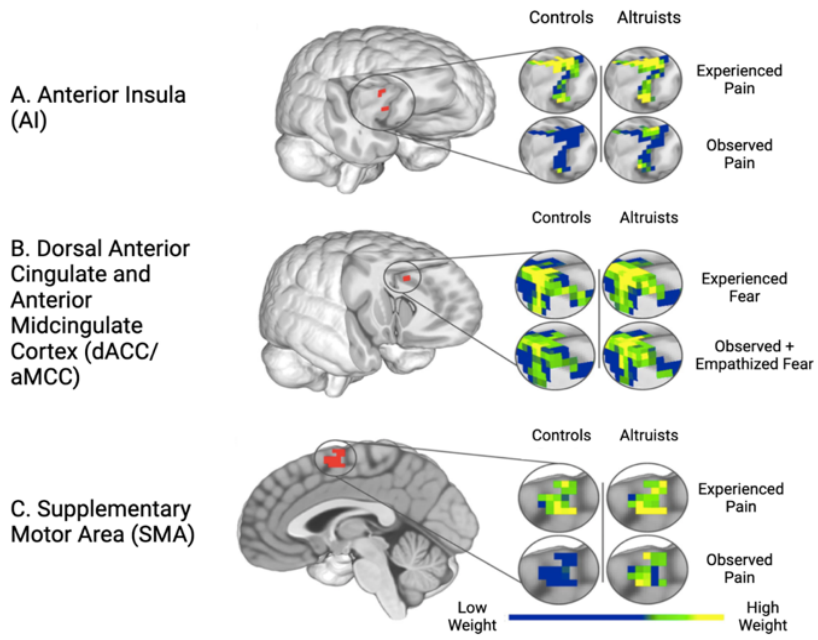
expressions may not be intrinsic to empathy itself. Instead, parochial bias may instead be rooted in the attitudes of individuals experiencing empathy. Accordingly, empathic expressions and behaviors tend to be more equitable in those who possess more equitable attitudes (Bruneau et al., 2017).

Another compelling line of evidence supporting the equitable utility of empathy derives from studies of extraordinary altruists who take on extreme risks and sacrifices on behalf of strangers. Research shows that individuals who altruistically donate organs or tissues to unknown recipients are not exceptionally rational. In fact, these individuals tend to exhibit heightened *empathic* responses to the suffering of strangers relative to typical adults, not less (Amormino, O'Connell, et al., 2022; Brethel-Haurwitz et al., 2018; Marsh et al., 2014; K. Vekaria et al., 2017; K. M. Vekaria et al., 2020). These altruists consistently show heightened patterns of empathic activation in various brain regions that support emotional responding (see Figure 1).

Specifically, extraordinary altruists exhibit heightened activity in the amygdala relative to the general population in response to cues signaling distress, like fearful facial expressions (Marsh, 2016a, 2016b; Marsh et al., 2014). They also exhibit increased self-other neural similarity in the empathy network (i.e., anterior insula and the dorsal anterior cingulate/anterior midcingulate cortex [dACC/aMCC]) when witnessing others experiencing pain (Brethel-Haurwitz et al., 2017, 2018; Crockett & Lockwood, 2018; Marsh, 2019). Results of structural imaging studies corroborate these functional imaging insights, revealing that extraordinary altruists show increased volume in the amygdala, but not other nearby subcortical regions (e.g., the hippocampus), further suggesting that heightened empathic abilities in particular are characteristic of this population. Extraordinary altruists may also exhibit more positively-valanced emotions, such as hope, when thinking about opportunities to help others, which

predicts less attrition from bone marrow donor registries (Brethel-Haurwitz et al., 2020; K. M. Vekaria et al., 2020). While altruistic organ donors might not always align with the principles of effective altruism–focusing explicitly on the prioritization of cost-effectiveness in charitable actions–effective altruists and altruistic organ donors together share a critical willingness to overlook parochial biases and sacrifice for strangers. Moreover, foundational scholars in the effective altruism movement have argued that donating organs to strangers also prioritizes effectiveness (Singer, 2015). Namely, providing a healthy organ to a stranger in failing health involves recognition that the organ would have a greater impact on welfare than keeping the organ for oneself when it isn't needed for survival. These commonalities, when taken together, hint that a common foundation may indeed underlie the cognitive, emotional, and moral architecture of both effective and extraordinary altruism. Nonetheless, this perspective has received little empirical attention (Law, Amormino, et al., 2024).

Figure 1. *Empathic neural activity in extraordinary altruism*



Note. *Contrary to claims favoring primarily rational mechanisms to counteract parochial bias in altruism, real-world extraordinary altruists, who donate organs to strangers, display heightened empathic neural activation. This includes elevated self-other overlap during experienced and observed pain (AI and SMA) and fear (dACC/aMCC), when compared to controls. This research used Multi-Voxel Cross-Classification (MVCC) with a Support Vector Machine (SVM) classifier to discern brain activity linked to pain and fear from non-aversive control trials. Higher SVM weights in voxel regions correspond to a stronger ability to differentiate experienced and observed pain and fear from non-aversive control trials using activity in those regions (Law, Amormino, et al., 2024; O'Connell et al., 2019). Altruists exhibit more overlap in activity in these high weight regions than matched control subjects. This figure has been adapted from Law, Amormino, et al. (2024) in accordance with author rights provided by Elsevier, which allow the author of this dissertation, who is also the author of the cited article, to include elements from the cited article without requiring additional permission (see https://www.elsevier.com/about/policies-and-standards/copyright#1-author-rights).*

In sum, current knowledge on the psychological factors that promote universal care and altruism is somewhat limited, deriving primarily from studies of living organ and marrow donors. Importantly, donating organs and tissue to strangers is only one form of distant altruism. At present it's unclear whether the mechanisms revealed in extraordinary altruists will also provide an account for other populations that practice universal impartiality in how they care for others, such as effective altruists who aim to rationally maximize gains in collective welfare through philanthropy. While prominent voices associated with the effective altruism movement have consistently advanced utilitarian reasoning over empathy as a better motivator for equitable altruism (Bloom, 2016; Caviola et al., 2021; Singer, 2015), emerging lines of peripheral evidence suggest that empathy and reasoning may not necessarily represent competing forces. Under

certain conditions, it's possible that empathy and deliberative reasoning may serve as complimentary avenues toward cultivating equitable and effective altruism and promote societal well-being and the greater good. Nevertheless, this perspective has been underemphasized and critically understudied in the existing literature at present.

1.3. The Current Research

In the current research, I integrate methodologies at the intersection of social and cognitive psychology with insights from ethical philosophy across three rounds of data collection involving three separate samples to address pivotal questions related to the underpinnings of equitable and effective altruism. These questions and the current aims of the present investigation are outlined in detail below:

1) Question 1: How do exceptionally caring populations and general population controls differ on measures of equitable and effective prosocial attitudes and behaviors (Question 1a) as well as on key psychological features that have been shown to underlie prosocial behavior in prior research (Question 1b)? This is explored in Phase 1, comprising Samples 1, 2 and 3.

2) Question 2: What psychological features are predictive of the equitable and effective altruistic tendencies of exceptionally caring populations? This is explored in Phase 2 in special population Samples 1 and 2.

3) Question 3: Do the same psychological features predict equitable and effective altruism in the general population? This is explored in Phase 2, in Sample 3, which comprises general population controls.

In Phase 1, to address Question 1a, I begin by conducting a series of analyses to validate two distinct altruistic populations–effective altruists (Sample 1) and extraordinary altruists (Sample 2)–as displaying a heightened emphasis on equity and effectiveness in their altruistic tendencies, relative to members of the general population. Specifically, I explore how these

samples score on a series of metrics and tasks assessing the prevalence of expansive prosocial attitudes and behaviors which either transcend typical parochial biases, prioritize effectiveness in resource allocations, or achieve both simultaneously. These scores are then compared between both special population samples and a sample of general population controls (Sample 3). These analyses serve two aims. The first is to assess whether the two special populations are indeed more equitable and effective in their prosociality relative to the general population, acting as a sanity check to confirm their exceptional altruistic tendencies. Additionally, as no empirical research has directly compared the prosocial architecture of these two groups, the second aim is to gain insight into how the scope of prosociality differs between them. To address Question 1b, I also compare the three participant groups on a host of psychological features that have been shown to predict prosocial behaviors in prior research. These features are outlined in detail below and their relationships with prosociality among the three populations are subsequently explored in Phase 2.

After investigating Questions 1a and 1b in Phase 1, in Phase 2, I begin by addressing Question 2. Specifically, I investigate the psychological architecture underlying the profound tendencies toward altruistic equity and effectiveness exhibited by effective and extraordinary altruists. The primary aim of Phase 2 is to examine associations between empathic ability, reasoning ability and their interaction with altruistic equity and effectiveness across the three samples. Nonetheless, a secondary aim is to evaluate whether and how moral beliefs and values as well as other-inclusive identity and other-oriented attitudes predict altruistic equity and effectiveness. The logic behind this secondary aim is spelled out in greater detail in section *1.4.3.* below.

Indeed, exploring the psychological profiles of exceptionally altruistic cohorts and examining the relationships between their psychological features and prosocial outcomes

provides insight into the foundations of equitable prosocial tendencies within altruistic populations. Yet, the methodological benefit of taking this special population approach extends to the insight it can offer into how altruistic equity might be better understood and later promoted in the general population. Prior research has employed similar special population methodologies to investigate the features which underlie variation in spatial reasoning ability in the general population by focusing on the special population of taxi drivers–experts in spatial navigation (Maguire et al., 2000). Similarly, much has been learned about the underpinnings of and normative variation in creativity by studying professionals in the arts and entertainment–experts in creative ability and imagination (Meyer et al., 2019; Orwig et al., 2023). Here, I apply a similar approach by studying experts in prosocial behavior. By doing so, I aim to (1) develop a deeper understanding of the features that characterize exceptionally altruistic minds and underpin altruistic equity and effectiveness within them, and (2) investigate whether these same psychological levers are present and potentially drive altruistic equity and the prioritization of effective impact in the general population.

As such, in Phase 2, I employ parallel methodology to explore the same relationships studied in the two special populations but within the general populace. This approach builds upon the findings from effective and extraordinary altruist samples to chart the psychological frameworks and underpinnings of equitable and effective altruism among the ordinary adults who comprise the majority, and thereby hold the greatest potential to significantly improve collective welfare through altruism. Specifically, by employing identical methodology across special and general population samples, I assess whether the features which guide prosocial behaviors in effective and extraordinary altruists are also prevalent in the broader population. The foundational premise of this research is that the psychological characteristics which underlie equitable and effective altruistic tendencies in exceptionally altruistic groups may be also be

present, although perhaps to a lesser degree, in the minds of every healthy adult. And, given the right circumstances, these features may be nurtured to cultivate more equitable and effective altruistic tendencies in the general population, ushering a future marked by greater care and compassion for all.

1.4. Hypotheses

Below I describe each of the primary hypotheses explored in the present investigation and their underlying motivation. Table 1 succinctly displays each of the hypotheses and how they relate to each research question and phase of analysis. Moreover, Table 1 describes the level of empirical support observed for each hypothesis in the current research.

Table 1

Hypotheses across phases and research questions and their levels of empirical support in the current investigation

Questions and Hypotheses	Level of Support
Phase 1	
Question 1a*: Do EAs, XAs, and Controls differ on signatures of equitable and effective altruism?*	
H1a: EAs and XAs, relative to controls, will score higher on measures of equitable and effective prosociality	**Mostly Supported:** EAs and XAs scored higher than controls on 10 out of 11 metrics, with XAs scoring lower than controls on effectiveness focus (EAIS).
H1b: EAs will score higher than XAs on prosociality measures capturing effectiveness prioritization	**Partially Supported:** This pattern was observed for 1 out of 2 outcomes, with EAs not differing significantly from XAs on the BDT.
Question 1b*: Do EAs, XAs, and Controls differ on signatures of psychological features associated with prosociality* (*reasoning, empathy, morality, other-Inclusive Identity)?*	

H2a: EAs and XAs, relative to controls, will score higher on measures of empathy (Alternative 1), reasoning (Alternative 2) or both (Alternative 3)	**Mostly Supported:** In general, the data supported **Alternative 1** for XAs and **Alternative 2** for EAs. Nonetheless, partial support for **Alternative 3** was observed on some metrics (e.g., HBT).
H2b (Open Hypothesis): No explicit predictions are made regarding differences between EAs and XAs on measures of empathy and reasoning	Where omnibus effects were observed, XAs scored higher on measures of empathy, while EAs scored higher on most (2 out of 3) measures of reasoning.
H2c: EAs and XAs, relative to controls, will score higher on prosocial and equitable moral values (care, fairness, impartial beneficence, moral expansiveness) and lower on moral values associated with ingroup favoritism (loyalty [MFQ], familial & group loyalty [MACQ])	**Partially Supported:** As predicted, EAs and XAs (relative to controls) scored higher on prosocial moral beliefs and values like moral concern on the MES and IB on the OUS, while scoring lower on the parochial moral value of loyalty on the MFQ. Intriguingly, however, (1) both special populations scored lower on the value of fairness on both MFQ and the MAC-Q while (2) XAs scored higher on group loyalty on the MAC-Q. No significant sample differences were noted on the MFQ value of Harm/Care.
H2d: EAs and XAs, relative to Controls, will score higher on measures of other inclusive identity (e.g., IWAH) and positive attitudes towards distant others (e.g., humanization)	**Mostly Supported:** For 9 out of the 14 relevant comparisons, EAs, XAs or both scored higher on measures of other inclusive identity and positive other-oriented values relative to controls.
Phase 2	
Question 2*: Which psychological features support equitable and effective altruism in EAs and XAs?*	
Question 3*: Which psychological features support equitable and effective altruism in Controls?*	
H3a: Across populations, empathy will predict higher scoring on measures of equitable and effective prosociality	**Mostly Supported:** Across samples, empathic ability was generally positively associated with equitable and effective prosociality.

H3b: Across populations, empathy will predict higher scoring on measures of equitable and effective prosociality	**Mostly Supported:** Across samples, reasoning ability was generally positively associated with equitable and effective prosociality.
H3c: Across populations, empathy and reasoning will have interactive effects on equitable and effective prosociality, such that the effects of each will be stronger at either higher (Alternative 1) or lower (Alternative 2) levels of the alternate capacity	**Partially Supported:** In some instances, reasoning ability and empathic ability interacted with respect to their associations with prosociality. Evidence for both **Alternatives 1 and 2** was observed.
H3d: Prosocial and equitable moral values (care, fairness, impartial beneficence, moral expansiveness) will associate positively, while moral values associated with ingroup favoritism (loyalty [MFQ], familial & group loyalty [MACQ]) will associate negatively with measures of equitable and effective prosociality	**Mostly Supported:** In most instances, the predicted effects were observed. Loyalty, however, at times showed unexpected positive relationships with prosocial outcomes.
H3e: Other-inclusive identity and positive attitudes towards distant others will associate positively with measures of equitable and effective prosociality	**Partially Supported:** Compassionate love for humanity (LHS) was the most consistent predictor of equitable and effective prosociality across samples. IWAH and tendencies to Humanize stigmatized groups (BDS) also predicted numerous prosocial outcomes across samples. Nonetheless, anthropomorphism (IDAQ), allo-inclusive identity (AIIS), and interdependence on others showed more nuanced relationships which at times differed between samples in magnitude and direction.

Note. *EA = "Effective Altruist", XA = "Extraordinary Altruist"*

1.4.1. Validating Special Populations as Prosocial Exemplars. Whereas effective and extraordinary altruists exhibit remarkable altruistic behaviors in daily life, I utilize diverse metrics in the research to follow as a means to quantitatively assess their pronounced prosocial tendencies in a controlled environment. The objective is to affirm the extraordinary prosocial tendencies of these groups. As such, I hypothesize that across these measures, both of the special population samples will score higher on average than the sample of general population controls

(Question 1a, H1a).

For instance, on one such metric, participants are asked to morally evaluate the philanthropic choices of fictional donors presented in narratives. In prior research utilizing tasks like this one, the broader population tends to find welfare-maximizing philanthropy benefitting a greater number of distant strangers less morally acceptable compared to philanthropy that benefits a smaller number of closer beneficiaries, such as relatives, friends, or fellow community members (Everett & Kahane, 2020; Law et al., 2022; McManus et al., 2020). Given the altruistic tendencies of the special populations in the real world, it is likely that the tendency observed in the general population with respect to these moral judgment vignettes will be eliminated or at least attenuated in the altruistic samples in the present investigation. That is, the two altruistic participant groups are expected to rate welfare-maximizing philanthropy as more morally acceptable than general population controls. Similarly, whereas the general population shows a tendency to discount the subjective value of outcomes for distant others on the social discounting task (Jones & Rachlin, 2006), extraordinary altruists who donate organs to strangers show a reduction in this effect (Vekaria et al., 2017), as do humanitarian aid workers and heroic rescuers (Rhoads et al., 2023). However, it remains unknown whether this attenuated social discounting effect is prevalent in the population of effective altruists as well, and whether both populations will also show greater expansive prosocial attitudes and behaviors on related tasks and measures.

I hypothesize that both exceptionally caring groups, relative to general population controls, will display an expansive inclination to help distant individuals (i.e., prioritize equity in altruism) and demonstrate beneficence toward causes that serve a greater number of beneficiaries per dollar donated (i.e., prioritize effectiveness in altruism), as captured by a diverse array of tasks and measures akin to the social discounting task which are described in detail in Chapter 2 (e.g., an in-house Behavioral Donation Task, the Equitable Altruism Vignettes Task [Law et al.,

2022], the Effective Altruism Interest Scale [EAIS; Caviola et al., 2022], the Longtermism Beliefs Scale [LBS; Syropoulos, Law, Kraft-Todd, et al., 2023]). Nonetheless, I also hypothesize that effective altruists will score higher on measures that uniquely capture the prioritization of effectiveness in prosociality (e.g., the Effectiveness Focus subscale of the EAIS) than extraordinary altruists, as this population is united by endorsing an explicit code of ethics which holds altruistic impact or effectiveness in the highest moral regard (Question 1a, H1b). In sum, the population validation measures I employ serve as a means to empirically capture the magnitude of equitable and effective altruistic tendencies in the two special populations relative to each other and to the general population. Moreover, as I elaborate on below, the inclusion of these metrics also provides a means to identify the cognitive, affective, and moral structures, as well as the features of other-inclusive identity and other-oriented attitudes, which most reliably predict these tendencies across and between populations.

1.4.2. Mapping the Cognitive and Affective Landscape of Equitable and Effective Altruism.

To date, most philosophers (Singer, 1981, 2015, 2016), economists (e.g., Mandler, 2020; Wade-Benzoni, 2017) and psychologists (e.g., Caviola et al., 2021, 2022; J. Greene, 2014; Huang et al., 2020) have put a primary focus on encouraging deliberative reasoning, emphasizing utilitarian goals in particular, as a means to reduce parochial bias and maximize collective welfare through altruism. This prevailing emphasis on reasoning is especially pronounced in the published philosophical and psychological literatures related to effective altruism. Some scholars go as far as framing emotion, and empathy in particular (Bloom, 2016; Prinz, 2011; Singer, 2015, 2016), as too inherently biased to be useful in expanding care and altruism to distant others and promoting collective welfare. In contrast, however, emerging laboratory research is beginning to show that another exceptionally caring population, extraordinary altruists who donate organs and body tissues to strangers, are not hyper-rational, but instead show enrichments

in emotions like empathy when compared to general population controls (Amormino, O'Connell, et al., 2022; Brethel-Haurwitz et al., 2018; Law, Amormino, et al., 2024; Marsh et al., 2014; K. M. Vekaria et al., 2020).

Research on altruistic behavior has been extensive, yet pivotal questions linger regarding the role of rational abilities in extraordinary altruism–the existing literature has predominantly focused on the emotional abilities of organ donors, but not their rational abilities. Furthermore, the psychological underpinnings of *effective* altruists' exceptional prosociality remain largely unexplored. Consequently, it's unknown whether the equitable and effective altruistic actions of these two special populations are driven *exclusively* by emotional responses–as seen in research on extraordinary altruists (which has largely overlooked the role of reasoning ability), the logical rationality championed by advocates of effective altruism (a population which has received little empirical attention to date), or a combination of both. By bringing both altruistic populations into the laboratory for the first time, I aim to gain novel insights into the drivers of both modalities of exceptional altruism, but also into the often-debated role of empathy in equitable prosociality more broadly (e.g., Bloom, 2016). That is, this research seeks to determine whether empathy *can* serve as a force for promoting equitable good or invariably serves as a factor that constrains altruism preferentially towards closer beneficiaries.

To start to develop a better understanding the psychological characteristics that drive altruistic equity in exceptionally altruistic individuals, in addition to comparing the three populations of interest on measures of equitable and effective prosociality (see section *1.4.1.* above), Phase 1 of the current investigation also involves comparing them on measures of empathic and reasoning ability, which have been shown to predict prosociality in prior research (Batson & Ahmad, 2009; J. Greene, 2014; Lindauer et al., 2020; Marsh, 2019; Schubert & Caviola, 2024). That is, in order to explore whether and how much empathy and reasoning *drive*

prosociality in these special populations, an initial starting point is first examining whether these capacities are in fact more pronounced in these populations relative to ordinary adults. I hypothesize that, when juxtaposed against a control group from the general population, both altruistic populations will manifest heightened capabilities in either reasoning (Question 1b, H2a, Alternative 1), empathy (Question 1b; H2a, Alternative 2), or both (Question 1b, H2a, Alternative 3). However, for reasons discussed below, it remains to be seen if the cognitive and emotional signatures will vary *between* these two altruistic populations (Question 1b, H2b [an open hypothesis]).

In Phase 2, I move on to explore Questions 2 and 3, which pertain to the relationships between empathic and reasoning ability (as well as the relationships between moral beliefs and values and differences in other-inclusive identity/other-oriented attitudes, as discussed below) with equitable and effective prosociality across the three populations of interest. Even despite prevailing perspectives on the constrictive impact of empathy on altruistic equity (Bloom, 2016), on the basis of extensive evidence suggesting that empathy supports stranger-directed altruism in organ donors (Marsh, 2019), I hypothesize positive relationships across populations between scores on empathic ability assessments with the suite of prosocial metrics outlined above in subsection *1.4.1.* (and in more detail in Chapter 2 below; Questions 2 and 3; H3a). In detail, I expect positive relationships between these variables even among members of the effective altruism movement, whose proponents routinely maintain that reasoning is a better-suited guide for equitable altruism than empathy (Caviola et al., 2021; Schubert & Caviola, 2024). I wager this hypothesis as the causes effective altruists regularly support often benefit *strangers*, which is a shared feature between effective and extraordinary altruists. And, extraordinary altruists' prosociality has been shown to be largely attributable to empathy (Marsh et al., 2019). Nonetheless, considering that organ donors prioritize equity in altruism–engaging in costly

prosocial acts that are impartial to the relatedness of beneficiaries–but not *necessarily* effectiveness, I argue that there is stronger preliminary support for the hypothesized positive relationships between empathic ability with prosocial outcomes that capture equity in altruism (e.g., the Expansive Altruism subscale of the EAIS, the Longtermism Beliefs Scale) than for prosocial outcomes that capture effectiveness (e.g., the Effectiveness Focus subscale of the EAIS, the Behavioral Donation Task). That is, the existing evidence, which comes primarily from the extraordinary altruist population of organ donors, speaks more directly to variations in willingness to help strangers than to variations in the weight and prioritization ascribed to the calculated impact of donations.

Similarly, on the basis of discourse in effective altruist philosophy (Singer, 2015) and suggestive evidence that reasoning ability is critical for choosing to prioritize cost-effectiveness and impact in altruistic resource allocations (Schubert & Caviola, 2024), I hypothesize positive relationships across populations between scores on reasoning ability assessments with the suite of prosocial metrics outlined above in subsection *1.4.1* above (Questions 2 and 3; H3b). In detail, I hypothesize positive relationships between these variables even among extraordinary altruists, who have primarily been studied in the context of their heightened empathic abilities. This is because donating an organ to a stranger can be viewed as involving an implicit prioritization of impact or effectiveness in altruism, with organ donors reasoning that the utility of an organ for a healthy individual is lower than the utility of the same organ for someone who needs it to survive. Indeed, this argument has in fact been advanced in popular writings related to effective altruism (Singer, 2015). Nonetheless, since preliminary evidence on the prosocial power of reasoning ability speaks more directly to the weight and prioritization ascribed to altruistic impact rather than the prioritization of equity or impartiality as to the relatedness of altruistic beneficiaries, I argue that there is stronger preliminary support for the hypothesized positive

relationships between reasoning ability and prosocial outcomes that capture effectiveness in altruism than for prosocial outcomes that capture equity.

To further understand the relationships between empathic and reasoning abilities on prosocial outcomes, I explore whether the two categories of predictors interact across the three populations. There are a handful of candidate possibilities for how such interactions may manifest. Specifically, the effect of reasoning ability on prosociality may depend on individuals already having a high capacity for empathy, as reasoning ability may help people identify effective ways to help others (Caviola et al., 2021), while empathy may motivate the desire to help in the first place or help equitably in particular (Marsh, 2019). Similarly, the impact of empathy on prosociality might rely on individuals already having a strong ability to engage in rational deliberation, as rational thinking may help individuals prioritize and act on their empathic emotions in more effective ways, leading to better prosocial outcomes. If either of these possibilities is supported, the effect of reasoning ability would be stronger at higher levels of empathic ability (Questions 2 and 3; H3c, Alternative 1). But an alternative possibility is that higher reasoning ability and empathic ability serve as protective factors against lower levels of the alternative capacity, such that the effects of either one would be stronger at higher levels of the alternative (Questions 2 and 3; H3c, Alternative 2). So, beyond merely examining how both emotional and rational processes relate to equitable and effective prosociality independently, this exploration also aims to uncover the nuanced ways these abilities might work together to enhance prosocial behavior.

Moreover, it's worth noting that the relative magnitude and predictive power of empathy and reasoning may differ between the two special populations. For instance, it's possible that extraordinary altruists only show enrichments in the empathic abilities on which they have been extensively assessed in prior research and that their altruistic tendencies are predicted by

variation in empathy to a greater extent than reasoning. Conversely, it's possible that effective altruists only show enrichments in the rational abilities upon which the effective altruism philosophy is premised and that their altruistic tendencies are predicted by variation in reasoning to a greater extent than empathy. If so, the present investigation would still hold implications for the role of both processes in altruistic behavior, suggesting that there may exist a plurality of routes towards equitable and effective altruism which manifest to different degrees in different individuals. Nevertheless, it's also possible that both empathic and rational ability are equally high in both special populations or that they both predict prosociality equally well across populations. I am open to each of these possibilities, but do not wager specific hypotheses as to which will be supported.

1.4.3. Assessing Associations between Moral Beliefs and Values as well as Other-Inclusive Identity and Other-Oriented Attitudes with Altruistic Equity and Effectiveness.

Beyond reasoning and empathy, the current investigation additionally allows for an assessment of the moral architecture of equitable and effective prosocial tendencies within exceptionally caring minds. Whereas much is known about the moral architecture behind prosociality in the general population, the proposed research serves as one of the first endeavors into the moral architecture of equitable altruism in exceptionally caring populations, with no empirical studies to my knowledge that have investigated the psychology of effective altruists on these dimensions in particular. As members of the two exceptionally caring populations examined in the present research show equitable and effective altruism through their prosocial behavior in vivo, I hypothesize that members of these populations will show an inclination towards equity and impartiality in their moral values as well (Question 1b, H2c).

Specifically, I hypothesize that members of the special populations, when compared to general population controls, will ascribe greater moral prioritization to the values of care and

fairness on the Moral Foundations Questionnaire (MFQ; Graham et al., 2011), fairness on the Morality as Cooperation Questionnaire (MACQ; Curry et al., 2019), impartial beneficence on the Oxford Utilitarianism Scale (OUS; Kahane et al., 2018), and possess larger circles of moral regard on the Moral Expansiveness Scale (MES; Crimston et al., 2016; Question 1b, H2c). I also expect to find that scores on these aforementioned metrics will correlate positively with scores on the prosociality measures described in section *1.4.1.* above (Questions 2 and 3, H3d). Conversely, I hypothesize that members of the special populations, compared to general population controls, will ascribe less moral prioritization to the values of loyalty on the MFQ, and familial and group loyalty on the MACQ, values which typically promote cooperation within groups, but at times can instead impede cooperation between them (Curry et al., 2019; Graham et al., 2011, 2017; Question 1b, H2c). Consequently, I hypothesize that scores on these aforementioned metrics will correlate negatively with scores on the prosociality measures described in section *1.4.1.* above (Questions 2 and 3, H3d).

Finally, I also included a battery of measures assessing the extent to which people identify with and harbor positive attitudes toward entities across a span of social and temporal distances (Eyal et al., 2008; Mentovich et al., 2016). In a general population sample (Law et al., 2022), my earlier research has shown that moral judgments of altruistic tradeoffs of social distance for gains in welfare track with scores on the Identification with All Humanity Scale (IWAHs; McFarland et al., 2012), which measures the extent to which people view all members of humanity as their ingroup. Specifically, those with higher scores on IWAHs find welfare-maximizing socially-distant altruism to be more acceptable than those with lower scores. Importantly, this effect holds even when helping distant others entails denying help to particularly socially closer others like family members and friends, suggesting that possessing a sense of self-identity that is inclusive of others may be fundamentally related to the way people

evaluate tradeoffs in social distance for gains in welfare. Similarly, work from other researchers has shown IWAH to be associated with more positive intergroup attitudes (Sparkman, & Hamer, 2020), and greater willingness to help migrants in need (Carmona et al., 2023).

Although other-inclusive identity has been shown to predict expansive altruism and positive attitudes towards care directed at distant others in the general population, no studies to date have investigated whether members of exceptionally caring populations feel a relatively stronger sense of identity with distant others. Furthermore, a number of other metrics in the existing literature tap into similar inclinations to identify with, feel connected to and harbor positive attitudes and feelings towards distant others. On the foundation of existing research connecting a sense of identity with all of humanity to expansive and equitable altruistic attitudes (e.g., Fowler et al., 2021; Law et al., 2022), I hypothesize that both special populations, relative to controls, will demonstrate more expansive other-inclusive identity and more positive attitudes towards distant others across the psychometric instruments I employ to measures these other-oriented phenomena (Question 1b; H2d). Further, I hypothesize that scores on these measures will be positively related to scores on the equitable and effective prosociality measures described in section *1.4.1.* above (Questions 2 and 3; H3e).

CHAPTER 2

METHODS AND ANALYSIS PLAN

Here, I first present the methods employed across Samples 1-3. Then, I continue on to provide an in-depth description of the analysis plan that was carried out across Phases 1 and 2 of the results, which are later presented in Chapter 3. To ensure adherence to open-science practices, the recruitment procedures, primary hypotheses and analysis plan have been pre-registered here: https://aspredicted.org/2LL_WYX. The deidentified data across Samples 1-3, a comprehensive "Analysis Script" featuring all code and output, and the comprehensive code used to create each graphic (where applicable, as Figure 16 was not created using code), are available for download on the Open Science Framework (OSF) website here: https://osf.io/gcy3s/?view_only=c263a61991b14ef29c1063f1251ec87a.

2.1. Methods

2.1.1. Participants. An a-priori power analysis determined that I would need to collect data from N = 319 subjects to detect an effect size of $r = 0.2$ (an effect size estimate commensurate with effect sizes found during piloting in a general population sample) with 95% power, assuming a two-tailed test and an alpha of 0.05. As documented in my prospectus, I set out to collect $N = 319$ usable cases of data for Samples 1 and 2, noting that due to the specialized nature of these two samples, it might prove untenable to obtain this many usable cases. As such, I planned to collect as many cases as possible in a 90-day period and to discontinue data collection once I had either (1) obtained $N = 319$ usable cases or (2) 90 days had passed from the start of data collection, adhering to whichever condition was satisfied first.

Recruitment for effective altruists was conducted through social media outreach via Facebook pages, Twitter accounts, Slack Channels, and LinkedIn groups associated with the effective altruism movement. Peter Singer, a prominent philosopher associated with the

movement, and Grace Adams, the Head of Marketing at Giving What You Can, an effective altruism-adjacent non-profit, agreed to launch social media campaigns to aid in the recruitment of members of the effective altruist population. At the beginning of the survey, participants were asked to indicate their level of familiarity with the effective altruism movement and whether they explicitly identified as effective altruists. Only participants who indicated explicit identification as effective altruists were invited to complete the entire survey using the Qualtrics platform. In a 90-day period, $N = 188$ self-identifying effective altruists from around the world completed the survey on Qualtrics and were rewarded $15 USD for their participation. Of these subjects, $N = 69$ were excluded based on a priori exclusion criteria: either failing to answer more than three of the 13 total attention checks correctly ($N = 52$) or possessing duplicate IP addresses ($N = 17$). Thus, Sample 1 comprised data from $N = 119$ self-identifying effective altruists. A sensitivity analysis revealed that this sample size achieves 95% power to detect an effect size of $r = 0.32$ or 80% power to detect an effect size of $r = 0.25$, specifying a two-tailed hypothesis test and an alpha of 0.05. Table 2 displays demographic information for all the samples recruited in the present research.

Recruitment for extraordinary altruists (i.e., altruistic organ and marrow donors) was similarly conducted online using the Qualtrics platform. However, instead of relying on social media outreach, participants from this population were recruited through an existing network of donors who have participated in research conducted by Abigail Marsh and her colleagues at Georgetown University. Only non-directed, living organ and tissue donors–participants who indicated having donated an organ (e.g., a kidney or portion of their liver) or bone marrow to strangers–were invited to complete the entire survey using the Qualtrics platform. In a 90-day period, $N = 66$ extraordinary altruists completed the survey on Qualtrics and were rewarded $15 USD for their participation. Of these subjects, only $N = 1$ was excluded for failing to answer

more than three of the 13 total attention checks correctly. No subjects possessed duplicate IP addresses. Thus, Sample 2 comprised data from $N = 65$ donors ($N = 57$ kidney donors; $N = 7$ kidney and liver double-donors; $N = 1$ kidney and marrow double-donor). A sensitivity analysis revealed that this sample achieves 95% power to detect an effect size of $r = 0.43$ or 80% power to detect an effect size of $r = 0.34$, specifying a two-tailed hypothesis test and an alpha of 0.05. While this sample may seem small, it is commensurate with the size of samples recruited in much of the existing and ongoing research targeting non-directed living donors and reflects the rarity of such profound displays of costly altruism (Amormino, O'Connell, et al., 2022; Brethel-Haurwitz et al., 2020; Law, Amormino, et al., 2024; Marsh, 2016a, 2019; Marsh et al., 2014).

After recruiting Samples 1 and 2, I set out to collect data from $N = 184$ English-speaking controls from the general population who (1) did not identify as effective altruists and (2) had not donated an organ or body tissue to a stranger in their lifetime. The target sample size of Sample 3 mirrored the combined size of the two special population samples (Samples 1 and 2). To ensure demographic similarity, recruitment for the control sample was stratified in two phases, as detailed below.

In the first recruitment phase, I aimed to collect data from $N = 119$ controls that were demographically similar to the effective altruist sample (Sample 1). To account for potential exclusions, I pre-registered that I would collect data from $N = 129$ subjects. Using the quota-sample functionality on Prolific Academic, I recruited a sample that was proportionately matched to Sample 1 on nationality (e.g., 38.6% U.S. American, 6.7% German, etc.), gender: 68.7% male, and age range: 18-60. One subject completed the survey without submitting for payment and was retained for analysis. Thus, $N = 130$ controls completed the survey. Of these subjects, N = 21 were excluded for failing to answer more than three of the 13 total attention checks correctly. No subjects possessed duplicate IP addresses. Thus, the portion of the control

sample matched to the effective altruist sample comprised data from $N = 109$ general-population controls.

For the second recruitment phase, I aimed to collect data from $N = 65$ subjects that were demographically similar to the extraordinary altruist sample (Sample 2), collecting data from $N = 75$ to allow for potential exclusions. Demographics quotas were set as 94% U.S. American and 6% Canadian; 66% Female; and an age range of 26-81. Because the extraordinary altruist sample (Sample 2) involved subjects from markedly fewer nations than the effective altruist sample (Sample 1), I was able to match on the additional dimension of race/ethnicity without exceeding Prolific's strata limit. Matching the proportions in Sample 2, the race/ethnicity quota was set to 91% White, 4.5% Asian, 3% Mixed/Multiracial, and 1.5% other. Again, one subject completed the survey without submitting for payment and was retained for analysis. Thus, $N = 76$ controls completed the survey. Of these subjects, N = 9 were excluded for failing to answer more than three of the 13 total attention checks correctly. No subjects possessed duplicate IP addresses. Thus, the portion of the control sample matched to the extraordinary altruist sample comprised data from $N = 67$ general-population controls.

In total, the control sample comprised data from $N = 176$ subjects. A sensitivity analysis revealed that this sample achieves 95% power to detect an effect size of $r = 0.27$ or 80% power to detect an effect size of $r = 0.21$, specifying a two-tailed hypothesis test and an alpha of 0.05. Across all samples combined, usable data were collected from $N = 360$ subjects ($N = 119$ effective altruists; $N = 65$ extraordinary altruists; $N = 176$ general population controls), affording 95% power to detect an effect size of $r = 0.19$ or 80% power to detect an effect size of $r = 0.15$, specifying a two-tailed hypothesis test and an alpha of 0.05. Table 2 displays a demographic breakdown across the three samples.

Table 2

Demographic information across Samples 1, 2 and 3.

Parameter	Sample 1: Effective Altruists	Sample 2: Extraordinary Altruists	Sample 3: Control
N_{Total}	119	65	176
N_{Male}	79	22	97
N_{Female}	36	42	77
N_{Other}	4	1	2
N_{White}	73	59	117
N_{Black}	17	0	22
N_{Asian}	13	3	19
$N_{Multiracial}$	8	2	11
$N_{OtherRace}$	8	1	7
$N_{UnitedStates}$	46	61	103
N_{Canada}	2	4	8
$N_{UnitedKingdom}$	13	0	17
$N_{AfricanContinent}$	21	0	13
$N_{AsianContinent}$	8	0	8
$N_{EuropeanContinent}$	21	0	19
$N_{SAmericanContinent}$	1	0	1
$N_{Australia\&NZ}$	7	0	7
M_{age}	31.9	53.0	37.2
SD_{age}	8.8	12.5	12.1
M_{income}[1]	3.0	4.5	3.0
SD_{income}	1.9	1.5	1.6
$M_{education}$[2]	7.9	8.0	6.9
$SD_{education}$	1.9	1.8	2.0

Note. *[1]Income was captured on a 1 "less than $25,000" – 6 "$150,000 or more" scale. On average, EAs and Controls fell into the "$50,000 – $74,999" bracket, while XAs fell into the "$75,000 – $100,000" bracket. [2]Education was captured on a 1 "no schooling completed" – 11 "doctorate degree" scale. On average, EAs and XAs obtained a bachelor's degree or equivalent, while controls obtained an associate's degree or equivalent.*

2.1.2. Design and Materials. Participants across samples were directed to a Qualtrics link containing the primary survey. After responding to a consent form, participants were presented with a battery of well-validated psychological metrics assessing empathic and reasoning ability, moral beliefs and values, and other-inclusive identity as well as other-oriented attitudes. These aforementioned measures served as the predictor variables across samples. As the outcome variables, participants also completed a battery of tasks and metrics that assessed equitable and effective prosociality (i.e., population validation measures). At the end of the survey, participants were asked a series of questions regarding demographic information and were debriefed (see Appendix E for the full text of demographic questionnaires and debriefing materials). The presentation order of the questionnaires in the survey, as well as the items in each questionnaire, was randomized to control for order effects.

To handle subject payments, participants in Samples 1 and 2 were presented with a passcode and a link to a payment survey. When participants visited the payment link, they were asked to provide an email address to which they wished their incentive (a $15 gift card) to be sent. Payment information was collected through a separate survey to ensure that any identifying information was not stored in the same dataset as participants' responses to the survey questions. Participants in Sample 3 were redirected to Prolific using a link that included information confirming completion of the survey as an embedded field and were remunerated directly through Prolific.

To uncover the cognitive, affective, moral, and attitudinal fingerprints of equitable and effective altruism across Samples 1-3, a large battery of measures assessing empathic ability, rational ability, moral values, and social attitudes (comprising other-inclusive identity and other-oriented attitudes) was employed. Table 3 presents key information on all of the measures that

were employed, including their names and abbreviations, sources, reliability (Cronbach's α), and scales of measurement.

Table 3

Key information on the measures that were employed in the current research across Samples 1-3 to capture empathic ability, reasoning ability, moral beliefs and values, other-inclusive identity/other-oriented attitudes as well as equitable and effective prosociality.

Empathic Ability (Predictor)	
Measure	***High Score Interpretation* (*Scale Points)***
Empathic Concern on the IRI IRI–EC; Davis, 1980 *IRI–EC Reliability:* α=0.86	Greater empathy for the suffering of others (1-7 Likert)
Emotionally Evocative Statements Task EEST; Marsh & Cardinale, 2012	Greater ability to identify emotions in statements (0-100; number of correctly identified emotions)
Toronto Alexithymia Scale TAS; Bagby et al., 1994 *TAS Reliability:* α=0.86	Greater ability to recognize one's own emotions (1-7 Likert)
Levenson Self-Report Psychopathy Scale LSRP; Levenson et al., 1995 *LSRP Primary Psychopathy Reliability*: α=0.88; *LSRP Secondary Psychopathy Reliability*: α=0.76	Lower primary and secondary psychopathy or Greater emotional ability (1-7 Likert)
Theories of Empathy Scale TES; Schumann et al., 2014 *TES Reliability:* α=0.91	Greater beliefs in the malleability of empathy (1-7 Likert)
Parochial Empathy Scale PES; Bruneau et al., 2017 *PES Reliability:* α=0.75	Greater empathy for distant targets (1-7 Likert)
Reasoning Ability (Predictor)	
Measure	***High Score Interpretation* (*Scale Points)***
Need for Cognition NFC; Cacioppo & Petty, 1982 *NFC Reliability:* α=0.90	Greater enjoyment from effortful cognitive endeavors (1-7 Likert)
Cognitive Reflection Test CRT; Frederick, 2005; Erlich et al., 2023	Greater tendency to override intuition and engage deliberative reasoning on word problems (1-7; number of correctly answered problems)

Heuristics and Bias Tasks HBT; Toplak et al., 2011	Greater utilization of algorithmic reasoning versus heuristics on word problems (1-7; number of correctly answered problems)
Rational Experiential Index REI; Pacini & Epstein, 1999 *REI Reliability: α*=0.89	Greater self-reports of reasoning ability (1-7 Likert)
Actively Open-Minded Thinking Scale AOT; Baron et al., 2022 *AOT Reliability: α*=0.85	Greater belief in changing one's mind on the basis of evidence (1-7 Likert)

Moral Beliefs and Values (Predictor)	
Measure	***High Score Interpretation (Scale Points)***
Moral Foundations Questionnaire MFQ; Graham et al., 2011 *MFQ Harm Reliability: α*=0.60 *MFQ Fairness Reliability*: *α*=0.68 *MFQ Loyalty Reliability*: *α*=0.77	Greater moral prioritization of the prosocial values of harm & fairness and the parochial value of loyalty (0-5 Likert)
Morality as Cooperation Questionnaire MAC-Q; Curry et al., 2019 *MAC-Q Fairness Reliability*: *α*=0.73 *MAC-Q Familial Loyalty Reliability*: *α*=0.91 *MAC-Q Group Loyalty Reliability*: *α*=0.86	Greater moral prioritization of the prosocial value of fairness and the parochial values of familial loyalty and group loyalty (0-100 Slider)
Moral Expansiveness Scale MES; Crimston et al., 2016 *MES Reliability*: *α*=0.95	Greater moral concern for targets across social distance or larger moral circles (0-4 Likert)
Oxford Utilitarianism Scale OUS; Kahane et al., 2018 *OUS Impartial Beneficence Reliability*: *α*=0.77 *OUS Instrumental Harm Reliability*: *α*=0.80	Greater moral endorsement of impartial beneficence (IB) and instrumental harm (IH), the sub features of utilitarianism–pre-registered hypotheses pertain to IB, while those pertaining to IH are exploratory (1-7 Likert)

Other-Inclusive Identity/Other-Oriented Attitudes (Predictor)	
Measure	***High Score Interpretation (Scale Points)***
Identification with All Humanity Scale IWAHs; McFarland et al., 2012 *IWAHs Reliability*: *α*=0.89	Greater sense of identity with all humans everywhere (1-5 Likert)
Compassionate Love for Humanity Scale LHS; Spreecher & Fehr, 2005 *LHS Reliability*: *α*=0.96	Greater feelings of compassion and love for all humans everywhere (1-7 Likert)
Self-Construal Scale SCS; Vignoles et al., 2016 *SCS Reliability*: *α*=0.81	Greater interdependence on others (1-5 Likert)

Allo-Inclusive Identity Scale AIIS; Leary et al., 2008 *AIIS Reliability*: α=0.91	Greater sense of connection to human and non-human entities across a range of relational distance (1-7 Graphic Scale)
Self-Other Four Immeasurables Scale SOFI; Kraus & Sears, 2009 *SOFI Reliability*: α=0.84	Greater positive feelings towards others over the course of the week preceding participation (1-5 Likert)
Individual Differences in Anthropomorphism IDAQ; Waytz et al., 2010 *IDAQ Reliability*: α=0.88	Greater tendencies to attribute mental states to non-humans (0-10 Likert)
Blatant Dehumanization Scale BDS; Kteily & Bruneau, 2017 *BDS Reliability*: α=0.90	Greater tendencies to humanize others or lesser tendencies to dehumanize others (0-100 Slider)

Measures of Equitable and Effective Prosociality (Outcome)

Measure	***High Score Interpretation* (*Scale Points)***
Moral Judgment Vignettes MJV; Law et al., 2022 *MJV Reliability*: α=0.93	Greater moral acceptability of welfare-maximizing altruism directed towards distant beneficiaries (**equitable and effective altruism captured**; 1-9 Likert)
Social Discounting Task SDT; Tuen et al., 2023	Greater prosociality towards socially distant others (**equitable and effective altruism captured**; 0-27; number of choices to direct resources to a target other than oneself)
Longtermism Beliefs Scale LBS; Syropoulos, Law, Kraft-Todd et al., 2023 *LBS Reliability*: α=0.98	Greater intergenerational concern (**equitable altruism captured**; 0-100 Slider)
Impact Legacy Motives Scale ILMS; Syropoulos & Markowitz, 2023 *ILMS Reliability*: α=0.89	Greater concern about impacting future generations positively (**equitable altruism captured**; 1-7 Likert)
Responsibility to Future Generations Scale RFG; Syropoulos & Markowitz, 2023 *RFG Reliability*: α=0.86	Greater perceived responsibility to protect future generations (**equitable altruism captured**; 1-7 Likert)
Behavioral Donation Task BDT; Law, 2024	Greater number of donations to effective vs. ineffective causes (**effective altruism captured**; 0-16; number of choices to donate actual resources [USD] to an effective versus ineffective cause)
Effective Altruism Interest Scale EAIS; Caviola et al., 2022 *EAIS Effectiveness Focus Reliability*: α=0.87 *EAIS Expansive Altruism Reliability*: α=0.82	Greater endorsement of **effective** (effectiveness focus subscale) and **equitable** (expansive altruism subscale) altruism (1-7 Likert)

Altruistic Enkrateia Scale AES; Law, 2024 *AES Reliability*: α=0.57	Greater strength of will in overcoming personal interest to prioritize equitable and effective altruism (**equitable and effective altruism captured**; 1-7 Likert)
Reported Real-World Charitable Action RWCA; Law, 2024	Greater proportions of income and time devoted to philanthropic causes (for income) and volunteerism (for time) in a given year (**neither equitable nor effective altruism captured**; 0%-100% Slider)

First, to assess the trait empathic ability, a multitude of validated psychometric instruments assessing empathic ability (and emotional ability more broadly) were included in the survey (see Table 3 for list of all measures and the Appendix for the full text of each; empathic ability measures can be found in Appendix A1). The Theories of Empathy Scale (TES; Schumann et al., 2014) was included to assess participants' lay theories regarding the malleability of empathy– the extent to which people believe empathic ability to be flexible and expandable. This scale included six items (e.g., Anybody can change how empathic a person they are.) and responses were recorded on a Likert-type agreement scale and averaged to form a composite score.

Scores on the Parochial Empathy Scale (PES; Bruneau et al., 2017) were assessed to measure the extent to which participants reported feeling empathy for relationally distant targets. Participants completed four items assessing perceptions of the amount of empathy they felt for the suffering of others in distant countries. Responses were recorded on a Likert-type agreement scale. Each participant's average was computed, with higher scores indicating lesser parochial bias or greater empathy for distant targets. Participants also responded to these items pertaining to others in their own country to serve as filler items, though no hypotheses were wagered for these items and they were not analyzed. Nonetheless, the data pertaining to these items are available on OSF.

The seven-item "empathic concern" (e.g., I am often quite touched by things that I see happen) subscale of the Interpersonal Reactivity Index (IRI; Davis, 1980) was assessed to measure participants' reported emotional empathy invoked by contemplating the suffering of others. Responses were recorded on a Likert-type agreement scale and averages were computed for each subject. The personal distress subscale was also included as filler, though no hypotheses were wagered for the present investigation and these items were not analyzed. Nonetheless, these data are available on OSF.

Furthermore, the Emotionally Evocative Statements Task (EEST; Marsh & Cardinale, 2012) was included in the survey to more directly capture participants' emotional ability. Participants read 100 statements, each designed to invoke a specific emotional response (e.g., Anger: "Don't you have any real friends?"; Disgust: "I haven't showered in days."; Fear: "I want to punch you."; Sadness: "I forgot your birthday."; Happiness: "You're amazing.") and indicated which emotion they believed corresponded best to each statement. The number of correct responses was calculated for each participant. This task measures trait ability to recognize emotions in text, and individuals higher in subclinical psychopathy tend to perform more poorly at this task compared to general population controls. As this measure taps emotional ability directly, it has advantages over self-report methodology which can be susceptible to social-desirability considerations. Values on the measure ranged from 0 (no emotions identified correctly) to 100 (100 emotions identified correctly).

In a similar fashion to the EEST, the Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994) was included in the survey to measure alexithymia–a reduced capacity to recognize one's own emotional states (e.g., "I am often confused about what emotion I am feeling."). Participants indicated their agreement with each item on a Likert-type scale and an average was computed

across the items. Prior to formal analysis the composite was reverse-coded to aid in interpretability, such that higher scores represent greater emotional ability.

Scores on the Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995) were captured to measure sub-clinical levels of "primary psychopathy" (i.e., manipulative behavior, interpersonal charm, reduced emotionality, inability to learn from past mistakes; e.g., "Making a lot of money is my most important goal.") and "secondary psychopathy" (i.e., similar to primary psychopathy but with additional propensities for emotional disturbances and maladaptive behavior; e.g., "I have been in a lot of shouting matches with other people."). Responses were recorded on Likert-type agreement scales and averages were computed for each 13-item subscale. In prior research, scores on these subscales have been shown to be robustly and negatively related to emotional ability and empathy in particular (see Marsh, 2019 for an extensive review of this literature). Like the TAS, scores on the subscales were reversed prior to analysis, such that higher scores correspond to lower psychopathy, or higher emotional ability.

In addition to capturing emotional ability, trait reasoning ability was assessed across Samples 1-3 with a battery of well-validated psychometric instruments (see Patil et al., 2018; reasoning ability measures can be found in Appendix A2). As a largely direct assessment of reasoning ability, responses to the Cognitive Reflection Test (CRT; Erlich et al., 2023; Frederick, 2005) were captured as the number of correct responses to seven word-problems. These word-problems required participants to override intuitive, incorrect solutions and instead engage deliberative reasoning to arrive upon the correct solutions, which are markedly less intuitive. Consider, for example, the following item from the CRT: "A racquet and ball cost $110 in total. The racquet costs $100 more than the ball. How much does the ball cost?" The intuitive response to this problem is the solution of $10. This response comes to mind readily yet is incorrect nonetheless. After all, $100 (the price of the racquet if the ball costs $10, assuming the total is

$110) is not $100 more than $10, but only $90 more. The correct solution to the problem is that the ball costs $5 and the racquet consequently costs $105, which is indeed $100 more than $5. This measure has been shown to be a robust and reliable measure of one's tendency to engage deliberative reasoning in problem-solving, with notable advantages over self-report measures which can be susceptible to social desirability influences (Bialek & Pennycook, 2018; Stagnaro et al., 2018).

The Need for Cognition Scale (NFC; Cacioppo & Petty, 1982). The 18-item NFC (e.g., "I would prefer complex to simple problems.") was included to measure participants' reported tendencies to enjoy engaging in effortful cognitive endeavors. Participants responded to each item on a Likert-type agreement scale and an average was computed for each participant.

Furthermore, the "reasoning ability" subscale of the Rational Experiential Index (REI; Pacini & Epstein, 1999) was included in the survey as well. Specifically, this 10-item subscale of the REI ("I have a logical mind.") assessed participants' self-reported ability to reason effectively. Participants responded to each item on a Likert-type agreement scale and an average was computed for each participant.

Scores on the Actively Open-Minded Thinking Beliefs Scale (AOT; Baron et al., 2022; Haran et al., 2013) were captured as yet another metric assessing reasoning ability. The 11-item AOT (e.g., "People should take into consideration evidence which goes against conclusions they favor.") measured the extent to which participants endorsed changing their minds to accommodate evidence which challenges their existing viewpoints. Participants responded to each item on a Likert-type agreement scale and an average was computed for each participant.

Finally, in a similar manner to the CRT, five Heuristics and Bias Tasks (Toplak et al., 2011) were included to capture reasoning ability in a more direct and objective manner than self-report methodology. Each of these longer-format word-problems were somewhat difficult to

solve, and, like the CRT, required overcoming intuitive responses with deliberative reasoning to do so correctly. The number of correct responses was assessed for each participant. Please see Appendix A2 for reference.

In addition to empathy and reasoning, a series of measures capturing moral beliefs and values were included as a third class of predictor variables (See Appendix B for full text of each measure). First, the 32-item Moral Foundations Questionnaire (MFQ; Graham et al., 2011) was employed to measure the extent to which participants reported morally valuing care (i.e., the reduction of harm), fairness, loyalty, respect for authority, and purity. Care and fairness are known as "individualizing" foundations, as they promote personal freedoms, and tend to be widely endorsed by most people. On the other hand, loyalty, authority and purity are known as "binding" foundations, as they promote intragroup stability, and tend to be endorsed more by people of conservative-leaning political ideology (Day et al., 2014; Waytz et al., 2013). The binding value of loyalty in particular has been shown to be associated with ingroup favoritism in prior research (Graham et al., 2017) Participants responded to each item on Likert-type scales and averages were computed for each foundation. Please see Appendix B for example items.

The Moral Expansiveness Scale (MES; Crimston et al., 2016), which measured the breadth and depth of participants' circles of moral regard (Singer, 1981), was included in the survey as well. Participants were asked to place a series of human and non-human entities at varying degrees of social distance within four boundaries of moral regard, placing the entities to whom they ascribe greater moral regard and more moral rights within the inner-most circles, while placing the entities to whom they ascribe comparatively lesser moral value and fewer moral rights in the outer-most circles. Typically, people place socially closer entities in the inner-most circles and socially-distant humans and non-human entities in the outer-most circles (Crimston et al., 2018), though this pattern is not universal (Rottman et al., 2021; Syropoulos,

Law, & Young, 2024d). Scores on the MES were aggregated across all entities on the scale for each participant, with higher scores representing more expansive moral circles.

In a similar fashion to the MFQ, scores on the subscales of the Morality as Cooperation Questionnaire (MAC-Q; Curry et al., 2019) were captured. This 42-item instrument measured the relative importance ascribed to the seven moral values encapsulated by the evolutionarily-informed Morality as Cooperation theory. Namely, this theory posits that human morality evolved primarily as a mechanism to promote ingroup cooperation and comprises the following values: familial loyalty, group loyalty, reciprocity, heroism, deference, fairness and property. Despite subtle distinctions, most of the moral values captured by the MAC-Q align with those assessed by the MFQ (Graham et al., 2011). Participants responded to each item on a Likert-type scale and averages were computed for each value.

To capture endorsement of utilitarian moral values, scores on the Oxford Utilitarianism Scale (Kahane et al., 2018) were assessed. Specifically, the five-item "impartial beneficence" (e.g., "It is morally wrong to keep money that one doesn't really need if one can donate it to causes that provide effective help to those who will benefit a great deal.") and four-item "instrumental harm" (e.g., "It is permissible to torture an innocent person if this would be necessary to provide information to prevent a bomb going off that would kill hundreds of people.") subscales were included to capture the perceived moral permissibility of two separate dimensions of utilitarianism: (1) helping others impartially for the greater good and (2) committing harm for the sake of the greater good. Participants responded to each item on a Likert-type scale and average composites were computed for each subscale separately.

A fourth battery of measures was included to capture a fourth and final class of predictors, broadly assessing the extent to which participants identify with and hold positive attitudes towards entities across a range of social distances (See Appendix C for full text of each

measure). In this vein, participants' tendency to identify with all of humanity (i.e., the extent to which they feel a part of and have concern for all humans everywhere) was measured using the Identification with all Humanity Scale (IWAHs; McFarland et al., 2012). Participants reported their agreement with nine items assessing their affiliation with all of humanity on a Likert-type scale. Averages for each were computed for each subject with higher scores corresponding to greater identification with all of humanity.

The 21-item Compassionate Love for Humanity Scale (LHS; Sprecher & Fehr, 2005) was measured to assess the extent to which participants reported feeling a compassionate sense of *love* for all of humanity. This measure was conceptually similar to the IWAHs, but exclusively captured attitudes towards all humans everywhere, rather than allowing for contrast between attitudes towards comparatively closer cohorts. Moreover, this measure captured reported feelings of compassion and love for all humans rather than a sense of identity. Participants responded to each item on a Likert-type scale and scores were averaged together.

Additionally, the 38-item Self-Construal Scale (SCS; Vignoles et al., 2016) was measured to capture the extent to which participants' reported interdependence between their own self-concepts and their relationships with others (e.g., "When someone in your family achieves something, you feel proud as if you had achieved something yourself."). Participants responded to each item on a Likert-type scale and scores on the scale were averaged together.

Furthermore, the 4-item Self-Other Four Immeasurables Scale (SOFI; Kraus & Sears, 2009) was included to capture the prevalence of positive emotions participants remembered feeling towards others over the course of the week preceding their participation in the study. Participants responded to each item on a Likert-type scale and scores were averaged together.

To assess the extent to which participants reported ascribing human-like capacities (e.g., a capability of engaging in insightful action and to experience emotions) to non-human entities,

scores on the Individual Differences in Anthropomorphism Questionnaire (IDAQ; Waytz et al., 2010) were assessed. Participants responded to each item on a Likert-type scale and scores were averaged together.

Conversely, scores on the Blatant Dehumanization Scale (Kteily & Bruneau, 2017) were assessed to capture participants' tendencies to ascribe fewer human-like qualities to out-group members. Participants were presented with an "ascent of man" diagram showing the evolutionary progression of hominin species from "ape-like" human ancestors to "advanced" modern humans. Participants were then asked to indicate on a slider scale where they would place six different groups of people spanning a range of racial, ethnic, and ideological classifications along the ascent of man continuum. Higher aggregate scores on the scale can be interpreted as representing a reduced tendency to dehumanize or a greater tendency to humanize stigmatized outgroups.

Finally, the Allo-Inclusive Identity Scale (AIIS; Leary et al., 2008) was included to measure the extent to which participants feel connected to a series of human and non-human entities spanning a range of social distances. Participants selected from a series of seven Venn diagrams displaying various degrees of overlap representing how connected they felt to each of the 16 entities on the scale. The Venn diagrams were coded from 1-7, with higher scores representing a greater sense of connection to others. Participants' scores were averaged to yield a composite measure.

I implemented diverse empirical measures capturing inclinations towards equitable and effective altruism in a laboratory environment (see Appendix D for full text of each measure) to (1) validate Samples 1 and 2 as exceptionally altruistic, (2) measure the frequency of equitable and effective altruistic tendencies in general population controls, and (3) serve as the primary outcomes of the present investigation to allow for an assessment of the psychological

architecture that underlies these tendencies across populations. Specifically, I sought to investigate whether both special populations would demonstrate an elevated degree of prosociality on these measures with respect to general population controls and whether effective altruists, relative to extraordinary altruists, would score higher on measures that uniquely capture the prioritization of effectiveness in altruism. I also assessed associations between the aforementioned measures of empathic ability, reasoning ability, morality, and other-inclusive identity/positive other-oriented attitudes with these measures to map the psychological underpinnings of equitable and effective altruism across populations.

As the first of the prosocial outcomes, participants were presented with short vignettes describing altruistic trade-offs involving hypothetical donors deciding whether to help socially distant others in severe need (e.g., strangers in distant countries undergoing greater hardship) or socially close others in comparatively less need (e.g., friends, family members, community members, and compatriots undergoing lesser hardship; Law et al., 2022). In all scenarios, hypothetical donors were depicted as choosing to help the distant strangers over the closer others. The task comprised 8 vignettes in total which varied on the basis of the social distance from the hypothetical donor of the person whom was denied aid. Specifically, the person denied aid was depicted as being a either compatriot, a community member, a friend, or a family member of the hypothetical donor, in two vignettes each. Participants were asked to rate the perceived moral acceptability of the hypothetical donor's decisions on a scale from 1 (completely unacceptable) to 9 (completely acceptable) and moral judgments were averaged across the eight items. Thus, this measure captured moral judgments of equitable (i.e., directed without regard to the social closeness of beneficiaries) and effective (i.e., aiding the more impactful cause) prosociality simultaneously.

In a similar manner to the task described above, the Social Discounting Task (Tuen et al., 2023) was included in the survey. On this task, participants chose between smaller rewards for themselves or larger rewards for others who varied on the dimension of social distance. The number of decisions (ranging from 1-27) participants made that favored larger rewards for others was computed to represent the extent of their tendencies to discount the subjective value of rewards to socially distant others relative to the subjective value they ascribed to rewards for themselves. Specifically, higher scores represent a *reduced* tendency to discount as a function of social distance. Like the moral judgment vignettes task described above, this measure simultaneously assessed tendencies towards equitable *and* effective altruism.

Participants' scores on the Longtermism Beliefs Scale (LBS; Syropoulos et al., 2023; e.g., "Positively influencing the long-term future is a key moral priority of our time.") were captured to measure endorsement of the principles of the longtermism ethical philosophy, which advocates beneficent concern for future generations regardless of their intergenerational distance from the present. This 7-item measure was displayed with respect to four different timeframes/timepoints varying across a range of temporal distances into the future (1,000, 10,000, 100,000, and 1,000,000 years from the present). Scores for the measure were captured on slider scales assessing agreement with each item. The average across all items and timeframes was computed, with higher scores representing a greater tendency to extend moral and ethical consideration to future generations in both the near and distal future (i.e., Intergenerational Concern; Law, Syropoulos, Coleman, et al., 2023; Law, Syropoulos, O'Connor, et al., 2024; Law, Syropoulos, & Young, 2023; Law, Syropoulos, Young, et al., 2024; Syropoulos, Law, & Young, 2024a, 2024b). Unlike the prior two outcomes, this measure uniquely captured tendencies towards equitable altruism without capturing effectiveness.

Similarly, concern for leaving behind a positive legacy after death was captured with the four-item Impact Legacy Motives Scale (ILMS; Syropoulos, Watkins, Goodwin, et al., 2023; Zaval et al., 2015; e.g., "I want to have an enduring positive effect on society.") and participants' sense of duty to protect future generations was captured with the four-item Responsibility to Future Generations Scale (RFG; Syropoulos & Markowitz, 2023; e.g., "People living today have an obligation to protect future generations, even if it means tightening our belts now."). Similar to the LBS, both measures addressed equitable prosocial intentions across temporal divides between donors and recipients. On both scales, participants responded to each item on a Likert-type scale and their scores were averaged across the items. Participants responses to the Reputation Legacy Motives Scale were also captured, though no hypotheses were wagered pertaining to these scores as this scale captures self-interest rather than prosociality. Data are available nonetheless on OSF.

Two outcomes were employed to specifically capture attitudes and behaviors aligning with effective altruism. First, I developed a Behavioral Donation Task to measure participants' preferences to donate real money to either effective charities (i.e., charities that have the potential to save lives) or ineffective charities (i.e., charities that have the potential to merely improve lives). I sourced four effective charities for use in the task from a recent catalog indexing those deemed most effective by experts who work at Give Well, an effective-altruism-aligned non-profit organization that evaluates charities on the basis of cost-effectiveness and impact (GiveWell, 2023). Specifically, I chose charities that have been vetted to save the greatest number of lives per dollar donated (e.g., the Against Malaria Foundation, Hellen Keller International). Additionally, I sourced four ineffective charities from recently published lists indexing those deemed to be the most ineffective compiled by Consumer Reports and Charity Navigator. Specifically, I chose charities that merely improve rather than save lives and have a

history of allocating a small percentage of funds to beneficiaries (e.g., the Make-A-Wish Foundation, Help Heal Veterans).

On the Behavioral Donation Task, participants completed 16 forced-choice trials. In each trial, they were presented with descriptions of two charities: one effective and one ineffective. These descriptions included information about each charity's effectiveness, specifically noting that one focused on saving lives while the other focused on merely improving them. Participants were given the choice to donate $1 per trial to either the effective or ineffective charity. They were informed that one of the 16 trials would be randomly selected, and a $1 donation would be made on their behalf by the research team to the organization they chose. The total number of donation decisions benefiting effective versus ineffective charities was then calculated. A greater number of donations to effective charities indicated a higher prioritization of effectiveness in the context of actual measured prosocial behavior.

As a self-report counterpart to the behavioral task described above, the "effectiveness focus" and "expansive altruism" subscales of the Effective Altruism Interest Scale (EAIS; Caviola et al., 2022) were included to capture alignment with two dimensions encapsulating the core principles of effective altruism: (1) the prioritization of cost-effectiveness in decisions to help others (i.e., effectiveness) and (2) the prioritization of relative need over social closeness considerations in decisions to help others (i.e., equitability). Participants responded to each item on Likert-type scales and their scores were averaged across the items on each subscale separately.

From the items included on the EAIS (Caviola et al., 2022), I also compute a third composite measure, aiming to capture the Aristotelian virtue of "Enkrateia" within the context of altruistic decision-making. This composite measure was based on two items from the "effectiveness focus" subscale and two items from the "expansive altruism" subscale.

"Enkrateia," often translated as "in-power," refers to a state of being in control over one's instincts or preferences (as described by Bobonich & Destrée, 2007; Gourinat, 2007). It signifies the ability to exercise self-control and make decisions aligned with one's values and goals, even in the face of conflicting desires or impulses. In contrast, "Akrasia" represents a weakness in willpower where individuals may act against their better judgment due to a lack of self-control or succumbing to immediate gratification. By including this third composite measure, I aimed to assess the degree to which participants exhibit "Enkrateia" in their altruistic decision-making, reflecting their ability to make principled altruistic choices that prioritize equitability and effectiveness, rather than succumbing to impulsive or conflicting desires.

The items included on the metric assessing altruistic Enkrateia were as follows: (1) "It would be the right choice to refrain from helping one person if that makes it possible to help a larger number of people," (2) "You should follow evidence and reason to do what is most effective, even if you emotionally prefer another option," (3) "I am willing to make significant sacrifices for people in need that I don't know and will never meet," (4) "We should put a lot of emphasis on the well-being of people who will live thousands of years from now, even relative to the well-being of people who live today." These statements reflect a disciplined approach to altruism, where one exercises self-control to make choices that are not necessarily emotionally satisfying but are rational and effective in doing the most good for the greatest number of people. This aligns with the concept of Enkrateia, involving control over personal inclinations for the sake of achieving a higher-order, more principled, and other-serving aim.

Finally, although not captured by metrics that have been validated in prior research, all subjects were asked at the end of the survey to indicate the percentage of income they donate to charity and the percentage of time they spend volunteering in a given year. Specifically, participants responded to the two following items: (1) "In a given year, what percentage (out of

100) of your yearly income do you donate to charity?" And (2), "In a given year, what percentage (out of 100) of your time do you devote towards volunteering to help others?" Responses were captured on 0-100 slider scales. Both of these metrics were included to capture reported levels of engagement in real-world charitable action, without regard to whether such actions are equitable or effective. As an exploratory hypothesis, I reasoned that both special populations would report greater levels of charitable action relative to general population controls.

At the end of the survey, a battery of standard demographic items was captured (see Appendix E for full text of the demographics survey).

2.2 Analysis Plan.

2.2.1. Phase 1 Analyses. To begin my formal investigation into the psychological architecture of equitable and effective prosociality, I compare the three populations of interest–self-identifying effective altruists who dedicate their lives to maximizing the effective impact of philanthropy, extraordinary altruists who have donated organs and tissues to strangers in need, and ordinary people from the general population who are matched on key demographic dimensions to members of the special populations–on laboratory metrics that capture either the prioritization of equity, effectiveness, or both in the context of prosocial attitudes and behaviors (Question 1a). The aims of addressing Question 1a are to validate both of the special populations as exceptionally altruistic relative to the general population and to further explore whether and how the two special populations differ in their prosocial tendencies from one another. I hypothesized that both of the special populations would score higher on the suite of metrics capturing effective and equitable prosociality relative to the general population (H1a) and that effective altruists would score higher than extraordinary altruists on measures specifically

capturing the prioritization of effectiveness or impact. A full list of each hypothesis is provided in Table 1 and information regarding each measure is provided in Table 3.

To address the primary hypotheses under Question 1a, I will first estimate one-way ANOVAs (Analysis of Variance) specifying Sample as the grouping variable (Sample 1 vs. Sample 2 vs. Sample 3) for each of the prosocial outcomes. Subsequently, I will examine pairwise comparisons using post-hoc t-tests which apply the Bonferroni alpha correction (i.e., dividing alpha by the number of post-hoc comparisons that will be run following each ANOVA) to guard against rising false positive rates. Since each outcome will require three follow-up tests, the desired familywise error rate of 0.05 will be divided by three, such that the significance level of each follow-up test will be set to 0.017. This will ensure that the probability of a false positive across all tests pertaining to each outcome will remain below the desired error rate of 0.05.

Two of the prosociality metrics also allow for finer-grained evaluation of how prosocial attitudes fluctuate as a function of beneficiary temporal (the Longtermism Beliefs Scale or LBS) or social (the Moral Judgment Vignettes Task) distance. On the LBS, my existing research has shown consistently that members of the general population tend to ascribe lesser beneficent intergenerational concern to targets in future generations as they become more distal from the present (Law, Syropoulos, & Young, 2023; Syropoulos, Law, et al., 2023; Syropoulos, Law, Kraft-Todd, et al., 2024b; Syropoulos, Law, Mah, et al., 2024c; Syropoulos, Law, & Young, 2024e, 2024d). Likewise, in my prior research employing the Moral Judgment Vignettes Task, subjects in the general population tend to report systematically more negative moral judgments of altruism directed towards distant strangers as a socially closer alternative recipient who is denied money becomes progressively socially closer to the hypothetical donor (Fowler et al., 2021; Law et al., 2022).

Because both special populations show profound equity and psychological distance-transcendent tendencies in their altruistic behaviors in daily life, it raises the question of whether these populations, relative to the general population, will show patterns of intergenerational concern and moral judgement that are less sensitive to the distance of beneficiaries on the two metrics discussed above. For exploratory purposes, I test these possibilities using 2 (Between-Subjects: Sample 1 vs. Sample 2 vs. Sample 3) by 4 (Within-Subjects: Beneficiary Distance) mixed-ANOVAs. For the LBS, Beneficiary Distance pertains to the temporal distance of the future generation beneficiaries that participants are asked to indicate their levels of concern for (1,000 vs. 10,000 vs. 100,000 vs. 1,000,000 years from the present). On the Moral Judgment Vignettes Task, Beneficiary Distance pertains to the social distance of the beneficiary who is denied aid in the commission of welfare-maximizing altruism directed towards a beneficiary that is more socially distant from the donor. My exploratory hypothesis is that I will observe an interaction between Sample and Beneficiary Distance on the outcome of intergenerational concern on the LBS and the outcome of moral acceptability on the Moral Judgment Vignettes task. Specifically, I expect that both special population samples will report concern and acceptability judgments that are more equitable (i.e., less sensitive) to variations in levels of Beneficiary Distance.

Also for exploratory purposes, I employ profile analysis using multivariate analysis of variance (MANOVA), a robust statistical technique that offers distinct advantages over traditional ANOVA-based approaches. MANOVA allows for a comprehensive assessment of group differences by simultaneously examining multiple dependent variables. Here, I specifically examine differences between the three participant groups of interest on prosociality, empathic ability and reasoning ability in a unified framework. This approach offers a holistic understanding of how the exceptionally caring populations (effective altruists and extraordinary

altruists) compare to each other as well as to general population controls across all of these outcomes taken together. In other words, profile analysis efficiently utilizes the collected data and enhances group comparisons by focusing on overall patterns of scores *across* variables, rather than isolated means for *each* variable.

Next, after examining the primary and exploratory hypotheses under Question 1a, I proceed to address the hypotheses under Question 1b. Here, I investigate whether and how the special populations differ from the general population and from each other on the suite of measures capturing psychological features that have been shown to predict prosociality in the general population (i.e., empathic ability, reasoning ability, moral beliefs and values, and other-inclusive identity/other-oriented attitudes). I test hypotheses H2a-H2d under Question 1b using the same procedure used to test the primary hypotheses under Question 1a described in detail above (i.e., one-way ANOVAs comparing the three samples with Bonferroni-corrected post-hoc t-tests for pairwise comparisons). As detailed in Table 1, I hypothesized both special populations would score higher than general population controls on measures of empathic ability (H2a, Alternative 1), reasoning ability (H2a, Alternative 2) or both (H2a, Alternative 3).

Moreover, I specified that I would investigate how the two special populations differ on these dimensions without wagering a specific hypothesis for how this may play out (Question 1b, H2b [An Open Hypothesis]). Namely, because effective altruists have yet to be explored on either dimension, it's at present unclear whether they will show similar enrichments to extraordinary altruists on the dimension of empathy (see Law, Amormino, Marsh, et al., 2024 for review) or enrichments in reasoning, as is advanced in effective altruist philosophical discourse (Schubert & Caviola, 2024; Singer, 2015). Moreover, because extraordinary altruists have been almost exclusively studied on the dimension of empathy, it's unclear whether they will show enrichments in reasoning ability. Nonetheless, proponents of the effective altruism philosophy

contend that reasoning is a better guide for equitable and effective altruism than empathy and some empirical research hints that this phenomena is at times observed in the general population (Caviola et al., 2021; Patil et al., 2021). Thus, this possibility remains open. Alternatively, it's also likely that both populations will score equally high on both of these psychological capacities relative to the general population but will not meaningfully differ from each other.

With regard to moral beliefs and values, I hypothesized that both special populations would score higher than the general population on measures of values that have been shown to predict expansive prosocial tendencies–the values of care and fairness on the Moral Foundations Questionnaire (MFQ; Graham et al., 2011), fairness on the Morality as Cooperation Questionnaire (MACQ; Curry et al., 2019), impartial beneficence on the Oxford Utilitarianism Scale (OUS; Kahane et al., 2018), and larger circles of moral regard on the Moral Expansiveness Scale (MES; Crimston et al., 2016). Likewise, I hypothesized that both special populations would score lower than the general population on measures of values the have been shown to be associated with cooperation within groups but often impede cooperation between them–the values of loyalty on the MFQ as well as familial and group loyalty on the MACQ. These hypotheses have been collectively labeled H2c. Finally, on the measures capturing other-inclusive identity and positive other-oriented attitudes, I hypothesized that both special populations would score higher than general population controls (Question 1b, H2d).

2.2.2. Phase 2 Analyses. After comparing the special and general populations on prosocial signatures and identifying the cognitive, affective, and moral structures present within the minds of each group, the objective of Phase 2 is to map the psychological architecture that underlies equitable and effective prosocial tendencies. Specifically, I investigate the extent to which these cognitive, affective, and moral structures as well as variation on measures of other-

inclusive identity and other-oriented attitudes, predict equitable and effective prosociality in the special populations (Question 2) and the general population (Question 3).

Phase 2 allows for a deep examination into the psychological *drivers* of equitable and effective altruism within the exceptionally caring minds, but also the general populace. By examining altruistic tendencies in ordinary individuals alongside those in exceptional altruists, I aim for a holistic perspective on the potentially multifaceted nature of altruistic behavior. Much of the published literature surrounding the psychology of effective altruism has hailed deliberative reasoning as most conducive to effectiveness in prosocial decisions (Caviola et al., 2021; J. D. Greene et al., 2004; Huang et al., 2020; Singer, 2015, 2016). Certain scholars in psychology and philosophy even recommend downregulating empathy to reduce biases in altruistic choices (Bloom, 2016; Caviola et al., 2022; Singer, 2015; Wilks et al., 2023). Yet, recent studies on extraordinary altruists–such as non-directed organ donors–challenge this view. These altruists tend to activate emotional neural pathways and show empathic responses when faced with strangers' suffering (e.g., Amormino et al., 2022; Marsh et al., 2014; O'Connell et al., 2019; Rhoads et al., 2023). Moreover, they consistently demonstrate heightened levels of empathic ability compared to typical adults.

Findings like those from the literature on extraordinary altruism converge with separate lines of evidence indicating that people weigh emotions more heavily into prosocial decisions than effectiveness-relevant information (Berman et al., 2018; Montealegre et al., 2020). Furthermore, recent research suggests empathy can be expanded beyond parochial boundaries (Fowler et al., 2021; Weisz & Cikara, 2021). Taken together, these insights allude to the possibility that empathy and reasoning might coexist harmoniously, rather than in opposition, in the context of efforts to amplify the equitability and effectiveness of altruistic resource allocations. Should the present data uphold this hypothesis in both exceptionally altruistic and

general population cohorts, it could reshape our understanding of altruism, suggesting that emotional and cognitive routes might synergistically cultivate expansive prosocial behavior, rather than exert competing forces that contract and expand the scope of beneficence, respectively.

Beyond theoretical implications, integrating data from the general population with data collected in altruistic populations offers the potential for tangible practical benefits. By gaining insight into whether and how capacities in emotional and rational ability, variation in moral beliefs, and expansive outlooks on self-identity and other-oriented attitudes underlie altruistic equity and effectiveness across the range of prosociality that humans are capable of, this research lays the groundwork for the development of future interventions which may revolutionize global welfare through altruism. Specifically, if I discern that the primary psychological characteristics that steer expansive altruism in exceptionally caring individuals mirror those underpinning expansive altruism in the general public, it implies these traits could be nurtured to amplify altruistic tendencies more broadly. The insights from this research would then pave the way for subsequent investigations targeting the design of scalable interventions grounded in the present findings.

I begin addressing the aims of Phase 2 by uncovering the affective and cognitive architecture of equitable and effective altruism across the three samples. Specifically, I first examine associations between scores on the rational and empathic ability measures with scores on the measures capturing variation in prosociality using bivariate correlational and multiple regression analyses. This approach allows an assessment of the extent to which cognitive and emotional ability predict altruistic behavior across a variety of domains in the three populations of interest. Moreover, the examination of these predictors in the context of multiple regression allows insight into which capacities *uniquely* account for variance in the outcomes. Here, I

hypothesized that across populations, both empathy (Questions 2 and 3, H3a) and reasoning (Questions 2 and 3, H3b) would predict higher scoring on measures of equitable and effective prosociality.

Next, before investigating potential independent and interactive effects of reasoning ability and empathic ability on the suite of prosocial outcomes, I first follow a systematic analytic approach to reduce the dimensionality of the large battery of predictors. As a preliminary step, I conduct exploratory factor analysis (EFA) to examine the factor structure among the predictors within each of the two categories and to determine which predictors should be retained in the subsequent step to maximize fit. Specifically, in this initial stage, I identify the predictors with factor loadings above 0.50, retaining these going forward. Moreover, I use the maximum likelihood extraction method with oblique rotation (direct oblimin method)–selected for its efficiency in estimating factor loadings and its ability to facilitate confirmatory factor analysis, even among correlated factors (Jackson, 2005).

Next, I conduct confirmatory factor analysis (CFA) using the maximum likelihood extraction method on the retained predictors within each category (i.e., empathic ability and reasoning ability), confirming the factor structure identified during EFA. To derive individual participant scores for each of the two expected factors, I first z-transform each component prior to composing averages for each subject in order to ensure that each component is measured along the same scale. And, finally, I conduct reliability analyses (*Cronbach's* α) to assess the internal consistency prior to utilizing the factor scores in the subsequent analyses. These methodological steps enable me to encapsulate the core aspects of reasoning and empathic ability elegantly, in two distinct composite measures. Thus, the primary objective of employing factor analysis is to reduce the possibility of problematic multicollinearity when examining the

independent and interactive effects of each predictor and to maximize the interpretability of these findings.

Once these two composites are established, I proceed with separate regression analyses to assess the individual contributions of empathic and reasoning ability on each prosocial outcome. These analyses allow me to identify any main effects associated with each predictor. Following the examination of main effects, I explore potential interaction effects by introducing interaction terms between the extracted factors capturing variation in empathic and reasoning ability. These interaction terms are included in the regression models to evaluate whether the combined influence of reasoning and empathy differs significantly from what would be expected based on their individual contributions. This comprehensive approach provides insights into the unique *and* combined impact of cognitive and affective ability on prosocial behavior, while the initial dimension reduction step merely enhances the clarity and interpretability of the findings by reducing the number of predictors included in these models. I hypothesized that, across populations, empathy and reasoning would have interactive effects on equitable and effective prosociality, such that the effects of each would be stronger at either higher (Questions 2 and 3; H3c, Alternative 1) or lower (Questions 2 and 3, H3c, Alternative 2) levels of the alternate capacity.

Next, to uncover the moral architecture of equitable and effective altruism across populations, I proceed by examining associations between scores on the moral beliefs and values measures with scores on the measures capturing variation in prosociality using bivariate correlational and subsequently multiple regression analysis. Here I assess the individual and unique contributions of each of these predictors on the suite of prosocial outcomes. I hypothesized that prosocial and equitable moral values (care, fairness, impartial beneficence, moral expansiveness) would associate positively, while moral values associated with ingroup

favoritism (loyalty [MFQ], familial & group loyalty [MACQ]) would associate negatively with measures of equitable and effective prosociality (Questions 2 and 3, H3d).

To assess whether including distant others in one's self-concept or harboring favorable attitudes towards distant others influences prosocial tendencies, I examine associations between scores on the other-inclusive identity/other-oriented attitude measures with scores on the measures capturing variation in prosociality, mirroring the procedure described above for the predictors related to moral beliefs. Here, I hypothesized that other-inclusive identity and positive attitudes towards distant others would associate positively with measures of equitable and effective prosociality (Questions 2 and 3, H3e).

CHAPTER 3

RESULTS

Chapter 3, the results section, proceeds by first evaluating questions under Phase 1, which pertain to comparisons across Samples 1, 2, and 3 on measures of equitable and effective prosociality, empathic and reasoning ability, moral beliefs and values, as well as other-inclusive identity and other-oriented attitudes. Then, in the second portion of Chapter 3, I evaluate questions related to the architecture that underlies prosociality within Samples 1, 2 and 3, evaluating empathic and reasoning ability, moral beliefs and values, as well as variation in other-inclusive identity and other-oriented attitudes as predictors of the prosocial outcomes.

3.1. Phase 1 Results[2]

3.1.1. Population Differences in Equitable and Effective Prosociality. To formally investigate the architecture of prosociality among exceptional altruists and the general population, I began by comparing the three groups on their reported engagement in real-world charitable actions. Effective altruists (EAs) aim to do good primarily through financial donations to impactful causes (MacAskill, 2016; Singer, 2015). On the other hand, extraordinary altruists (XAs), who earn their designation through substantial sacrifices such as donating parts of their bodies to strangers, are not necessarily unified by a commitment to financially benefiting others like EAs (Marsh, 2019). Moreover, altruism and philanthropy take on numerous forms in ordinary adults, from financial contributions to volunteering time for the benefit of others. To capture prosociality in a way that does not rely solely on financial contributions and acknowledges the diverse ways people engage in it in the real world, I asked subjects across

[2] The analyses for Phase 1 in the main text compare each special population to the entire control sample. An exploratory series of analyses, available in the "Analysis Script" on the OSF, also compares each special population to the subset of the control sample matched to each special population's demographics. The general patterns remain consistent across both series of analyses, but the findings presented here focus on the full control sample due to the greater statistical power these comparisons afford.

samples to report both the proportion of income they allocate to charitable causes and the proportion of time they spend volunteering in a given year.

One-way ANOVAs with Bonferroni corrected pairwise comparisons were estimated to compare the three samples across the metrics capturing real-world charitable action. The results empirically confirmed the exceptional altruism displayed by both special populations in vivo. Specifically, EAs and XAs reported (1) donating significantly more financial resources to charitable causes and (2) volunteering a greater amount of time for the benefit of others. See Table 4 for summary statistics from the inferential tests and Figure 2 for a graphical depiction of these findings. Since monetary contributions to charity and the availability of time for volunteering may depend on income, I conducted exploratory analyses, this time including income variation as a statistical control. The "Analysis Script" on the OSF displays the results, showing that the patterns remain unchanged whether or not income variation is controlled.

The effects described above were not only significant, but large, with *Cohen's d*s ranging from 0.580 to 1.013. Furthermore, EAs reported donating more income than XAs (an average of 15.5% for EAs versus 10.5% for XAs), although there was no significant difference in the amount of time they reported volunteering. These initial findings serve as a sanity check, providing evidence that both special populations are indeed more generous than controls. They also suggest that the exceptional altruism of XAs is not limited to engaging in a single profound act of altruism, such as donating an organ. Rather, it runs deeper, as they report devoting substantially more of their time and money to benefit others than controls in their daily lives.

Table 4

Results from one-way ANOVAs with post-hoc tests employing the Bonferroni alpha correction comparing Samples 1-3 on measures capturing prosociality.

Sample	**Prosociality**

Percentage of Income Donated (RWCA, Overall Prosociality)[1]

Omnibus	$F(2, 357) = 37.2$, $p < .001$, $\eta^2_p = .172$
EA vs. Control	$t(357) = 8.53$, $p < .001$, *Cohen's d* = 1.013
XA vs. Control	$t(357) = 4.00$, $p < .001$, *Cohen's d* = 0.580
EA vs. XA	$t(357) = 2.80$, $p = .016$, *Cohen's d* = 0.432

Percentage of Time Spent Volunteering (RWCA, Overall Prosociality)[1]

Omnibus	$F(2, 357) = 23.9$, $p < .001$, $\eta^2_p = .118$
EA vs. Control	$t(357) = 6.51$, $p < .001$, *Cohen's d* = 0.772
XA vs. Control	$t(357) = 4.36$, $p < .001$, *Cohen's d* = 0.633
EA vs. XA	$t(357) = 0.90$, $p > .999$, *Cohen's d* = 0.139

Moral Judgment Vignettes (MJV, Equitable & Effective Prosociality)

Omnibus	$F(2, 357) = 20.20$, $p < .001$, $\eta^2_p = 0.102$
EA vs. Control	$t(357) = 5.45$, $p < .001$, *Cohen's d* = 0.647
XA vs. Control	$t(357) = 4.90$, $p < .001$, *Cohen's d* = 0.711
EA vs. XA	$t(357) = -0.42$, $p > .999$, *Cohen's d* = -0.065

Social Discounting (SDT, Equitable & Effective Prosociality)

Omnibus	$F(2, 357) = 40.40$, $p < .001$, $\eta^2_p = 0.185$
EA vs. Control	$t(357) = 5.55$, $p < .001$, *Cohen's d* = 0.659
XA vs. Control	$t(357) = 8.51$, $p < .001$, *Cohen's d* = 1.235
EA vs. XA	$t(357) = -3.73$, $p < .001$, *Cohen's d* = -0.576

Altruistic Enkrateia (AES, Equitable & Effective Prosociality)

Omnibus	$F(2, 357) = 38.1$, $p < .001$, $\eta^2_p = .176$
EA vs. Control	$t(357) = 8.32$, $p < .001$, *Cohen's d* = 0.987
XA vs. Control	$t(357) = 0.24$, $p > .999$, *Cohen's d* = 0.035
EA vs. XA	$t(357) = 6.17$, $p < .001$, *Cohen's d* = 0.952

Intergenerational Concern (LBS, Equitable Prosociality)

Omnibus	$F(2, 357) = 6.02$, $p = .003$, $\eta^2_p = 0.033$
EA vs. Control	$t(357) = 3.15$, $p = .005$, *Cohen's d* = 0.373
XA vs. Control	$t(357) = 2.42$, $p = .048$, *Cohen's d* = 0.351
EA vs. XA	$t(357) = 0.14$, $p > .999$, *Cohen's d* = 0.022

Impact Legacy Motives (ILMS, Equitable Prosociality)

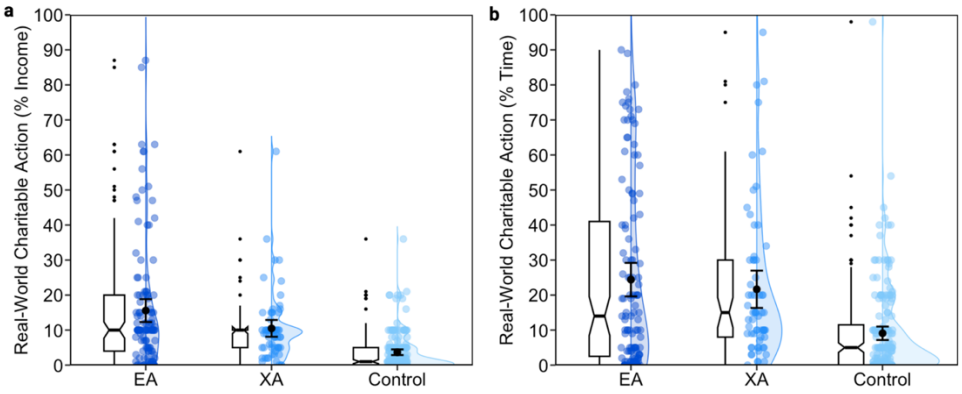
Omnibus	$F(2, 357) = 16.80$, $p < .001$, $\eta^2_p = 0.086$
EA vs. Control	$t(357) = 5.35$, $p < .001$, *Cohen's d* = 0.635
XA vs. Control	$t(357) = 3.86$, $p < .001$, *Cohen's d* = 0.560
EA vs. XA	$t(357) = 0.49$, $p > .999$, *Cohen's d* = 0.075

Responsibility to Future Generations (RFG, Equitable Prosociality)

Omnibus	$F(2, 357) = 21.60$, $p < .001$, $\eta^2_p = 0.108$
EA vs. Control	$t(357) = 6.51$, $p < .001$, *Cohen's d* = 0.773
XA vs. Control	$t(357) = 3.01$, $p = .009$, *Cohen's d* = 0.436
EA vs. XA	$t(357) = 2.18$, $p = .089$, *Cohen's d* = 0.337
Expansive Altruism (EAIS, Equitable Prosociality)	
Omnibus	$F(2, 357) = 31.80$, $p < .001$, $\eta^2_p = .151$
EA vs. Control	$t(357) = 7.92$, $p < .001$, *Cohen's d* = 0.940
XA vs. Control	$t(357) = 3.51$, $p = .002$, *Cohen's d* = 0.509
EA vs. XA	$t(357) = 2.79$, $p = .017$, *Cohen's d* = 0.431
Effectiveness Focus (EAIS, Effective Prosociality)	
Omnibus	$F(2, 357) = 59.30$, $p < .001$, $\eta^2_p = .249$
EA vs. Control	$t(357) = 8.92$, $p < .001$, *Cohen's d* = 1.058
XA vs. Control	$t(357) = -2.96$, $p = .010$, *Cohen's d* = -0.429
EA vs. XA	$t(357) = 9.64$, $p < .001$, *Cohen's d* = 1.488
Behavioral Donations to Effective Causes (BDT, Effective Prosociality)	
Omnibus	$F(2, 357) = 19.8$, $p < .001$, $\eta^2_p = .100$
EA vs. Control	$t(357) = 6.14$, $p < .001$, *Cohen's d* = 0.729
XA vs. Control	$t(357) = 3.32$, $p = .003$, *Cohen's d* = 0.482
EA vs. XA	$t(357) = 1.60$, $p = .329$, *Cohen's d* = 0.247

Note. [1]These models were estimated without controlling for variation in income. To rule out the possibility that the observed Sample effects on real-world charitable action owe to differences in income between the three samples, exploratory models controlling for income can be found in the analysis script posted on OSF. In short, these exploratory models revealed the same patterns as those presented in the main text.

Figure 2. *Reported real-world charitable action in effective altruists, extraordinary altruists and general population controls*



Note. *Raincloud plots displaying the levels of financial* (***a***) *and temporal* (***b***) *resources devoted to the benefit of others reported by EAs, XAs and demographically-similar controls. Plots display individual data-points, jittered for readability, with overlaid split-violins to illustrate the shape of the underlying probability distributions. Means and error bars depicting 95% CIs are also included, as well as box plots with notches to convey 95% CIs around the medians.*

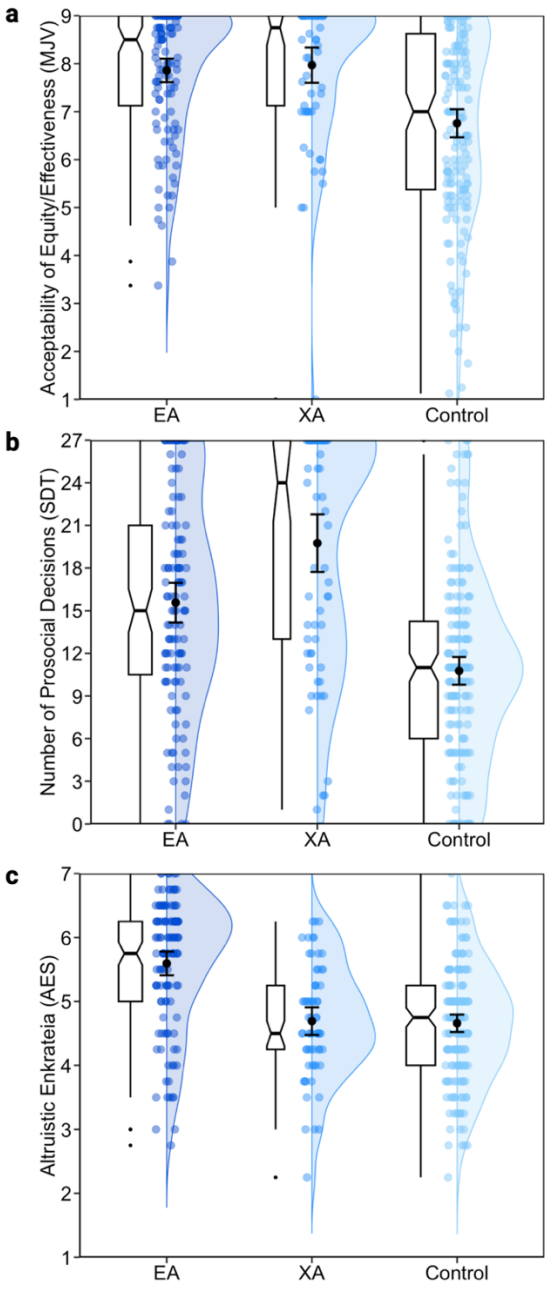
The data presented above shed light on how the three samples differ on real-world prosocial behavior overall. Nonetheless, the particular focus of the present investigation is to shed light on the architecture of prosocial attitudes and actions that are *equitable*, in that they are not constrained by parochial biases that routinely limit everyday prosocial displays towards close, identifiable or otherwise similar beneficiaries (Bloom, 2016; Caviola et al., 2021; Cikara et al., 2011; Fowler et al., 2021; Law et al., 2022; Law, Amormino, et al., 2024; Schubert & Caviola, 2024), and *effective*, in that they have a high potential for maximizing welfare gains. To mimic the types of tradeoffs between social closeness and gains in welfare that are common in the real world, where resources–from the perspective of the affluent–can do greater good when donated to distant strangers, a subset of the prosociality metrics captured attitudes towards altruism that is simultaneously equitable and effective.

On each of these metrics, there was a significant omnibus effect of Sample, with 10% to 18% of the variance in these outcomes being attributable to differences between the three subject groups (see Table 4). On the Moral Judgment Vignettes (MJV) task, both EAs and XAs reported more positive third-party moral judgments of welfare-maximizing (i.e., effectiveness) altruism directed towards distant strangers that came at the expense of not being able to help a less-effective but socially closer alternative (i.e., equitability). Subjects from the two special populations did not differ from one another on average in their judgments of these altruistic decisions. Similarly, both special populations made a greater number of decisions in the Social Discounting Task (SDT) to allocate larger monetary rewards (i.e., effectiveness) to other individuals across a range of social distances (i.e., equitability) rather than keeping smaller rewards for themselves. Especially intriguing is that XAs on this task demonstrated greater prosociality than EAs. Conversely, on the composite metric capturing the endorsement of Altruistic Enkrateia, a measure that, broadly construed, captures attitudes in favor of setting aside personal commitments to prioritize equitability and effectiveness in altruism, EAs scored higher than both XAs and general population controls, while XAs did not differ from the general population. See Figure 3 for a graphical depiction of these results.

Overall, these findings provide support that the moral judgments and prosocial decisions of both special populations favor equitability and effectiveness in altruism in the context of real-world tradeoffs–where supporting the most effective causes also happens to require a departure from the parochial biases which generally constrain prosociality towards those who are close. Yet, only EAs explicitly endorse exerting strength of will to overcome personal commitments to engage in these actions to a greater extent than controls. Because XAs and EAs alike embody effectiveness in their moral judgments (on the MJV task) and decision-making (on the SDT), a possible interpretation of these findings is that XAs may simply feel that helping others in an

effective and equitable manner is in line with their personal commitments and thus doesn't require setting them aside. Indeed, prior research has shown that XAs are characterized by robust patterns of empathic responding towards the suffering of complete strangers (Law, Amormino, et al., 2024) and that feeling greater empathy towards targets is associated with more pronounced feelings of personal commitment to improve their welfare (Batson & Ahmad, 2009; Bruneau et al., 2017; Lamm et al., 2007; Montealegre et al., 2020). Moreover, these findings provide the first evidence that XAs, whose prosociality is driven by empathy rather than reasoning, tend to prioritize both altruistic equity *and* effectiveness in their helping behaviors. This contrasts with prevailing perspectives that view empathy as a constrictive force in efforts to promote these virtues (Bloom, 2016; Caviola et al., 2021). Thus, I find initial support that empathy, or at least people who show robust enrichments in empathic capacities, may be a force for equitable *and* effective good.

Figure 3. *Differences between effective altruists, extraordinary altruists and general population controls on measures capturing equitable and effective altruism simultaneously*



Note. *Raincloud plots displaying moral judgments in the MJV task* (***a***)*, the number of prosocial decisions made in the SDP* (***b***)*, and endorsements of altruistic Enkrateia on the AES* (***c***) *between EAs, XAs and demographically-similar controls. Plots display individual data-points, jittered for readability, with overlaid split-violins to illustrate the shape of the underlying probability*

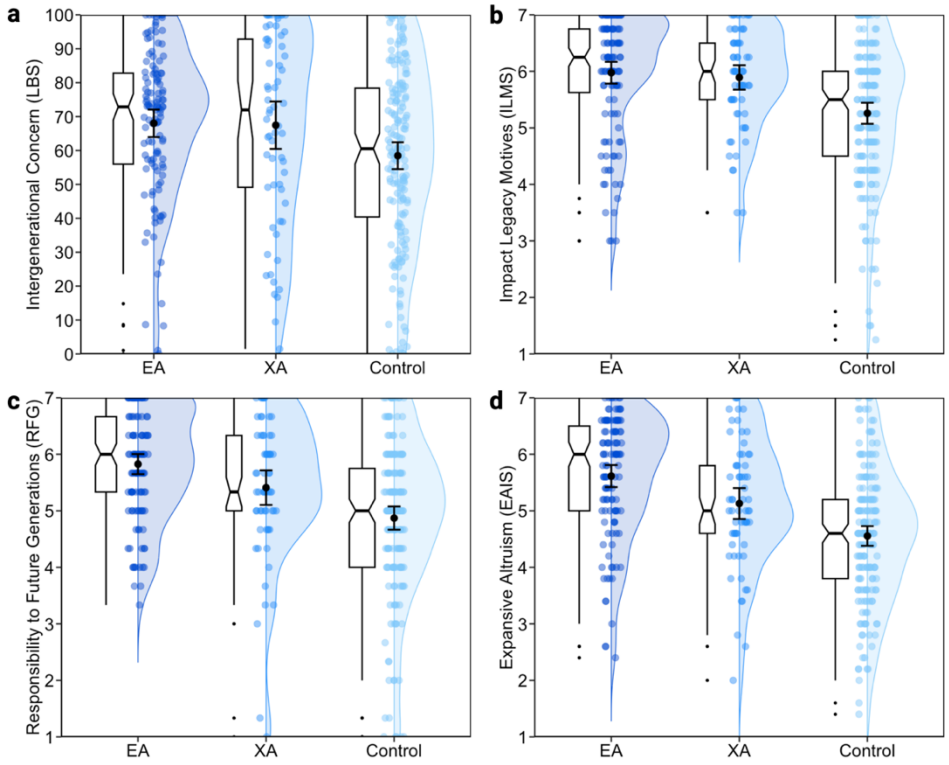
distributions. Means and error bars depicting 95% CIs are also included, as well as box plots with notches to convey 95% CIs around the medians.

The present investigation also included measures of equitable and effective prosociality that captured these two facets in isolation, to disentangle whether and how the three samples differ on each facet when considered separately. On the expansive altruism subscale of the EAIS, which captures impartiality regarding the relational proximity of prosocial beneficiaries–a measure that uniquely captures altruistic equity, both special populations scored higher on average than general population controls and EAs scored higher on average than XAs. Refer to Table 4 above for inferential statistics. See Figure 4 below for a graphical depiction of results related to sample differences on the metrics uniquely capturing altruistic equity.

I also included three metrics capturing similar attitudes of altruistic equity across time–impartiality to the *temporal* distance of altruistic beneficiaries. I chose to include these metrics on account of growing concern across philosophy (see the ethical philosophy of Longtermism; Greaves & MacAskill, 2019; MacAskill, 2022; Ord, 2021), international politics (Bose & Shepardson, 2023; Nations, 2022) and the natural sciences (Aitken, 1998; Barnosky et al., 2011) over rising levels of existential threat from sources such as pandemics, climate change, and AI–sources that are expected to have their most devastating impacts on distant future generations. Specifically, combating these future challenges requires collective action from present society. Much of my research shows that people exhibit parochial bias toward temporally distant beneficiaries that closely resemble the biases they show across social distance in the present (Law, Syropoulos, & Young, 2023; Syropoulos, Law, Kraft-Todd, et al., 2024b; Syropoulos, Law, & Young, 2024e, 2024d). With regard to these metrics capturing altruistic equity across time, EAs and XAs scored higher on average than controls, but did not differ meaningfully from one another, on intergenerational concern, as measured by the LBS, as well as legacy related

concerns to leave a positive impact on society, as measured by the ILMS. Both also scored higher on reported feelings of personal responsibility to help and protect future generations, with EAs scoring higher than XAs on this measure in particular. In sum, with regard to the measures capturing altruistic equity alone, both special populations score remarkably higher than ordinary adults on each, with EAs on some, namely expansive altruism and responsibility to future generations, scoring higher than XAs.

Figure 4. *Differences between effective altruists, extraordinary altruists and general population controls on measures capturing equitable altruism in isolation*



Note. *Raincloud plots displaying intergenerational concern on the LBS* (***a***)*, legacy motivations on the ILMS* (***b***)*, perceived personal duty to help and protect future generations on the RFG* (***c***)*, and expansive altruism on the EAIS* (***d***) *between EAs, XAs and demographically-similar controls. Plots display individual data-points, jittered for readability, with overlaid split-violins to illustrate the shape of the underlying probability distributions. Means and error bars depicting*

95% CIs are also included, as well as box plots with notches to convey 95% CIs around the medians.

Venturing further into how prosociality across Samples 1, 2 and 3 differ as a function of beneficiary distance, I conducted two mixed-ANOVAs specifying Sample as the between-subjects factor with denied beneficiary social distance (on the MJV task) and beneficiary temporal distance (on the LBS) as the within-subjects factor. On the MJV, participants rated the moral acceptability of welfare-maximizing altruism directed by a hypothetical donor towards a distal beneficiary. What differed within-subjects, across trials, was the social closeness to the hypothetical donor of an alternative beneficiary who was denied aid as a consequence of helping the beneficiary that was simultaneously more-effective and more distant (i.e., the denied recipient was either a compatriot, townsman, friend, or family member of the donor). My prior research has shown that general population subjects show markedly less positive moral judgments of welfare-maximizing altruism when denied beneficiaries are depicted as being socially closer compared to more distant from hypothetical donors (Law et al., 2022).

In the present research, despite significant main effects of Sample (EAs and XAs reported more positive moral judgments than controls) and denied beneficiary social distance (moral judgments were more positive for more distal denied beneficiaries), the main effects were qualified by a significant interaction. Examining the simple main effect of social distance within each of the three samples revealed that membership to the two special populations substantially attenuated the effect of distance on moral judgments. In other words, both EAs and XAs found effective prosociality more effective even when it required more equitable or impartial care for the suffering of others (i.e., regardless of how socially close the denied beneficiary was to the hypothetical donor). Table 5 displays relevant inferential statistics and Figure 5 provides a graphical depiction of these results.

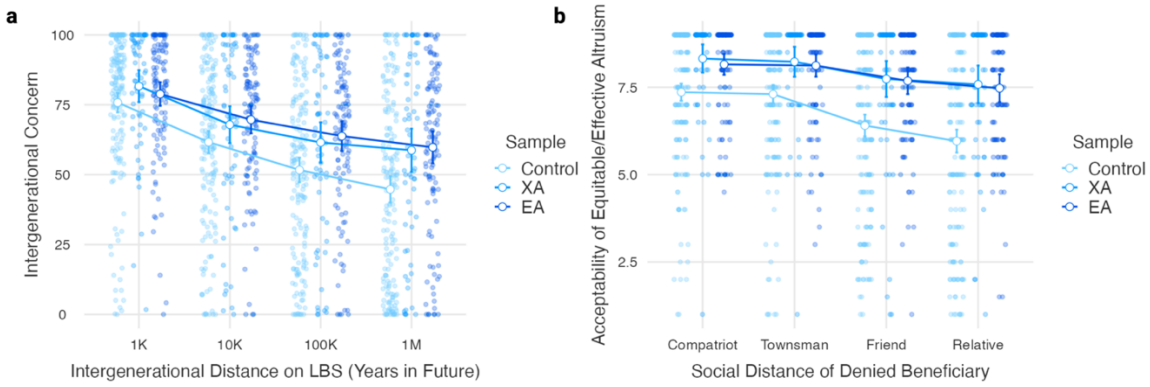
Table 5

Results from mixed ANOVA with post-hoc tests employing the Bonferroni alpha correction evaluating the main effects of Sample and denied beneficiary distance as well as their interaction on moral judgments of welfare-maximizing altruism directed towards distant beneficiaries.

Social Distance Comparison	Moral Acceptability of Equitable and Effective Altruism
Moral Judgment Vignettes (MJV, Equitable & Effective Prosociality)	
Sample	$F(2, 357) = 20.20$, $p < .001$, $\eta^2_p = 0.102$
Social Distance of Denied Beneficiary	$F(3, 1071) = 55.23$, $p < .001$, $\eta^2_p = 0.134$
Sample X Social Distance	$F(6, 1071) = 4.60$, $p < .001$, $\eta^2_p = 0.025$
Simple Effect of Denied Beneficiary Social Distance on Moral Acceptability in EAs	
Compatriot vs. Townsman	$t(357) = 0.45$, $p > .999$
Compatriot vs. Friend	$t(357) = 3.24$, $p = .086$
Compatriot vs. Relative	$t(357) = 4.18$, $p = .002$
Townsman vs. Friend	$t(357) = 2.97$, $p = .208$
Townsman vs. Relative	$t(357) = 3.93$, $p = .007$
Friend vs. Relative	$t(357) = 1.63$, $p > .999$
Simple Effect of Denied Beneficiary Social Distance on Moral Acceptability in XAs	
Compatriot vs. Townsman	$t(357) = 0.91$, $p > .999$
Compatriot vs. Friend	$t(357) = 2.98$, $p = .206$
Compatriot vs. Relative	$t(357) = 3.35$, $p = .059$
Townsman vs. Friend	$t(357) = 2.48$, $p = .910$
Townsman vs. Relative	$t(357) = 2.90$, $p = .263$
Friend vs. Relative	$t(357) = 0.88$, $p > .999$
Simple Effect of Denied Recipient Social Distance on Moral Acceptability in Controls	
Compatriot vs. Townsman	$t(357) = 0.92$, $p > .999$
Compatriot vs. Friend	$t(357) = 8.00$, $p < .001$
Compatriot vs. Relative	$t(357) = 10.45$, $p < .001$
Townsman vs. Friend	$t(357) = 7.43$, $p < .001$
Townsman vs. Relative	$t(357) = 9.92$, $p < .001$
Friend vs. Relative	$t(357) = 4.21$, $p = .002$

Note. *The results presenting the simple effect of distance at each level of Sample are depicted here and the full decomposition can be found in the analysis script on the OSF.*

Figure 5. *The interaction of sample and beneficiary distance on intergenerational concern and moral judgments of altruism*



Note. *Line graphs displaying the interaction between sample and beneficiary distance on intergenerational concern on the LBS (**a**)and moral judgments on the MJV task (**b**). On the LBS, participants indicated their levels of beneficent concern for future generations depicted at four timeframes of varying distance from the present. Moral judgments on the MJV were reported with respect to the decisions of hypothetical donors who opted to aid an effective but socially distant cause instead of a less-effective but socially closer alternative. The closeness to the donor of the denied beneficiary was manipulated within-subjects. Plots display individual data-points, jittered for readability. Means and error bars depicting 95% CIs are also included.*

On the LBS, participants rated their levels of beneficent concern for future generations depicted at four timeframes progressively more distal from the present (i.e., 1,000, 10,000, 100,000, and 1,000,000 years from the present). My prior research has shown that general population subjects show substantially less intergenerational concern for future generations that are farther away in time . Mirroring the findings from the MJV task above, significant main effects of Sample (EAs and XAs reported greater intergenerational concern) and beneficiary distance (intergenerational concern was higher for proximal compared to more distal beneficiaries) were qualified by a significant interaction. Here, I examined the simple main

effects of Sample at each level of beneficiary temporal distance. At the most distal timeframe, EAs reported significantly greater intergenerational concern than controls, suggesting partial attenuation. Table 6 displays relevant inferential statistics and Figure 5 (above) provides a graphical depiction of these results.

Table 6

Results from mixed ANOVA with post-hoc tests employing the Bonferroni alpha correction evaluating the main effects of Sample and beneficiary distance as well as their interaction on intergenerational concern on the LBS. The results presenting the simple effect of Sample at each level of distance are depicted here and the full decomposition can be found in the analysis script on the OSF. Moreover, an additional exploratory analysis can be found on the OSF where Bonferroni corrections were not applied, given the large number of comparisons required to decompose the interaction. The exploratory analyses, which show a more pronounced effect of attenuation, should be interpreted with caution. Further replications of these findings in larger samples are warranted.

Intergenerational Distance Comparison	Intergenerational Concern
Intergenerational Concern (LBS, Equitable Prosociality)	
Sample	$F(2, 357) = 6.02$, $p = .003$, $\eta^2_p = 0.033$
Intergenerational Distance of Beneficiary	$F(3, 1071) = 198.05$, $p < .001$, $\eta^2_p = 0.357$
Sample X Intergenerational Distance	$F(6, 1071) = 6.16$, $p < .001$, $\eta^2_p = 0.033$
Simple Effect of Sample on Intergenerational Concern for Targets 1K Years from the Present	
EA vs. Control	$t(357) = 1.10$, $p > .999$
XA vs. Control	$t(357) = 1.74$, $p > .999$
EA vs. XA	$t(357) = -0.79$, $p > .999$
Simple Effect of Sample on Intergenerational Concern for Targets 10K Years from the Present	
EA vs. Control	$t(357) = 2.51$, $p = .837$
XA vs. Control	$t(357) = 1.60$, $p > .999$
EA vs. XA	$t(357) = 0.43$, $p > .999$
Simple Effect of Sample on Intergenerational Concern for Targets 100K Years from the Present	

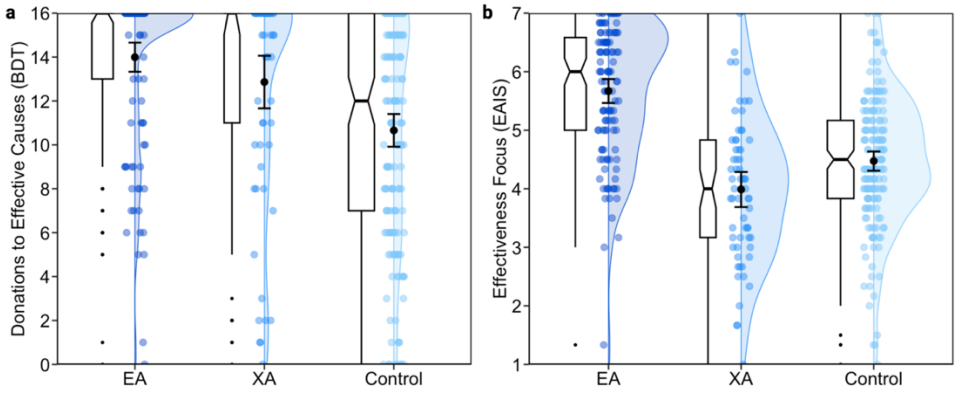
EA vs. Control	$t(357) = 3.45,\ p = .041$
XA vs. Control	$t(357) = 2.29,\ p > .999$
EA vs. XA	$t(357) = 0.50,\ p > .999$
Simple Effect of Sample on Intergenerational Concern for Targets 1M Years from the Present	
EA vs. Control	$t(357) = 4.01,\ p = .005$
XA vs. Control	$t(357) = 3.05,\ p = .164$
EA vs. XA	$t(357) = 0.22,\ p > .999$

Important to note is that EAs and XAs descriptively showed the same pattern of greater concern relative to controls at each level of beneficiary distance, despite these comparisons not reaching statistical significance when applying the Bonferroni alpha correction. Notably, the Bonferroni correction drastically reduces statistical power in the present analysis, given the large number of comparisons required to decompose the interaction. Exploratory analyses not applying the correction can be found in the analysis script on the OSF, which show that the intergenerational concern of both EAs and XAs significantly differed from that of controls at three out of the four timeframes when considered at the .05 significance level. Nonetheless, these exploratory findings warrant interpretative caution and further replication efforts with larger sample sizes are required to more definitively elucidate these patterns.

Finally, I also included two metrics that capture altruistic effectiveness in isolation. Here, I reasoned that EAs and XAs alike would score higher than general population controls. Nonetheless, I also reasoned it was possible that EAs, who place a particular emphasis on utilitarian impact and effectiveness in altruism, may score higher than XAs, whose real-world prosociality is more definitively aligned with equitability. On the effectiveness focus subscale of the EAIS, an intriguing pattern was observed. Specifically, EAs scored higher than both general population controls *and* XAs, with XAs scoring even lower than members of the general population. Yet, on a behavioral metric of effectiveness-prioritization, where subjects allocated actual resources between effective versus ineffective causes across 16 trials, both EAs and XAs

chose the effective charitable option on a greater number of trials relative to the control sample (see Table 4 and Figure 6). These findings dovetail with the findings above, where XAs displayed more effective decision-making on the SDT and moral judgments on the MJV task, but less effective explicit attitudes on the AES. Taken together, they suggest that extraordinary altruists may indeed embody a leaning towards effective causes in their altruistic *behaviors*, *decisions*, and *judgments*, but a leaning that does not always manifest in their *explicit attitudes*.

Figure 6. *Differences between effective altruists, extraordinary altruists and general population controls on measures capturing effective altruism in isolation*



Note. *Raincloud plots displaying the number of donations made to effective causes vs. ineffective causes on the BDT (**a**) and explicit attitudes in alignment with endorsing effectiveness prioritization in altruistic decisions on the EAIS (**b**). Plots display individual data-points, jittered for readability, with overlaid split-violins to illustrate the shape of the underlying probability distributions. Means and error bars depicting 95% CIs are also included, as well as box plots with notches to convey 95% CIs around the medians.*

In summary, the findings above largely supported the predictions under Question 1a (see Table 1). EAs and XAs scored higher on measures of equitable and effective prosociality (and prosociality more broadly) than members of the general population on 10 out of the 11 measures that were employed. Nonetheless, while EAs outscored XAs on effectiveness focus, this pattern

was not observed on the BDT, a behavioral metric of effective prosociality. Thus, from the present findings, it is clear that both effective and extraordinary enaltruists' prosociality is not only evident in their behaviors in vivo, but also in their attitudes and behaviors in the lab. Moreover, although XAs' modality of real-world altruism (i.e., organ donorship to strangers in need) aligns more closely with the principle of altruistic equity in overcoming parochial bias, this population also seems to favor causes that are more impactful, at least in the context of their behaviors, decisions and judgments. Given that the altruism of XAs has been shown to be driven primarily by emotional psychological processes like empathy (Amormino, O'Connell, et al., 2022; Law, Amormino, et al., 2024; Marsh, 2019), these findings provide preliminary and novel evidence that theories of empathy as a constrictive force in the context of equitable and effective altruism may warrant revisiting.

3.1.2. Population Differences in Empathic and Reasoning Ability. After finding substantial evidence that both EAs and XAs display numerous signatures of equitable and effective prosociality in the lab, I next explored differences between Samples 1, 2, and 3 on a host of psychological features that have been shown to predict prosocial attitudes and behaviors in the general population. To start, I compared the three samples on measures of empathic and reasoning ability, features that have been central to fervent debates in the context of research and discourse on prosociality across psychology and philosophy.

In brief, people routinely help those who are relationally close or similar to themselves (Berg et al., 1995; Hamilton, 1964; Trivers, 1971). In these contexts, empathy is a robust predictor of prosocial behavior (Batson et al., 1983, 1997; Batson & Ahmad, 2009). Yet, people also show parochial biases in prosociality, favoring closer beneficiaries *over* distant ones (Bruneau et al., 2017; Dovidio et al., 2010; Patil et al., 2021; Sherif et al., 1961), even in contexts where helping distant others can do comparatively greater good (e.g., Law et al., 2022). Indeed,

as evidenced by the research cited above, the goals of much psychological inquiry over the past half century, particularly in the subfield of social psychology, has been devoted to overcoming these parochial biases to promote equity in altruism. Intriguingly, however, a number of researchers and philosophers, including those associated with the effective altruism philosophy, have advocated overcoming parochial biases by downregulating empathy–which at times is subject to parochial bias–to instead engage a more rational and deliberative process when deciding who deserves help (Bloom, 2016; J. Greene, 2014; Prinz, 2011; Schubert & Caviola, 2024).

Despite prevailing skepticism regarding empathy's potential to promote altruism at a distance, considerable emerging evidence suggests that empathy may still indeed be a force for equitable and effective good (Fowler et al., 2021; Zaki, 2018). Of particular relevance to the present investigation, empathy has been shown to drive prosocial attitudes and behaviors even towards distant strangers among altruistic organ donors (i.e., extraordinary altruists or XAs; Law, Amormino, et al., 2024). Moreover, this exceptionally altruistic population has been shown to possess profound enhancements in empathic ability compared to general population controls. While proponents of effective altruism advocate reasoning over empathy as a better guide for equitable and effective altruism, the research on XAs raises the question of whether effective altruists (EAs)–who have been empirically understudied to date–also show similar enhancements in empathic ability, and whether this empathy drives their prosocial engagement. It is also possible that reasoning ability is heightened among EAs exclusively, or among EAs and XAs collectively, and represents a distinct or complementary route, alongside empathy, in fostering altruistic equity and effectiveness.

To begin addressing the questions above, I began by first comparing whether and how EAs, XAs and general population controls differ on metrics of empathic and reasoning ability.

Then, in subsequent sections, I proceeded to investigate how these processes and their interactions influence equitable and effective prosociality across the three subject groups. On the basis of the extant research and discourse taken collectively, I hypothesized that EAs and XAs, relative to general population controls, would score higher on empathic ability, reasoning ability, or perhaps both.

With regard to empathic ability, members of the three populations differed on several measures: empathic concern (EC) on the IRI, capacities in correctly identifying emotional content in written statements on the EEST, alexithymia on the TAS, and primary and secondary psychopathy on the LSRP. For the latter three measures, higher scores (reverse-coded) correspond to greater emotional ability. The three populations did not differ significantly on beliefs regarding the within-person malleability of empathy on the TES, nor on feelings of outgroup empathy on the PES (though results trended toward significance on the PES, with XAs scoring marginally higher than controls[3]). Confirming the findings from research on altruistic organ donors (see Marsh, 2019 for review), XAs, relative to controls, scored higher on EC on the IRI, reverse-coded alexithymia on the TAS, and reverse coded primary and secondary psychopathy on the LSRP. Moreover, on each of these metrics, as well as the EEST, XAs scored higher than EAs, though there was no significant difference in performance on the EEST between XAs and controls. Perhaps most in intriguing, however, is that EAs, relative to controls, showed deficits in emotional recognition on the EEST, an effect consistent with prior findings in the special populations of individuals high in psychopathy and callus unemotional traits (Marsh & Cardinale, 2012; O'Connell et al., 2019). Also of note is that EAs differed from controls on no other metric, and in fact trended lower on most (see Table 7 and Figure 7).

[3] Despite not reaching significance, this effect had a *Cohen's d* of 0.315, which is greater than negligible. Nonetheless, further research with a larger sample size is required to elucidate whether this effect is robust or more likely to be an artifact. As such, these findings are inconclusive and should be interpreted with caution.

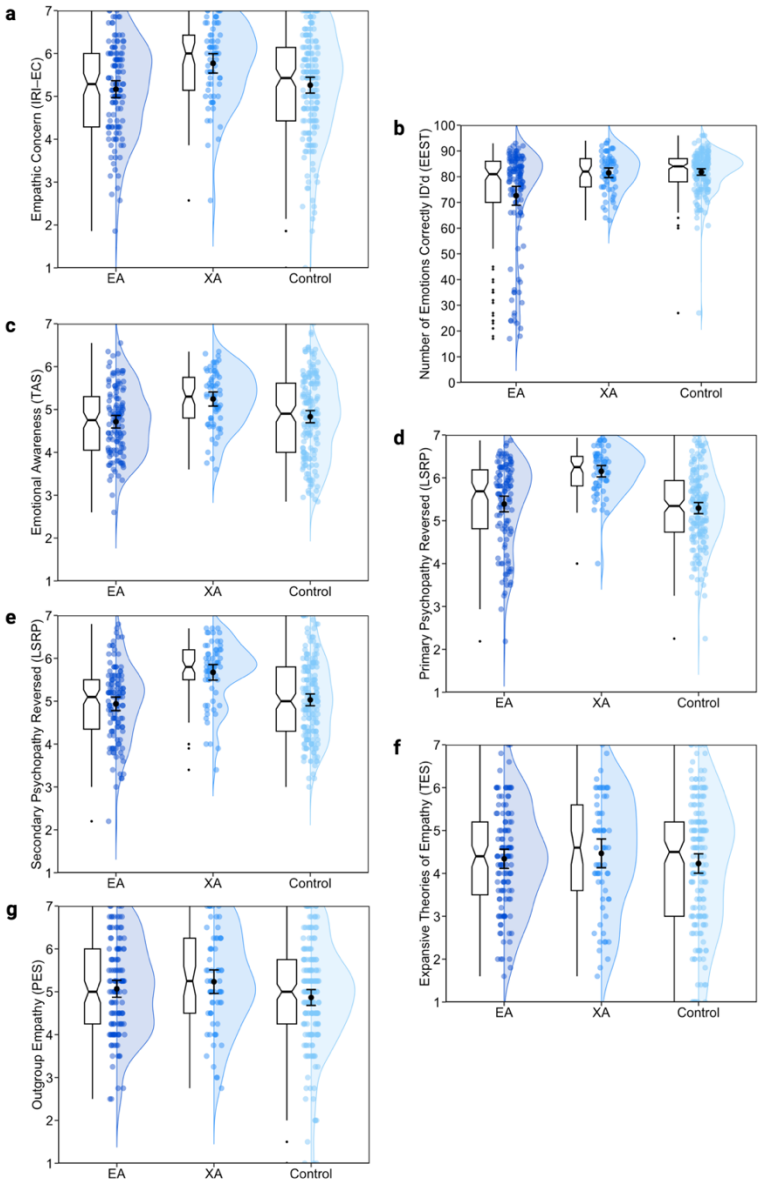
Table 7

Results from one-way ANOVAs with post-hoc tests employing the Bonferroni alpha correction comparing Samples 1-3 on measures capturing empathic and reasoning ability.

Sample	Empathic Ability
Empathic Concern (IRI–EC)	
Omnibus	$F(2, 357) = 6.34$, $p = .002$, $\eta^2p = .034$
EA vs. Control	$t(357) = -0.69$, $p > .999$, *Cohen's d* = -0.082
XA vs. Control	$t(357) = 3.07$, $p = .007$, *Cohen's d* = 0.445
EA vs. XA	$t(357) = -3.42$, $p = .002$, *Cohen's d* = -0.527
Ability to Identify Emotions in Statements (EEST)	
Omnibus	$F(2, 357) = 18.70$, $p < .001$, $\eta^2p = .095$
EA vs. Control	$t(357) = -5.82$, $p < .001$, *Cohen's d* = -0.691
XA vs. Control	$t(357) = -0.17$, $p > .999$, *Cohen's d* = -0.025
EA vs. XA	$t(357) = -4.32$, $p < .001$, *Cohen's d* = -0.666
Emotional Awareness (TAS)	
Omnibus	$F(2, 357) = 8.22$, $p < .001$, $\eta^2p = .044$
EA vs. Control	$t(357) = -1.14$, $p = .762$, *Cohen's d* = -0.136
XA vs. Control	$t(357) = 3.30$, $p = .003$, *Cohen's d* = 0.479
EA vs. XA	$t(357) = -3.98$, $p < .001$, *Cohen's d* = -0.614
Reverse-Coded Primary Psychopathy (LSRP Primary Psychopathy)	
Omnibus	$F(2, 357) = 24.00$, $p < .001$, $\eta^2p = .118$
EA vs. Control	$t(357) = 0.92$, $p > .999$, *Cohen's d* = 0.109
XA vs. Control	$t(357) = 6.78$, $p < .001$, *Cohen's d* = 0.984
EA vs. XA	$t(357) = -5.67$, $p < .001$, *Cohen's d* = -0.875
Reverse-Coded Secondary Psychopathy (LSRP Secondary Psychopathy)	
Omnibus	$F(2, 357) = 16.40$, $p < .001$, $\eta^2p = .084$
EA vs. Control	$t(357) = -0.90$, $p > .999$, *Cohen's d* = -0.107
XA vs. Control	$t(357) = 5.03$, $p < .001$, *Cohen's d* = 0.731
EA vs. XA	$t(357) = -5.43$, $p < .001$, *Cohen's d* = -0.838
Beliefs that Empathy is Malleable (TES)	
Omnibus	$F(2, 357) = 0.72$, $p = .486$, $\eta^2p = .004$
EA vs. Control	$t(357) = 0.66$, $p > .999$, *Cohen's d* = 0.078
XA vs. Control	$t(357) = 1.17$, $p = .731$, *Cohen's d* = 0.170
EA vs. XA	$t(357) = -0.59$, $p > .999$, *Cohen's d* = -0.092
Outgroup Empathy (PES)	
Omnibus	$F(2, 357) = 2.67$, $p = .071$, $\eta^2p = .015$
EA vs. Control	$t(357) = 1.46$, $p = .433$, *Cohen's d* = 0.174
XA vs. Control	$t(357) = 2.17$, $p = .092$, *Cohen's d* = 0.315

EA vs. XA	$t(357) = -0.92$, $p > .999$, *Cohen's d* = -0.141
Sample	**Reasoning Ability**
Need for Cognition (NFC)	
Omnibus	$F(2, 357) = 10.20$, $p < .001$, $\eta^2p = .054$
EA vs. Control	$t(357) = 4.51$, $p < .001$, *Cohen's d* = 0.536
XA vs. Control	$t(357) = 1.49$, $p = .414$, *Cohen's d* = 0.216
EA vs. XA	$t(357) = 2.07$, $p = .116$, *Cohen's d* = 0.320
Ability to Override Intuition on Word Problems (CRT)	
Omnibus	$F(2, 357) = 19.40$, $p < .001$, $\eta^2p = .098$
EA vs. Control	$t(357) = 4.34$, $p < .001$, *Cohen's d* = 0.515
XA vs. Control	$t(357) = -2.78$, $p = .017$, *Cohen's d* = -0.403
EA vs. XA	$t(357) = 5.96$, $p < .001$, *Cohen's d* = 0.919
Tendency to Use Algorithmic vs. Heuristic Thinking on Word Problems (HBT)	
Omnibus	$F(2, 357) = 10.80$, $p < .001$, $\eta^2p = .057$
EA vs. Control	$t(357) = 4.37$, $p < .001$, *Cohen's d* = 0.520
XA vs. Control	$t(357) = 2.91$, $p = .012$, *Cohen's d* = 0.422
EA vs. XA	$t(357) = 0.63$, $p > .999$, *Cohen's d* = 0.098
Self-Reported Reasoning Ability (REI)	
Omnibus	$F(2, 357) = 0.08$, $p = .920$, $\eta^2p < .001$
EA vs. Control	$t(357) = -0.41$, $p > .999$, *Cohen's d* = -0.048
XA vs. Control	$t(357) = -0.16$, $p > .999$, *Cohen's d* = -0.023
EA vs. XA	$t(357) = -0.16$, $p > .999$, *Cohen's d* = -0.025
Actively Open-Minded Thinking (AOT)	
Omnibus	$F(2, 357) = 2.22$, $p = .110$, $\eta^2p = .012$
EA vs. Control	$t(357) = 0.81$, $p > .999$, *Cohen's d* = 0.096
XA vs. Control	$t(357) = 2.11$, $p = .108$, *Cohen's d* = 0.306
EA vs. XA	$t(357) = -1.36$, $p = .522$, *Cohen's d* = -0.210

Figure 7. *Differences between effective altruists, extraordinary altruists and general population controls on measures capturing empathic ability*

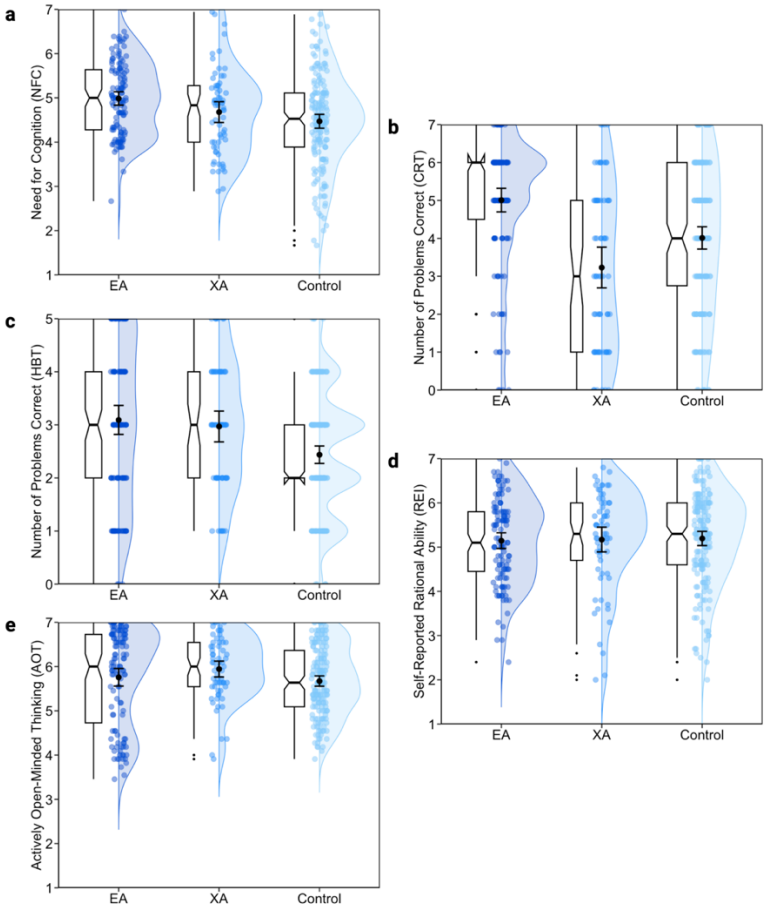


Note. *Raincloud plots displaying empathic concern on the IRI (**a**), the number of correctly identified emotions on the EEST (**b**), alexithymia reverse coded to convey emotional awareness on the TAS (**c**), primary and secondary psychopathy reverse coded to convey emotional ability on the LSRP (**d-e**), beliefs that empathy is malleable on the TES (**f**), and outgroup empathy on the PES (**g**). Plots display individual data-points, jittered for readability, with overlaid split-*

violins to illustrate the shape of the underlying probability distributions. Means and error bars depicting 95% CIs are also included, as well as box plots with notches to convey 95% CIs around the medians.

Regarding reasoning ability, the three populations differed meaningfully on need for cognition (NFC) and tendencies to engage in deliberative processes to arrive at correct answers to challenging word problems on both the cognitive reflection test (CRT) and a battery of heuristics and bias tasks (HBT). No significant differences were observed among the populations in self-reported reasoning ability on the rational experiential index (REI), nor in beliefs about changing one's mind to accommodate evidence that challenges existing viewpoints on the AOT scale. Intriguingly, and largely consistent with perspectives advanced among philosophers and researchers associated with the effective altruism movement, EAs scored higher than controls on NFC, the CRT, and the HBT. Moreover, XAs scored lower than EAs and controls on the CRT, but higher than controls on the HBT. This suggests that while EAs exhibit a strong tendency towards deliberative reasoning across various measures, XAs may engage in greater reasoning relative to controls in specific contexts, highlighting distinct cognitive profiles between these groups (see Table 7 and Figure 8). Since reasoning ability may be partly influenced by level of education or income, exploratory ANCOVAs were conducted to estimate sample differences on each of the three significant reasoning outcomes, controlling for both education and income levels. These analyses confirmed that the same patterns are observed with or without controlling for these demographics. The results from these exploratory ANCOVAs are available on the OSF.

Figure 8. *Differences between effective altruists, extraordinary altruists and general population controls on measures capturing reasoning ability*

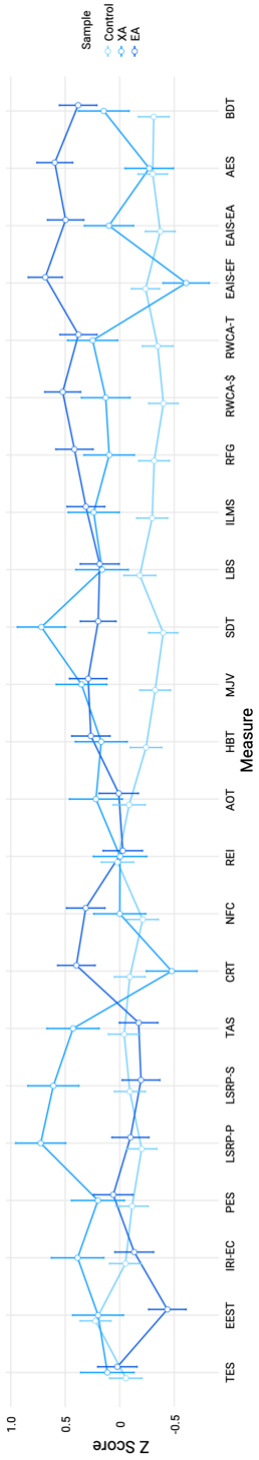


Note. *Raincloud plots displaying need for cognition on the NFC (**a**), the number of correctly solved problems on the CRT (**b**) and HBT (**c**), self-reported rational ability on the REI (**d**), and actively open-minded thinking (AOT; **e**). Plots display individual data-points, jittered for readability, with overlaid split-violins to illustrate the shape of the underlying probability distributions. Means and error bars depicting 95% CIs are also included, as well as box plots with notches to convey 95% CIs around the medians.*

Finally, for exploratory purposes, to test the overall cognitive, affective and prosocial profiles of the three populations when considering variation across all of the outcomes in a unified model, I conducted a profile analysis using MANOVA. First, I standardized scores on

each of the measures capturing prosociality, empathic ability and rational ability by transforming them into z-scores. This step ensured that each was measured on the same scale. Next, I entered Sample as a between-subjects predictor and each of the measures across the three categories above as separate levels of a within-subjects factor. The analysis revealed a significant interaction term between Sample with the cognitive, affective and prosocial profile, $F(44, 7854) = 13.76$, $p < .001$, $\eta^2p = 0.072$, suggesting that scores on these measures, taken collectively, differed across subject groups to a meaningful extent. Graphical analysis revealed patterns identical to those presented above when each outcome was assessed individually. See Figure 9 for a graphical depiction of these results and the analysis script on the OSF for the exhaustive output pertaining to this analysis.

Figure 9. *The cognitive, affective and* (*equitable and effective) prosocial profiles of effective altruists, extraordinary altruists, and general population controls*



Note. *Profile plot comparing Samples 1-3 on measures of empathic ability, rational ability, and equitable/effective prosociality. Error bars correspond to 95% CIs around the mean.*

To summarize, the findings thus far largely align with earlier ones from extant research on the XA population, which shows altruistic organ donors to possess psychological profiles that are characterized by heightened emotional capacities in empathic ability (see Marsh, 2019 for review). Nonetheless, they offer novel insights into the cognitive profiles of this cohort, which have yet to be studied until now. Here, I find mixed results. On one hand, XAs are comparable to the general population on their reported enjoyment engaging in effortful cognitive endeavors (i.e., NFC). They also score more poorly on the cognitive reflection test (CRT), which requires overcoming intuitive responses that initially seem correct to engage deliberative processes to arrive upon the correct solution. Yet, they score higher than the general population on heuristics and bias tasks, which are arguably more difficult to solve than the problems on the CRT, yet the incorrect responses are less intuitive (Frederick, 2005; Marzilli Ericson et al., 2015; Stagnaro et al., 2018; Toplak et al., 2011). Thus, it may be that where the cognitive profiles of EAs and XAs differ the most is not reasoning *ability* but instead the tendency to *enjoy* thinking hard to solve problems. In other words, while both groups may demonstrate unique cognitive strengths, their motivations and preferences for engaging in cognitive tasks appears to differ significantly.

Furthermore, the present findings shed novel light on the cognitive and affective profiles of EAs, which have yet to be studied until now. Namely, these data serve as the first ever evidence that EAs exhibit heightened rational ability relative to controls and diminished empathic ability relative to XAs. Moreover, the empathic ability of EAs appears to dip below levels in the general population at times, particularly with respect to capacities in emotion identification. To conclude, on the basis of the present findings, it seems that there may indeed exist a plurality of avenues to equitable and effective altruism. I explore this possibility more directly by examining the predictive power of empathic and rational ability on prosocial outcomes within in each sample in Phase 2.

3.1.3. Population Differences in Moral Beliefs and Values, Other-Inclusive Identity and Other-Oriented Attitudes.

Beyond investigating the affective and cognitive profiles of each sample, additional measures were included to capture predictors outside the domains of empathy and reasoning that have known relationships with prosociality and equity in the general population. The first set of these additional metrics was included to capture moral beliefs and values[4].

First, I hypothesized moral values that have been shown in prior research to track positively with prosociality and equitability, such the values of harm-reduction (i.e., harm) and fairness on the MFQ (Graham et al., 2011, 2013; Haidt, 2007), fairness on the MAC-Q (Curry et al., 2019), moral concern for the well-being of others on the MES (Crimston et al., 2016, 2018; Syropoulos, Crimston, Markowitz, et al., 2024a), and values in line with the utilitarian principle of impartial beneficence (IB) on the OUS (Everett & Kahane, 2020; Kahane et al., 2018), would be more pronounced in the exceptionally caring subject groups relative to controls. Conversely, I hypothesized that moral values associated with ingroup favoritism, such as loyalty on the MFQ, as well as familial and group loyalty on the MAC-Q, would be less pronounced in the exceptionally caring subject groups relative to controls.

The hypotheses above were partially supported (see Table 8 and Figure 10). With respect to prosocial moral values, as predicted, EAs and XAs both scored higher than controls on moral concern (MES) and IB (OUS), with EAs scoring higher than XAs on IB in particular. These findings suggest that both of the special populations include a greater number of entities within their circles of moral regard and endorse utilitarian values of maximizing welfare through prosocial acts. Yet, against my predictions, EAs and XAs both scored lower than controls on the

[4] Only moral values for which there were motivated hypotheses are presented in the main text. Tendencies to morally value authority and purity on the MFQ as well as reciprocity, deference, heroism, and property rights were also captured. Analyses pertaining to these metrics can be found in the "Analysis Script" posted on the OSF.

measure of fairness from the MFQ, with EAs scoring lower on the measure of fairness from the MAC-Q as well. Moreover, no sample differences were observed on the moral foundation of harm/care from the MFQ.

The findings related to harm are not entirely surprising, as members of the general population tend to uniformly score remarkably high on this value–most people believe it is wrong to harm others and right to reduce suffering (Graham et al., 2011). However, at first blush, the principle of fairness seems to generally align with the equitable behaviors observed in special populations both in real life and–as of now–in laboratory settings. This raises the possibility that these unexpected findings might be explained by the way fairness has been measured. For instance, the Moral Foundations Questionnaire (MFQ) and the Morality as Cooperation Questionnaire (MAC-Q) conflate two dimensions of fairness: (1) equitability–the idea that everyone should receive an equal share of resources, and (2) proportionality–the idea that resources should be distributed based on individual contributions (Atari et al., 2023). Recent measurement tools better distinguish these principles (see emerging findings from Atari and colleagues cited above), and it is possible that both EAs and XAs may score higher than controls on measures that uniquely capture equitability, but not those that uniquely capture proportionality or conflate both dimensions. Further research employing more sensitive metrics should investigate this possibility.

Table 8

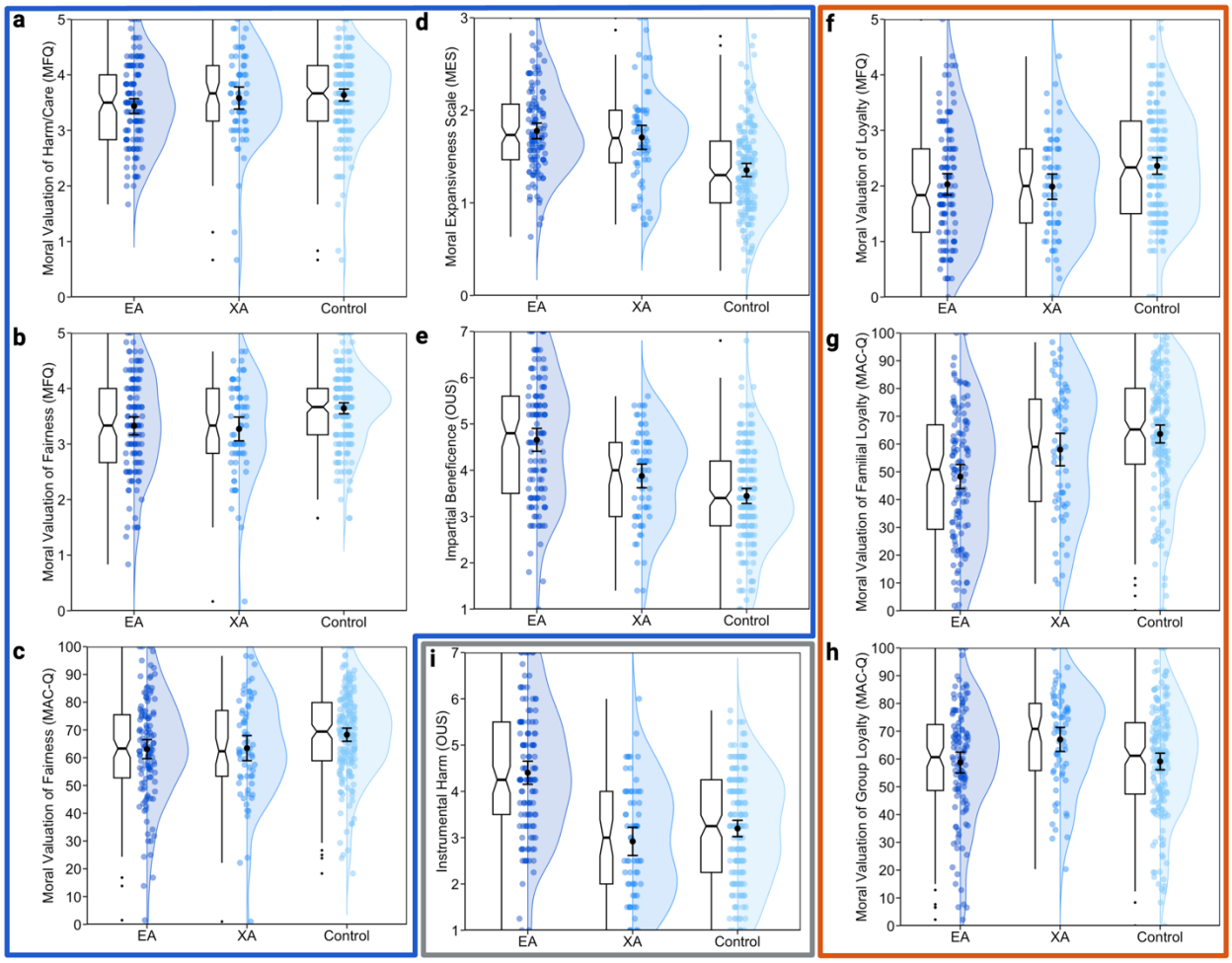
Results from one-way ANOVAs with post-hoc tests employing the Bonferroni alpha correction comparing Samples 1-3 on measures capturing moral beliefs and values.

Sample	Prosocial and Equitable Moral Beliefs and Values
Moral Valuation of Care and Harm Reduction (MFQ–Harm)	
Omnibus	$F(2, 357) = 2.57$, $p = .078$, $\eta^2p = .014$
EA vs. Control	$t(357) = -2.25$, $p = .075$, *Cohen's d* = -0.267
XA vs. Control	$t(357) = -0.50$, $p > .999$, *Cohen's d* = -0.073

EA vs. XA	$t(357) = -1.26$, $p = .624$, *Cohen's d* = -0.195
Moral Valuation of Fairness (MFQ–Fairness)	
Omnibus	$F(2, 357) = 8.52$, $p < .001$, $\eta^2p = .046$
EA vs. Control	$t(357) = -3.44$, $p = .002$, *Cohen's d* = -0.408
XA vs. Control	$t(357) = -3.29$, $p = .003$, *Cohen's d* = -0.478
EA vs. XA	$t(357) = 0.45$, $p > .999$, *Cohen's d* = 0.070
Moral Valuation of Fairness (MAC-Q–Fairness)	
Omnibus	$F(2, 357) = 3.85$, $p = .022$, $\eta^2p = .021$
EA vs. Control	$t(357) = -2.52$, $p = .036$, *Cohen's d* = -0.300
XA vs. Control	$t(357) = -1.92$, $p = .167$, *Cohen's d* = -0.279
EA vs. XA	$t(357) = -0.13$, $p > .999$, *Cohen's d* = -0.021
Moral Concern (MES)	
Omnibus	$F(2, 357) = 30.90$, $p < .001$, $\eta^2p = .147$
EA vs. Control	$t(357) = 7.36$, $p < .001$, *Cohen's d* = 0.874
XA vs. Control	$t(357) = 5.02$, $p < .001$, *Cohen's d* = 0.728
EA vs. XA	$t(357) = 0.94$, $p > .999$, *Cohen's d* = 0.146
Valuation of Utilitarian Impartial Beneficence (OUS–IB)	
Omnibus	$F(2, 357) = 37.90$, $p < .001$, $\eta^2p = .175$
EA vs. Control	$t(357) = 8.70$, $p < .001$, *Cohen's d* = 1.032
XA vs. Control	$t(357) = 2.54$, $p = .035$, *Cohen's d* = 0.368
EA vs. XA	$t(357) = 4.30$, $p < .001$, *Cohen's d* = 0.664
Sample	**Parochial Moral Beliefs and Values**
Moral Valuation of Ingroup Loyalty (MFQ–Loyalty)	
Omnibus	$F(2, 357) = 5.35$, $p = .005$, $\eta^2p = .029$
EA vs. Control	$t(357) = -2.77$, $p = .018$, *Cohen's d* = -0.329
XA vs. Control	$t(357) = -2.56$, $p = .033$, *Cohen's d* = -0.371
EA vs. XA	$t(357) = 0.27$, $p > .999$, *Cohen's d* = 0.042
Moral Valuation of Familial Loyalty (MAC-Q, Familial Loyalty)	
Omnibus	$F(2, 357) = 16.30$, $p < .001$, $\eta^2p = .083$
EA vs. Control	$t(357) = -5.70$, $p < .001$, *Cohen's d* = -0.676
XA vs. Control	$t(357) = -1.70$, $p = .269$, *Cohen's d* = -0.247
EA vs. XA	$t(357) = -2.78$, $p = .017$, *Cohen's d* = -0.429
Moral Valuation of Group Loyalty (MAC-Q, Group Loyalty)	
Omnibus	$F(2, 357) = 4.43$, $p = .013$, $\eta^2p = .024$
EA vs. Control	$t(357) = -0.17$, $p > .999$, *Cohen's d* = -0.020
XA vs. Control	$t(357) = 2.75$, $p = .019$, *Cohen's d* = 0.399
EA vs. XA	$t(357) = -2.72$, $p = .021$, *Cohen's d* = -0.419
Sample	**Instrumental Harm**
Valuation of Utilitarian Instrumental Harm (OUS–IH)	

Omnibus	$F(2, 357) = 42.40$, $p < .001$, $\eta^2p = .192$
EA vs. Control	$t(357) = 8.07$, $p < .001$, *Cohen's d* = 0.958
XA vs. Control	$t(357) = -1.53$, $p = .379$, *Cohen's d* = -0.222
EA vs. XA	$t(357) = 7.65$, $p < .001$, *Cohen's d* = 1.180

Figure 10. *Differences between effective altruists, extraordinary altruists and general population controls on measures moral beliefs and values*



Note. *Raincloud plots displaying levels of endorsement of prosocial and equitable moral values* (***a-e****, within the blue boundary), parochial moral values* (***f-h****, within the red boundary), and the utilitarian principle of instrumental harm* (***i****, within the gray boundary) in Samples 1-3. Plots display individual data-points, jittered for readability, with overlaid split-violins to illustrate the*

shape of the underlying probability distributions. Means and error bars depicting 95% CIs are also included, as well as box plots with notches to convey 95% CIs around the medians.

With respect to moral values that have been shown to promote ingroup favoritism, in line with my predictions, EAs and XAs both scored lower than controls on the moral foundation of loyalty on the MFQ, EAs scored lower than controls and XAs on the value of familial loyalty on the MAC-Q, while XAs scored higher on the value of group loyalty on the MAC-Q compared to both EAs and controls. That is, both special populations place less moral emphasis on showing preferential commitment to those close to them when broadly construed. However, XAs do tend to show greater preferential commitment to their group members than both EAs and ordinary adults alike (see relevant discussion of these findings in Chapter 4).

For exploratory purposes, I was also interested in investigating whether EAs in particular showed greater endorsement of the utilitarian principle of Instrumental Harm (IH on the OUS)–sacrificing the few for the benefit of the many–relative to XAs and controls. Specifically, prominent voices within EA have asserted that the philosophy's alignment with utilitarian ethics is primarily reserved to the sub features of utilitarianism that promote IB–maximizing welfare in the context of prosociality (MacAskill, 2018; Singer, 2015). Moreover, foundational texts associated with the EA philosophy contend EA does *not* advocate for IH. However, the emerging philosophy and growing social movement has recently come into the public spotlight, not for its efforts in alleviating global suffering, but due to significant acts of financial misconduct by one of its most prominent advocates (Cohen & Godoy, 2024). Notably, this individual's criminal defense argued that these financial misdeeds were justified on the basis of the defendant committing them for the purpose of instrumental harm–to raise funds to allocate for the greater good of society. While the EA community maintains that IH is not a foundational objective of the movement, no published research has empirically investigated whether or not this value is

widely supported among members of the EA community. The current design afforded the ability to investigate this possibility. Intriguingly, regardless of the extent to which the principle of instrumental harm is embedded within the foundations of EA philosophy, EAs not only endorse this value, but score 0.96 and 1.18 standard deviations higher on average than controls and XAs, respectively.

Finally, in addition to moral beliefs and values, I included a host of metrics capturing other-inclusive identity[5] and positive other-oriented attitudes. In line with my predictions, EAs and XAs scored higher than controls on both the Identification with All Humanity Scale (IWAH), the Compassionate Love for Humanity Scale (LHS), and the SOFI, which measured reported positive feelings towards others over the course of the week preceding the study. Moreover, consistent with the EA movement's close alignment with the animal welfare movement (see Singer, 1981), EAs endorsed more positive attitudes towards animals, attributing greater capacities of mind to animals on the Individual Differences in Anthropomorphism Scale (IDAQ) and indicating greater inclusion of non-human entities within their self-concepts on the Allo-Inclusive Identity Scale (AIIS). While XAs did not show differences relative to the general population on attitudes pertaining to non-human targets, they did show greater tendencies to humanize stigmatized groups (e.g., terrorists; the Blatant Dehumanization Scale [BDS], reverse-coded). Intriguingly, neither special population differed from the general population on the measure capturing interdependence on others (SCS). In sum, these findings suggest that both EAs and XAs show expansive positive attitudes towards others across spans of relational distance, from socially close humans, to outgroup members, to all humans everywhere, with EAs

[5] Other-inclusive identity pertains to the extent that one's self concept is enmeshed with the social context in which they are embedded. Participants were measured on the following constructs under this umbrella: Identification with All Humanity (IWAH), Compassionate Love for Humanity (LHS), Interdependence on Others (i.e., the Self Construal Scale, SCS), and Allo-Inclusive Identity–the tendency to include humans, non-humans, and all living beings within one's self concept (AIIS).

in particular extending their sense of identification to include even entities within the natural environment (see Table 9 and Figure 11).

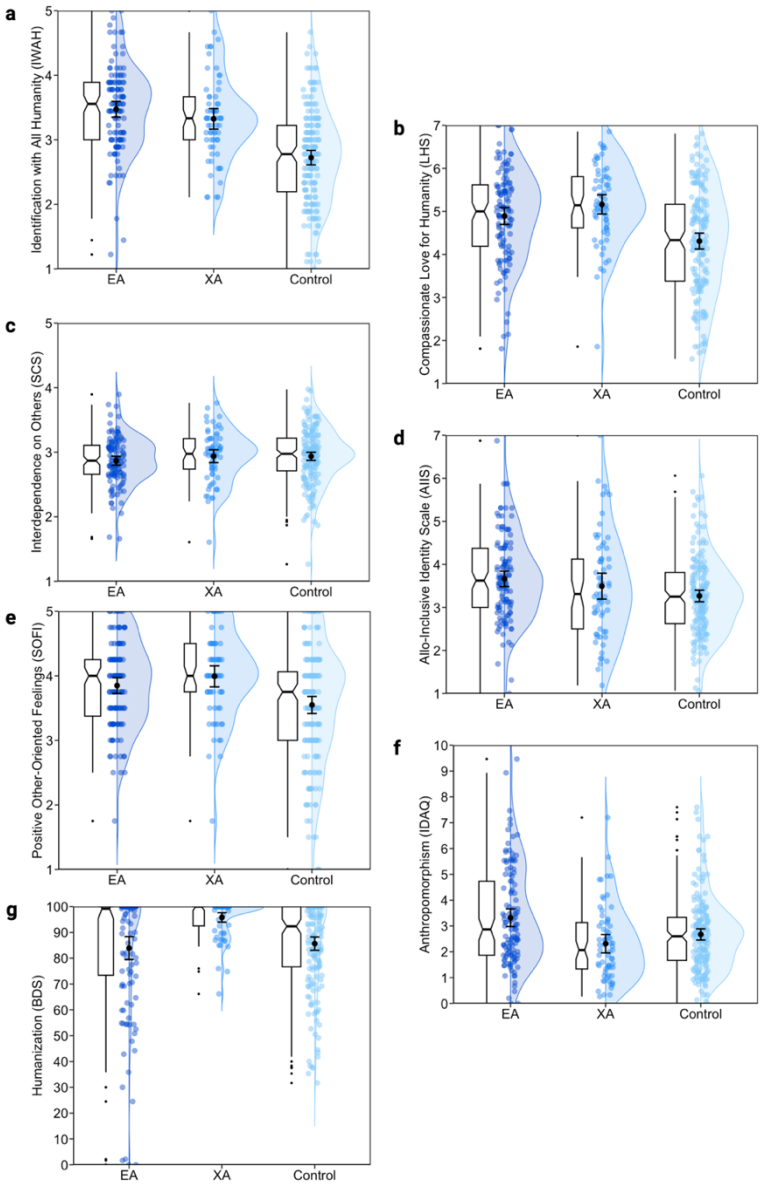
Table 9

Results from one-way ANOVAs with post-hoc tests employing the Bonferroni alpha correction comparing Samples 1-3 on measures other-inclusive identity and positive other-oriented attitudes.

Sample	Other-Inclusive Identity
Identification with All Humanity (IWAHs)	
Omnibus	$F(2, 357) = 44.60$, $p < .001$, $\eta^2p = .200$
EA vs. Control	$t(357) = 8.93$, $p < .001$, *Cohen's d* = 1.059
XA vs. Control	$t(357) = 5.85$, $p < .001$, *Cohen's d* = 0.849
EA vs. XA	$t(357) = 1.36$, $p = .523$, *Cohen's d* = 0.210
Compassionate Love for Humanity (LHS)	
Omnibus	$F(2, 357) = 17.50$, $p < .001$, $\eta^2p = .089$
EA vs. Control	$t(357) = 4.36$, $p < .001$, *Cohen's d* = 0.517
XA vs. Control	$t(357) = 5.21$, $p < .001$, *Cohen's d* = 0.756
EA vs. XA	$t(357) = -1.55$, $p = .367$, *Cohen's d* = -0.239
Interdependence on Others (SCS)	
Omnibus	$F(2, 357) = 1.20$, $p = .302$, $\eta^2p = .007$
EA vs. Control	$t(357) = -1.45$, $p = .446$, *Cohen's d* = -0.172
XA vs. Control	$t(357) = 0.05$, $p > .999$, *Cohen's d* = 0.007
EA vs. XA	$t(357) = -1.16$, $p = .743$, *Cohen's d* = -0.179
Inclusion of All Beings in Self-Concept (AIIS)	
Omnibus	$F(2, 357) = 5.79$, $p = .003$, $\eta^2p = .031$
EA vs. Control	$t(357) = 3.37$, $p = .003$, *Cohen's d* = 0.400
XA vs. Control	$t(357) = 1.57$, $p = .354$, *Cohen's d* = 0.227
EA vs. XA	$t(357) = 1.12$, $p = .792$, *Cohen's d* = 0.173
Sample	**Positive Other-Oriented Attitudes**
Positive Feelings Towards Others (SOFI)	
Omnibus	$F(2, 357) = 9.70$, $p < .001$, $\eta^2p = .052$
EA vs. Control	$t(357) = 3.22$, $p = .004$, *Cohen's d* = 0.383
XA vs. Control	$t(357) = 3.90$, $p < .001$, *Cohen's d* = 0.566
EA vs. XA	$t(357) = -1.19$, $p = .709$, *Cohen's d* = -0.183
Individual Differences in Anthropomorphism (IDAQ)	
Omnibus	$F(2, 357) = 9.73$, $p < .001$, $\eta^2p = .052$

EA vs. Control	$t(357) = 3.39$, $p = .002$, *Cohen's d* = 0.403
XA vs. Control	$t(357) = -1.54$, $p = .373$, *Cohen's d* = -0.224
EA vs. XA	$t(357) = 4.06$, $p < .001$, *Cohen's d* = 0.626
Humanization of Stigmatized Groups (BDS)	
Omnibus	$F(2, 357) = 9.20$, $p < .001$, $\eta^2p = .049$
EA vs. Control	$t(357) = -0.77$, $p > .999$, *Cohen's d* = -0.092
XA vs. Control	$t(357) = 3.73$, $p < .001$, *Cohen's d* = 0.541
EA vs. XA	$t(357) = -4.10$, $p < .001$, *Cohen's d* = -0.633

Figure 11. *Differences between effective altruists, extraordinary altruists and general population controls on measures of other inclusive identity and positive other-oriented attitudes*



Note. *Raincloud plots displaying levels of other-inclusive identity (**a-d**) and positive other-oriented attitudes (**e-g**) across Samples 1-3. Plots display individual data-points, jittered for readability, with overlaid split-violins to illustrate the shape of the underlying probability distributions. Means and error bars depicting 95% CIs are also included, as well as box plots with notches to convey 95% CIs around the medians.*

3.2. Phase 2 Results

Phase 1 revealed that both effective altruists (EAs) and extraordinary altruists (XAs)–who engage in notably prosocial acts in the real world–also score higher on well-validated laboratory tasks and measures that capture equitable and effective prosocial attitudes and behaviors. Beyond validating these two populations as exceptionally caring, it's especially noteworthy that XAs, whose real-world altruism and explicit *attitudes*–as measured in the present research (see findings above pertaining to the EAIS)–align more closely with the principle of altruistic equity than altruistic effectiveness, also prioritized effectiveness on laboratory metrics capturing prosocial *behavior* to a greater extent than controls. Moreover, the altruistic engagement of XAs, rather than being constrained to singular profound acts of altruism, appears to run through their lives more comprehensively, as evidenced by this population reporting greater engagement in real-world philanthropy and volunteerism relative to controls, even when taking into account differences in levels of income.

Phase 1 also allowed insight into differences across the three subject groups on other psychological properties that have known associations with prosociality in the general population (i.e., empathic ability and reasoning ability, moral beliefs and values, other-inclusive identity, and positive other-oriented attitudes; e.g., Batson & Ahmad, 2009; Caviola et al., 2021; Graham et al., 2017; B. A. Jones, 2022). With some exceptions (see results above and Table 1), both special populations scored higher on (1) moral beliefs and values associated with equitable attitudes and prosocial behaviors in the general population (e.g., moral concern on the MES, impartial beneficence (IB) on the OUS), (2) lower on those associated with parochiality and ingroup favoritism in the general population (e.g., loyalty on the MFQ), and (3) higher on measures of other-inclusive identity (e.g., IWAH, the LHS) and positive other-oriented attitudes (e.g., positive feelings towards others on the SOFI, humanization on the BDS). Consistent with

existing literature on the affective abilities of XAs (Law, Amormino, et al., 2024; Marsh, 2019), altruistic organ donors scored higher than general population controls on measures of empathic ability. Furthermore, despite existing literature having been centrally focused on the affective but not reasoning abilities of this population to date, I find that the XAs outperform the general population on certain deliberative reasoning tasks (e.g., Heuristics and Bias Tasks), while showing lower performance on the Cognitive Reflection Test and no significant difference on Need for Cognition.

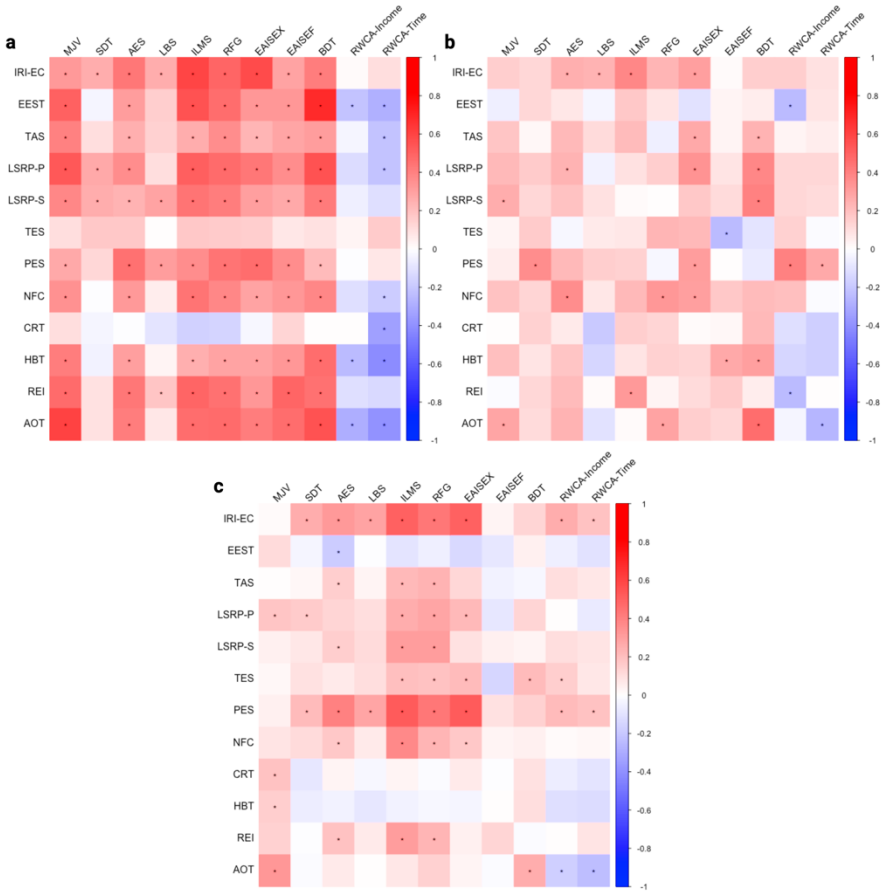
On the other hand, EAs–who have been woefully understudied in the extant literature and whose proponents champion deliberative reasoning over empathy as a better-suited tool for driving equitable and effective altruistic action (Schubert & Caviola, 2024; Singer, 2015, 2016)–do show some enrichments in reasoning ability relative to ordinary adults. Moreover, this population at times shows deficits in emotional ability, particularly in the capacity to accurately identify emotional content in written statements. However, countering the perspective that empathy invariably impedes equitable and effective altruism, a perspective that has proliferated in much of the recent psychological literature (Bloom, 2016; Bruneau et al., 2017; J. Greene, 2014; Prinz, 2011), the present findings suggest instead that these two distinct psychological capacities may both serve as viable routes to these ends. Namely, in studying these two populations together for the first time, I find that the minds of EAs differ most from ordinary individuals on capacities in reasoning, whereas the minds of XAs differ most on capacities in empathy. In other words, based on the divergence observed in trait levels of these affective and cognitive capacities across the two exceptionally altruistic populations, it seems that reasoning and empathy are not necessarily countervailing, but may both be forces that can serve to support equitable and effective good.

Here, in Phase 2, by examining each of the psychological features above as predictors of prosocial outcomes within samples 1, 2, and 3, I further build out our current understanding of the possibility raised above–that a plurality of cognitive *and* affective routes towards altruistic equity and effectiveness may be viable. Specifically, tested how well individual differences in the empathic and rational abilities of EAs, XAs, and controls can account for individual differences in prosociality within each subject group. Moreover, as a more secondary aim, I investigated how well individual differences in morality, other-inclusive identity, and other-oriented attitudes predict differences in prosociality within each sample. Thus, the overarching goal of Phase 2 was to investigate the array of psychological features that underlie equitable and effective prosociality in exceptionally caring and ordinary subjects alike. These insights help to reconcile ongoing debates over the roles of affect and cognition in altruistic behavior (Bloom, 2016; Fowler et al., 2021; Zaki, 2018). They may also serve as a starting point going forward for research to develop interventions cultivating equitable and effective altruism in the general population, targeting the most predictive psychological capacities identified here, should doing so be deemed a societal objective, to realize the applied benefits of the present insights, or an academic objective, to shed light on the causal roles of empathy and reasoning in the context of the present findings.

3.2.1. The Cognitive and Affective Architecture of Equitable and Effective Altruism. I began by conducting an exploratory and exhaustive set of bivariate correlational analyses between each measure of empathic and reasoning ability with each prosocial outcome within each sample (see Figure 12). Generally speaking, across samples, most measures of empathic and reasoning ability were associated positively with most attitudes and behaviors in line with equitable and effective prosociality. Intriguingly, these bivariate relationships were strongest among effective altruists, even with respect to empathic predictors. However, relationships with

measures of *overall* real-world prosociality, without regard to equitability or effectiveness, were generally weaker and at times negative, suggesting affective and cognitive abilities may be more closely associated with the architecture of equitable and effective altruism than altruism when broadly construed. It is also worth noting that in the general population, empathy appeared to be a stronger predictor of equitable and effective altruism than reasoning. These finding stand in contrast to perspectives on the inherent parochiality of empathy (Bloom, 2016) and the primacy of reasoning (Schubert & Caviola, 2024) in the prosocial context. They suggest instead that our ability to empathize with others may live among the most critical tools at our disposal–rather than among our greatest limitations–for promoting the greater good through beneficent action. The numerical results from these analyses are presented in full in the "Analysis Script" on the OSF.

Figure 12. *Bivariate relationships between empathic ability and reasoning ability with equitable, effective and overall prosociality among EAs, XAs, and controls*



Note. *Heatmaps displaying correlation coefficients from -1 (blue) to +1 (red) in EAs (**a**), XAs (**b**), and controls (**c**). Asterisks correspond to statistically significant relationships.*

For the focal analysis, I conducted a series of 11 multiple regression models per sample (one per prosocial outcome, 33 in total), entering each measure of empathy and reasoning as a simultaneous predictor. The goal of estimating these models was to ascertain which predictors accounted for unique variance in the outcomes, above and beyond the effects of the other predictors. As pre-registered, because of the vast number of number of predictors included, I evaluated the variance inflation factor (VIF) for each predictor, which ranged from 1.11 to 2.34,

indicating the predictors were only moderately correlated with one another and that multicollinearity was not problematic (Thompson et al., 2017).

Within each sample, these results largely confirmed that both empathy and reasoning ability can be a force for good (see Table 10). Notably, empathic capacities were positively associated with: (1) generosity on the Social Discounting Task for EAs and XAs; (2) altruistic enkrateia (AES) for EAs and controls; (3) intergenerational concern on the LBS, RFG and ILMS for all three samples; (4) attitudes aligned with expansive altruism (altruistic equity) on the EAIS for EAs and controls; and intriguingly, with (5) attitudes aligned with effectiveness prioritization on the EAIS for EAs, (6) behavioral donations to effective charitable causes across all three samples, (7) real world monetary charitable contributions (as the percentage of income donated in a given year) for XAs and controls, and (8) with real-world time invested volunteering (as the percentage of time spent volunteering in a given year) in XAs. Beyond the findings presented in Phase 1, which demonstrate that exceptionally caring individuals show enhancements in empathy, these findings suggest that among exceptional altruists and ordinary adults alike, greater empathy and emotionality often *predicts* greater engagement and attitudes in line with altruistic equity, effectiveness, and real-world charitable action, above and beyond differences in reasoning ability. Critically, they provide clear evidence that arguments against empathy may be misguided (e.g., Bloom, 2016; Singer, 2016), as greater ability to empathize with others appears to underlie rather than inhibit altruistic attitudes and actions that transcend parochial boundaries and maximize impact.

Table 10

Results from multiple linear regression models evaluating associations between individual variation in empathy and reasoning with individual variation in equitable, effective, and real-world prosociality within Samples 1, 2, and 3.

Outcome Predictor	**EA** (R^2 = 0.47)		**XA** (R^2 = 0.20)		**Control** (R^2 = 0.16)	
MJV	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	-0.013	0.898	0.255	0.168	-0.081	0.424
EEST	0.103	0.387	-0.116	0.402	0.028	0.727
TAS	0.080	0.447	-0.030	0.859	-0.154	0.130
LSRP-P	0.072	0.610	-0.087	0.632	0.129	0.234
LSRP-S	0.024	0.832	0.214	0.217	-0.028	0.802
TES	0.036	0.653	0.073	0.621	0.020	0.790
PES	0.159	0.086	-0.105	0.564	0.111	0.223
NFC	-0.139	0.199	0.157	0.336	-0.142	0.169
CRT	0.033	0.700	-0.108	0.508	0.105	0.193
HBT	0.011	0.924	0.255	0.109	0.052	0.542
REI	**0.237**	**0.029**	-0.141	0.367	**0.216**	**0.043**
AOT	**0.368**	**0.013**	0.158	0.440	**0.254**	**0.005**

Outcome Predictor	**EA** (R^2 = 0.23)		**XA** (R^2 = 0.28)		**Control** (R^2 = 0.11)	
SDT	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	0.143	0.261	-0.216	0.218	0.166	0.111
EEST	**-0.438**	**0.003**	0.085	0.518	-0.032	0.694
TAS	-0.156	0.218	-0.168	0.292	-0.076	0.466
LSRP-P	**0.426**	**0.013**	0.126	0.463	0.141	0.207
LSRP-S	0.247	0.067	0.067	0.683	-0.009	0.938
TES	0.021	0.827	-0.006	0.967	0.020	0.801
PES	-0.061	0.584	**0.603**	**< .001**	0.069	0.464
NFC	-0.120	0.357	-0.088	0.567	0.143	0.178
CRT	0.009	0.932	0.155	0.318	-0.071	0.393
HBT	0.047	0.740	-0.025	0.869	-0.016	0.852
REI	0.151	0.247	0.184	0.218	-0.114	0.299
AOT	-0.013	0.939	0.165	0.396	-0.069	0.460

Outcome Predictor	**EA** (R^2 = 0.43)		**XA** (R^2 = 0.27)		**Control** (R^2 = 0.25)	
AES	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	**0.265**	**0.016**	0.168	0.341	0.174	0.070
EEST	-0.088	0.472	-0.037	0.780	**-0.248**	**0.001**
TAS	0.066	0.544	-0.016	0.922	0.061	0.523
LSRP-P	-0.116	0.426	0.045	0.796	-0.012	0.906
LSRP-S	-0.002	0.987	-0.031	0.853	-0.042	0.685
TES	0.047	0.565	-0.133	0.351	-0.059	0.409

PES	**0.307**	**0.002**	0.214	0.225	**0.311**	**< .001**
NFC	0.037	0.742	0.250	0.114	-0.068	0.486
CRT	-0.025	0.778	-0.008	0.957	0.079	0.298
HBT	**0.268**	**0.028**	0.086	0.569	0.005	0.953
REI	0.185	0.099	0.108	0.470	0.147	0.146
AOT	0.114	0.450	0.177	0.369	0.077	0.368

Outcome Predictor	**EA** (R^2 = 0.23)		**XA** (R^2 = 0.17)		**Control** (R^2 = 0.12)	
LBS	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	0.193	0.131	**0.397**	**0.039**	**0.214**	**0.040**
EEST	0.061	0.668	0.012	0.929	0.011	0.890
TAS	0.007	0.954	0.008	0.962	-0.128	0.218
LSRP-P	-0.307	0.073	-0.263	0.159	-0.025	0.825
LSRP-S	**0.353**	**0.01**	0.099	0.576	0.090	0.426
TES	-0.103	0.285	0.149	0.328	0.029	0.706
PES	**0.295**	**0.009**	-0.167	0.373	**0.191**	**0.042**
NFC	0.008	0.949	0.191	0.254	-0.087	0.407
CRT	-0.010	0.926	-0.229	0.172	0.027	0.742
HBT	0.124	0.377	0.003	0.985	0.000	0.997
REI	0.007	0.958	-0.041	0.795	0.070	0.521
AOT	-0.073	0.678	-0.037	0.858	-0.017	0.853

Outcome Predictor	**EA** (R^2 = 0.62)		**XA** (R^2 = 0.36)		**Control** (R^2 = 0.42)	
ILMS	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	**0.422**	**< .001**	**0.566**	**0.001**	**0.258**	**0.002**
EEST	**0.256**	**0.012**	0.078	0.528	-0.123	0.065
TAS	-0.171	0.057	0.157	0.299	-0.051	0.550
LSRP-P	-0.018	0.882	-0.063	0.700	0.022	0.805
LSRP-S	0.149	0.117	-0.186	0.232	0.087	0.344
TES	-0.030	0.657	0.093	0.487	0.048	0.449
PES	0.051	0.514	-0.162	0.326	**0.323**	**< .001**
NFC	0.145	0.116	0.228	0.122	0.084	0.328
CRT	-0.144	0.052	0.147	0.317	0.074	0.269
HBT	0.122	0.220	0.170	0.230	0.023	0.749
REI	**0.233**	**0.012**	0.108	0.439	0.140	0.116
AOT	-0.081	0.515	-0.246	0.184	-0.035	0.641

Outcome Predictor	**EA** (R^2 = 0.52)		**XA** (R^2 = 0.43)		**Control** (R^2 = 0.30)	
RFG	*β*	*p*	*β*	*p*	*β*	*p*

IRI-EC	0.198	0.050	**0.624**	**< .001**	**0.218**	**0.020**
EEST	0.064	0.569	0.106	0.366	-0.119	0.105
TAS	0.083	0.408	-0.278	0.054	0.008	0.932
LSRP-P	-0.096	0.473	-0.134	0.384	0.011	0.913
LSRP-S	0.127	0.233	0.023	0.877	0.134	0.186
TES	-0.015	0.845	**0.316**	**0.015**	0.076	0.273
PES	**0.290**	**0.001**	**-0.409**	**0.011**	**0.262**	**0.002**
NFC	0.114	0.267	**0.521**	**< .001**	-0.055	0.561
CRT	**-0.176**	**0.035**	-0.165	0.233	0.001	0.992
HBT	0.172	0.122	0.255	0.059	0.055	0.478
REI	0.116	0.258	-0.218	0.104	0.092	0.344
AOT	0.107	0.444	0.043	0.802	0.082	0.325

Outcome Predictor	**EA** (R^2 = 0.50)		**XA** (R^2 = 0.33)		**Control** (R^2 = 0.40)	
EAIS-EX	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	**0.492**	**< .001**	0.204	0.229	**0.336**	**< .001**
EEST	-0.113	0.324	-0.151	0.235	**-0.171**	**0.012**
TAS	-0.074	0.466	0.112	0.465	0.075	0.381
LSRP-P	-0.089	0.513	0.242	0.148	0.065	0.476
LSRP-S	0.152	0.158	-0.079	0.617	-0.139	0.134
TES	-0.031	0.683	0.171	0.211	0.068	0.284
PES	**0.264**	**0.004**	0.096	0.566	**0.349**	**< .001**
NFC	0.098	0.350	0.241	0.110	-0.004	0.962
CRT	-0.005	0.955	0.011	0.941	0.115	0.090
HBT	**0.374**	**0.001**	0.212	0.145	0.063	0.378
REI	0.015	0.889	-0.001	0.993	-0.056	0.535
AOT	0.012	0.932	-0.120	0.521	-0.022	0.773

Outcome Predictor	**EA** (R^2 = 0.40)		**XA** (R^2 = 0.17)		**Control** (R^2 = 0.09)	
EAIS-EF	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	0.079	0.482	0.046	0.804	0.057	0.587
EEST	-0.053	0.673	-0.092	0.513	-0.068	0.410
TAS	0.057	0.607	-0.165	0.334	-0.153	0.150
LSRP-P	-0.063	0.676	0.022	0.904	-0.156	0.168
LSRP-S	0.042	0.723	0.084	0.632	0.126	0.270
TES	-0.005	0.948	-0.277	0.072	**-0.175**	**0.027**
PES	**0.274**	**0.006**	0.101	0.587	0.143	0.133
NFC	-0.048	0.678	0.120	0.469	-0.092	0.390
CRT	0.083	0.371	-0.008	0.960	0.004	0.961

HBT	0.099	0.424	0.190	0.238	0.037	0.677
REI	**0.288**	**0.013**	0.114	0.476	0.195	0.079
AOT	**0.245**	**0.117**	0.075	0.720	0.054	0.565

Outcome Predictor	**EA** (R^2 = 0.57)		**XA** (R^2 = 0.39)		**Control** (R^2 = 0.15)	
BDT	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	**0.211**	**0.028**	0.207	0.203	0.086	0.399
EEST	**0.423**	**< .001**	-0.119	0.325	-0.002	0.980
TAS	-0.174	0.069	-0.028	0.847	-0.105	0.306
LSRP-P	0.054	0.673	0.063	0.692	-0.039	0.720
LSRP-S	0.117	0.247	**0.327**	**0.034**	0.050	0.651
TES	0.014	0.843	-0.106	0.415	**0.175**	**0.023**
PES	-0.001	0.994	-0.119	0.457	0.116	0.207
NFC	-0.065	0.509	0.080	0.574	-0.087	0.404
CRT	-0.047	0.544	0.173	0.227	0.029	0.720
HBT	**0.215**	**0.042**	0.234	0.093	0.099	0.248
REI	**0.201**	**0.041**	-0.133	0.330	-0.039	0.718
AOT	-0.005	0.972	0.158	0.379	**0.268**	**0.004**

Outcome Predictor	**EA** (R^2 = 0.15)		**XA** (R^2 = 0.33)		**Control** (R^2 = 0.15)	
RWCA ($)	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	0.083	0.533	-0.034	0.842	**0.264**	**0.010**
EEST	-0.034	0.821	-0.163	0.200	0.004	0.964
TAS	0.179	0.179	-0.134	0.386	0.086	0.404
LSRP-P	0.149	0.403	0.120	0.468	-0.125	0.252
LSRP-S	-0.075	0.596	0.113	0.476	0.074	0.506
TES	0.022	0.828	-0.025	0.852	0.108	0.159
PES	0.033	0.779	**0.338**	**0.047**	0.068	0.462
NFC	0.122	0.373	**0.313**	**0.039**	-0.050	0.632
CRT	0.151	0.169	0.039	0.795	0.028	0.728
HBT	-0.123	0.402	-0.065	0.651	0.022	0.797
REI	-0.033	0.806	-0.262	0.071	-0.038	0.725
AOT	**-0.484**	**0.010**	-0.060	0.750	**-0.188**	**0.040**

Outcome Predictor	**EA** (R^2 = 0.29)		**XA** (R^2 = 0.24)		**Control** (R^2 = 0.14)	
RWCA (Time)	*β*	*p*	*β*	*p*	*β*	*p*
IRI-EC	0.222	0.070	-0.230	0.204	0.190	0.065
EEST	-0.075	0.583	0.151	0.267	-0.029	0.722

TAS	-0.027	0.823	-0.038	0.815	0.036	0.724
LSRP-P	0.045	0.783	**0.383**	**0.034**	-0.141	0.199
LSRP-S	-0.068	0.595	0.046	0.784	0.060	0.587
TES	0.061	0.509	-0.063	0.663	0.038	0.623
PES	-0.076	0.476	0.249	0.167	0.109	0.241
NFC	0.056	0.657	0.062	0.696	-0.090	0.390
CRT	-0.203	0.045	0.037	0.816	-0.005	0.950
HBT	-0.073	0.590	-0.166	0.283	0.025	0.770
REI	0.104	0.403	0.023	0.883	0.125	0.246
AOT	**-0.398**	**0.020**	-0.384	0.059	**-0.215**	**0.020**

Note. *Bolded effects indicate statistical significance at $p < .05$.*

Furthermore, these results partially align with assertions raised in discourse related to the effective altruism movement (MacAskill, 2018) and with some earlier empirical findings (Caviola et al., 2021) suggesting that reasoning ability underlies the prioritization of altruistic equity and effectiveness. Namely, variation in reasoning ability was positively associated with: (1) moral judgments of equitable and effective prosociality on the MJV task for EAs and controls; (2) altruistic enkrateia (AES) for EAs, (3) intergenerational concern on the ILMS for EAs and on the RFG for XAs; (4) the prioritization of expansiveness (equity) and effectiveness in altruism on the EAIS for EAs; (5) behavioral donations to effective causes for EAs and Controls; and (6) real-world monetary contributions to charity in XAs. However, it is noteworthy that the associations between reasoning ability with equitable and effective prosociality were most pronounced among members of the EA subject group, who explicitly emphasize applying reasoning skills to guide altruistic decision-making.

Taken together, these findings largely support that empathy *and* reasoning, rather than one and not the other, are associated with greater altruistic equity, effectiveness, and real-world charitable action in most cases. Additionally, these findings *do not* support that empathy is invariably a constraint on the scope of equity and impact. Nonetheless, in some cases, affective capacities did associate negatively with prosociality. Namely, although not a measure of empathy

in particular, but emotional ability more broadly, the ability to correctly identify emotions in written statements on the EEST was negatively associated with generosity on the SFT for EAs, as well as with altruistic enkrateia and the prioritization of expansiveness in altruism for controls. Intriguingly, scores on this particular emotional capacity were markedly lower for EAs compared to the other two subject groups (see Phase 1 results), suggesting that high levels of this emotional capacity may indeed constrain certain prosocial attitudes. Conversely, prior research has shown scores on this measure to be *negatively* associated with psychopathy and callus unemotional traits in adolescents, which are well known to predict *antisocial* behavior (Marsh & Cardinale, 2012). Thus, further investigation is warranted to better elucidate these unexpected findings.

Also of note is that, among XA subjects, greater outgroup empathy predicted lower scoring on the RFG, which measures one's sense of duty to safeguard the welfare of future generations. This suggests that for members of this subject group in particular, feeling greater empathy for those who are suffering today may come at the expense of feeling less responsible for the suffering of those who will come tomorrow. This is particularly intriguing, as this pattern is reversed in both general population and EA subjects. Moreover, this finding counters insights from my previous research in the general population, which demonstrates across highly-powered samples that ordinary adults who feel greater concern for future generations also tend to feel greater concern for socially distant others, minoritized groups, friends, and family members in the present day (Law, Syropoulos, Coleman, et al., 2023; Syropoulos, Law, Amormino, et al., 2024a; Syropoulos, Law, & Young, 2024d).

Finally, it's worth mentioning that controls who reported lay theories of empathy as being more malleable within individuals scored significantly lower in their reported prioritization of effectiveness in altruism. Prior research has shown that individuals who espouse more malleable versus fixed theories of empathy generally exert greater empathic effort towards distant and

stigmatized targets (Schumann et al., 2014). Thus, one might expect expansive views of empathy to associate *positively* with effectiveness prioritization. Further research might address whether this unexpected finding could be indicative of something akin to an empathic bystander effect (Darley & Latane, 1968), where holding stronger beliefs that empathy is malleable allows people to diffuse personal responsibility for mass suffering more easily to others. In other words, because this measure captures not only beliefs that empathy is malleable for oneself, but for others as well, it's possible that people who score higher on this measure tend to *overestimate* the empathy of others, anticipating that others will prioritize effectiveness in their stead.

Deviations from the expected results were observed not only with regard to measures of empathic ability but reasoning ability as well. Specifically, scores on the Cognitive Reflection Test (CRT) were negatively associated with a sense of responsibility to future generations among EAs. Additionally, Actively Open-Minded Thinking (AOT)–willingness to change one's mind when faced with new evidence–was moderately negatively associated with real-world charitable donations and volunteerism for EAs and weakly negatively associated for controls. These findings suggest that, in certain contexts, reasoning abilities may backfire, perhaps by increasing focus on the opportunity costs of giving, particularly when giving equitably, even among those who typically prioritize effectiveness in their altruistic efforts. Nonetheless, these findings were exceptions to the general trend observed across samples and measures, where both reasoning ability and empathic ability were typically positively associated with equitable and effective prosocial attitudes and behaviors.

To explore these patterns in greater depth, I next investigated whether reasoning and empathy interact within each population concerning their associations with measures of prosociality. This analysis aimed to determine if the combined influence of reasoning and empathy contributes to prosocial behaviors differently than either trait alone. As pre-registered,

prior to these analyses, I reduced the dimensionality of both categories of predictors (see pages 149-154 of the "Analysis Script" on the OSF for the full results from the dimension reduction analyses. First, I conducted Exploratory Factor Analysis (EFA) using the maximum likelihood extraction method in combination with oblimin rotation on the battery of measures of empathic ability and reasoning ability, separately. For empathic ability, the measures loaded onto a single factor, but outgroup empathy on the PES and scores on the EEST had weak loadings, below 0.5, and thus were excluded from the next stage of analysis. Likewise, for reasoning ability, each measure loaded onto a single factor, but scores on the CRT had loadings below 0.5, and thus were excluded from the next stage of analysis.

Afterward, I conducted Confirmatory Factor Analysis (CFA) on the retained measures, employing the maximum likelihood extraction method to confirm their factor structure. The CFAs revealed a CFI of 0.93 for empathic ability, indicating good fit, and a CFI of 0.84 for reasoning ability, indicating acceptable fit (Kim et al., 2016). Thus, for empathic ability, scores on the TAS (reverse-coded), IRI-EC, LSRP-P (reverse-coded), and LSRP-S (reverse-coded) were retained, z-transformed, and averaged (*Cronbach's* $\alpha = 0.79$). For reasoning ability, scores on the NFC, IRI, AOT, and HBT were retained, z-transformed, and averaged (*Cronbach's* $\alpha = 0.69$). After reducing dimensionality, 33 multiple regression analysis mirroring those above were estimated, this time specifying as predictors the empathic ability composite, reasoning ability composite, and their interaction. See Table 11 for the results from these models.

Table 11

Results from multiple linear regression models evaluating associations between individual variation in the empathic ability factor, the reasoning ability factor, and their interaction with individual variation in equitable, effective, and real-world prosociality within Samples 1, 2, and 3.

Outcome Predictor	**EA** (R^2 = 0.41)		**XA** (R^2 = 0.18)		**Control** (R^2 = 0.07)	
MJV	*β*	*p*	*β*	*p*	*β*	*P*
Empathic Ability	**0.370**	**0.001**	0.168	0.177	-0.046	0.570
Reasoning Ability	-0.150	0.158	0.156	0.201	**0.283**	**< .001**
Interaction	0.004	0.964	**-0.244**	**0.027**	0.032	0.693
Outcome Predictor	**EA** (R^2 = 0.09)		**XA** (R^2 = 0.04)		**Control** (R^2 = 0.04)	
SDT	*β*	*p*	*β*	*p*	*β*	*P*
Empathic Ability	0.198	0.076	0.094	0.481	**0.203**	**0.015**
Reasoning Ability	**0.277**	**0.031**	0.130	0.322	-0.087	0.299
Interaction	-0.076	0.398	-0.066	0.574	0.072	0.387
Outcome Predictor	**EA** (R^2 = 0.25)		**XA** (R^2 = 0.19)		**Control** (R^2 = 0.10)	
AES	*β*	*p*	*β*	*p*	*β*	*P*
Empathic Ability	**0.234**	**0.021**	0.216	0.083	**0.221**	**0.006**
Reasoning Ability	**0.336**	**< .001**	**0.302**	**0.015**	0.023	0.781
Interaction	0.003	0.973	-0.055	0.612	**0.225**	**0.006**
Outcome Predictor	**EA** (R^2 = 0.06)		**XA** (R^2 = 0.04)		**Control** (R^2 = 0.04)	
LBS	*β*	*p*	*β*	*p*	*β*	*P*
Empathic Ability	**0.268**	**0.018**	0.192	0.153	**0.208**	**0.013**
Reasoning Ability	-0.035	0.745	-0.093	0.481	-0.076	0.366
Interaction	-0.009	0.931	0.100	0.393	0.024	0.772
Outcome Predictor	**EA** (R^2 = 0.41)		**XA** (R^2 = 0.10)		**Control** (R^2 = 0.19)	
ILMS	*β*	*p*	*β*	*p*	*β*	*P*
Empathic Ability	**0.371**	**< .001**	0.160	0.219	**0.383**	**< .001**
Reasoning Ability	**0.273**	**0.002**	0.180	0.160	0.101	0.190
Interaction	-0.131	0.097	-0.102	0.366	-0.043	0.578

Outcome Predictor	**EA** (R^2 = 0.37)		**XA** (R^2 = 0.09)		**Control** (R^2 = 0.17)	
RFG	*β*	*p*	*β*	*p*	*β*	*P*
Empathic Ability	**0.369**	**< .001**	0.052	0.687	**0.391**	**< .001**
Reasoning Ability	**0.276**	**0.002**	**0.295**	**0.024**	0.046	0.556
Interaction	-0.079	0.334	0.057	0.617	0.003	0.966
Outcome Predictor	**EA** (R^2 = 0.28)		**XA** (R^2 = 0.17)		**Control** (R^2 = 0.12)	
EAIS-EX	*β*	*p*	*β*	*p*	*β*	*P*
Empathic Ability	**0.385**	**< .001**	**0.315**	**0.014**	**0.326**	**< .001**
Reasoning Ability	**0.217**	**0.024**	0.170	0.167	-0.083	0.302
Interaction	0.020	0.815	-0.036	0.738	**0.185**	**0.021**
Outcome Predictor	**EA** (R^2 = 0.28)		**XA** (R^2 = 0.07)		**Control** (R^2 = 0.04)	
EAIS-EF	*β*	*p*	*β*	*p*	*β*	*p*
Empathic Ability	0.157	0.112	0.004	0.976	-0.055	0.507
Reasoning Ability	**0.415**	**< .001**	0.257	0.050	0.056	0.501
Interaction	-0.027	0.752	-0.046	0.687	**0.211**	**0.012**
Outcome Predictor	**EA** (R^2 = 0.50)		**XA** (R^2 = 0.26)		**Control** (R^2 = 0.02)	
BDT	*β*	*p*	*β*	*p*	*β*	*p*
Empathic Ability	**0.218**	**0.009**	**0.292**	**0.015**	0.032	0.705
Reasoning Ability	**0.347**	**< .001**	**0.278**	**0.019**	0.136	0.107
Interaction	**-0.324**	**< .001**	-0.128	0.216	-0.018	0.827
Outcome Predictor	**EA** (R^2 = 0.13)		**XA** (R^2 = 0.05)		**Control** (R^2 = 0.07)	
RWCA ($)	*β*	*p*	*β*	*p*	*β*	*p*
Empathic Ability	0.182	0.095	0.159	0.235	**0.248**	**0.003**
Reasoning Ability	**-0.239**	**0.023**	-0.142	0.279	**-0.195**	**0.019**
Interaction	**0.279**	**0.004**	-0.105	0.369	-0.117	0.153
Outcome Predictor	**EA** (R^2 = 0.13)		**XA** (R^2 = 0.09)		**Control** (R^2 = 0.04)	
RWCA (Time)	*β*	*p*	*β*	*p*	*β*	*p*
Empathic Ability	0.086	0.429	0.149	0.253	0.163	0.051
Reasoning Ability	**-0.380**	**< .001**	-0.237	0.067	-0.146	0.081
Interaction	0.050	0.600	-0.172	0.131	-0.121	0.147

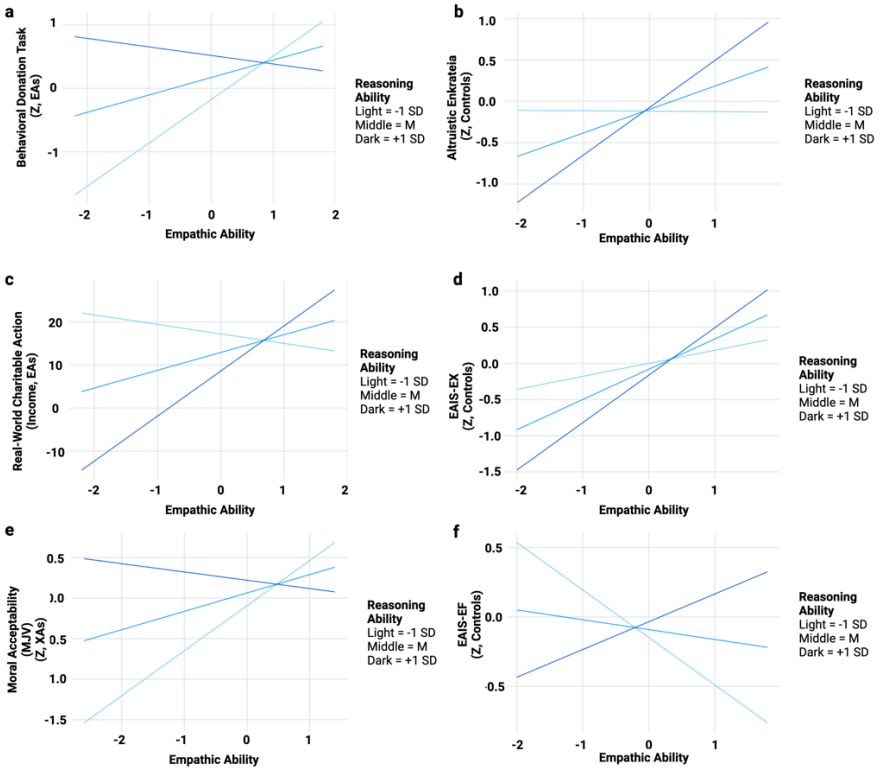
Note. *Bolded effects indicate statistical significance at p < .05.*

Here, largely in line with the findings above where each facet of empathic ability was considered separately scores on the empathic ability factor were associated with greater: (1) moral acceptability judgments of equitable and effective altruism on the MJV task for EAs; (2) generosity on the SDT for controls; (3) altruistic enkrateia (AES) for EAs and controls; (4) intergenerational concern on the LBS, RFG and ILMS for EAs and controls; (5) attitudes aligned with expansive altruism (altruistic equity) on the EAIS for all three subject groups; and intriguingly, with (6) behavioral donations to effective charitable causes across EAs and XAs and (7) real-world monetary charitable contributions for controls. Likewise, scores on the reasoning ability factor we associated with greater: (1) generosity on the SDT for controls; (2) altruistic enkrateia for all three subject groups; (3) intergenerational concern on the ILMS for EAs and on the RFG for EAs and XAs; (4) prioritization of expansiveness and effectiveness on the EAIS in EAs; and (5) behavioral donations to effective causes in EAs and XAs. Yet, in a few instances, the reasoning ability factor showed negative associations with prosociality. Namely, reasoning ability was associated negatively with real-world monetary contributions to charity for EAs and Controls and with real-world time invested volunteering for EAs. Also notable is that empathic ability in no instances associated significantly and negatively with any aspect of prosociality in any sample.

There were, however, a few instances where significant interactions were observed between empathic and reasoning ability (see Figure 13). Specifically, these two capacities interacted with respect to their associations with behavioral donations (BDT) and real-world monetary charitable contributions among EAs, with moral judgments on the MJV task among XAs, and with altruistic enkrateia as well as the prioritization of expansiveness and effectiveness in altruism among controls. Specifically, with regard to the BDT and judgments on the MJV, empathic ability served as a protective factor against lower levels of reasoning ability among

EAs and XAs, respectively. In other words, at lower levels of reasoning ability, greater empathy associated more strongly and positively with prosociality. Yet, with regard to real-world charitable action (among EAs) as well as altruistic enkrateia and the prioritization of expansiveness (equity) in altruism (among controls), empathic ability was more strongly and positively related with prosociality at *higher* levels of reasoning ability.

Figure 13. *Interactions between empathic and reasoning ability on equitable, effective, and real-world prosociality*



Note. **(*a-e*)** *Plots displaying the interaction between empathy and reasoning on prosociality for EAs on the BDT and RWCA–Income, for XAs on the MJV, and for controls on the AES, and EAIS. Plots display predicted values for y across values of x.*

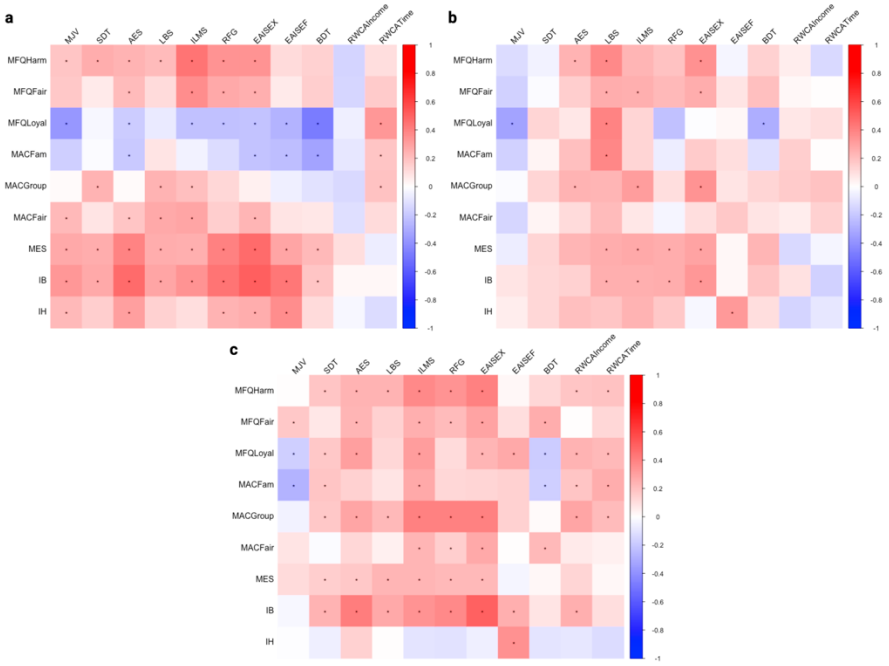
While no main effects were observed for either empathic or reasoning ability on the prioritization of effectiveness in altruism for the control sample, a significant crossover interaction was identified. Specifically, at lower levels of empathic ability, greater reasoning

ability was associated with less prioritization of effectiveness in altruism compared to lower reasoning ability. Conversely, at higher levels of empathic ability, greater reasoning ability was linked to a greater prioritization of effectiveness. Similarly, at lower levels of reasoning ability, empathy was negatively related to effectiveness prioritization, whereas it was positively associated at higher levels of reasoning ability. Taken together, these findings related to the independent and combined effects of empathy and reasoning suggest that both are crucial for fostering altruistic equity and effectiveness. They often work together to enhance prosocial behavior across contexts, including prioritizing welfare-maximizing impact, engaging and holding attitudes in line with providing aid to distant targets, and engaging in real-world charitable actions. However, the interplay between these traits is nuanced: in the general population, reasoning is predictive of greater impact prioritization only when empathic ability is high, and empathy influences impact prioritization more positively when reasoning ability is also high. This underscores the importance of developing both reasoning and empathy as complementary capacities that, when applied together, are more effective in promoting altruistic behavior than either on its own.

3.2.2. The Roles of Morality, Other-Inclusive Identity, and Other-Oriented Attitudes in Equitable and Effective Altruism. After assessing the cognitive and affective architecture of prosociality across the three subject groups, I proceeded to investigate relationships between each of the measures capturing moral beliefs and values and other-inclusive identity/positive other-oriented attitudes with prosociality. I began by focusing on the battery of predictors related to morality. Mirroring the procedures above, bivariate relationships were first assessed for exploratory purposes (see Figure 14 for a graphical depiction of these findings and the "Analysis Script" on the OSF for full results). Confirming my predictions, most equitable and prosocial moral values were associated positively with prosociality across samples. However, moral values

that have been associated with parochial bias and ingroup favoritism in prior research (e.g., see Graham et al., 2017 for review) showed more complex patterns of results than anticipated. While loyalty on the MFQ and familial loyalty on the MAC-Q associated negatively–as predicted–with most prosociality measures among EAs, group loyalty was comparatively more weakly associated, and at times positively, with prosociality among this subject group. Although these findings diverge with my initial predictions, they align with the perspective that loyalty, particularly at the group level, may not necessarily constrain the scope of altruism among EAs, whose groups routinely comprise other members of the EA movement. This perspective is elaborated upon in Phase 1 Results and again below. Nonetheless, positive bivariate associations with loyalty were also observed among members of the other two subject groups for some prosocial outcomes, suggesting that perhaps at a broader level, ascribing moral importance to upholding obligations to those who are close does not necessarily come at the cost of sidelining the needs of those who are distant and causes that are the most impactful.

Figure 14. *Bivariate relationships between moral beliefs and values with equitable, effective and overall prosociality among EAs, XAs, and controls*



Note. *Heatmaps displaying correlation coefficients from -1 (blue) to +1 (red) in EAs (**a**), XAs (**b**), and controls (**c**). Asterisks correspond to statistically significant relationships.*

After examining bivariate relationships, I conducted the focal analyses–33 multiple linear regression models (one per prosocial outcome within each sample)[6]. Where significant effects were observed, they largely confirmed my predictions, with some exceptions (see Table 12). Among EAs, concern for harm was linked to greater intergenerational concern, while loyalty generally associated negatively with the prioritization of equitable and effective prosocial behaviors. Moreover, greater moral concern on the MES predicted greater prioritization of expansiveness on the EAIS, while greater endorsement of impartial beneficence on the OUS predicted greater prioritization of effectiveness on the same scale. Nonetheless, group loyalty in particular (on the MAC-Q) among EAs predicted greater intergenerational concern, while loyalty on the MFQ associated positively with volunteerism.

Table 12

Results from multiple linear regression models evaluating associations between individual variation in moral beliefs and values with individual variation in equitable, effective, and real-world prosociality within Samples 1, 2, and 3.

Outcome Predictor	**EA** (R^2 = 0.29)		**XA** (R^2 = 0.22)		**Control** (R^2 = 0.16)	
MJV	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	0.062	0.637	-0.022	0.907	-0.066	0.518
MFQ-Fair	0.096	0.443	-0.222	0.243	**0.202**	**0.035**
MFQ-Loyal	**-0.409**	**< .001**	**-0.472**	**0.011**	-0.038	0.684
MACQ-Family	0.007	0.961	0.002	0.992	**-0.377**	**< .001**
MACQ-Group	0.117	0.349	0.242	0.180	0.182	0.077
MACQ-Fair	0.082	0.496	0.031	0.855	0.039	0.665
MES	0.044	0.656	-0.134	0.302	0.107	0.165

[6] This same procedure was followed with respect to moral beliefs and values and other-inclusive identity/positive other-oriented attitudes, separately. The VIFs for each predictor on all multiple regression models were well below 5, suggesting that multicollinearity would not be problematic.

IB	0.148	0.221	0.213	0.111	-0.062	0.451
IH	0.130	0.198	0.065	0.600	0.059	0.426

Outcome Predictor	**EA** (R^2 = 0.23)		**XA** (R^2 = 0.10)		**Control** (R^2 = 0.12)	
SDT	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	0.267	0.052	-0.132	0.514	0.021	0.841
MFQ-Fair	-0.083	0.522	-0.065	0.747	0.035	0.721
MFQ-Loyal	-0.068	0.551	0.209	0.282	0.038	0.695
MACQ-Family	-0.176	0.209	-0.233	0.365	0.145	0.150
MACQ-Group	**0.436**	**0.001**	0.179	0.354	0.030	0.779
MACQ-Fair	-0.178	0.157	0.079	0.667	-0.130	0.154
MES	0.125	0.230	0.136	0.328	0.109	0.165
IB	0.090	0.476	0.139	0.331	**0.188**	**0.025**
IH	0.090	0.394	0.104	0.433	-0.080	0.293

Outcome Predictor	**EA** (R^2 = 0.32)		**XA** (R^2 = 0.20)		**Control** (R^2 = 0.27)	
AES	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	0.035	0.786	0.134	0.483	-0.066	0.488
MFQ-Fair	0.111	0.367	0.062	0.746	**0.193**	**0.031**
MFQ-Loyal	-0.076	0.481	-0.183	0.321	**0.221**	**0.012**
MACQ-Family	-0.158	0.232	0.274	0.259	-0.110	0.230
MACQ-Group	0.127	0.299	0.052	0.776	0.126	0.188
MACQ-Fair	0.056	0.637	-0.071	0.682	-0.028	0.737
MES	0.175	0.076	0.211	0.113	0.115	0.107
IB	0.220	0.065	0.035	0.795	**0.283**	**< .001**
IH	0.166	0.096	**0.280**	**0.029**	0.128	0.065

Outcome Predictor	**EA** (R^2 = 0.20)		**XA** (R^2 = 0.39)		**Control** (R^2 = 0.15)	
LBS	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	0.058	0.677	0.182	0.280	0.077	0.451
MFQ-Fair	-0.056	0.672	0.162	0.338	0.063	0.513
MFQ-Loyal	-0.163	0.161	0.217	0.180	-0.009	0.923
MACQ-Family	0.100	0.482	0.309	0.148	-0.068	0.490
MACQ-Group	0.202	0.129	-0.224	0.162	0.199	0.055
MACQ-Fair	0.109	0.395	-0.088	0.566	-0.110	0.221
MES	0.166	0.119	**0.299**	**0.012**	**0.175**	**0.024**
IB	0.121	0.345	0.057	0.628	0.156	0.058
IH	0.097	0.366	**0.258**	**0.022**	0.023	0.763

Outcome Predictor	**EA** (R^2 = 0.39)		**XA** (R^2 = 0.30)		**Control** (R^2 = 0.28)	
ILMS	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	**0.305**	**0.014**	0.048	0.791	0.057	0.542
MFQ-Fair	0.224	0.055	0.239	0.187	0.078	0.379
MFQ-Loyal	**-0.393**	**< .001**	0.070	0.685	0.156	0.071
MACQ-Family	-0.079	0.524	-0.084	0.712	0.010	0.917
MACQ-Group	**0.374**	**0.002**	0.284	0.098	**0.197**	**0.039**
MACQ-Fair	-0.142	0.206	-0.217	0.186	0.049	0.552
MES	-0.046	0.617	0.202	0.104	**0.151**	**0.034**
IB	0.136	0.226	0.127	0.315	0.149	0.050
IH	0.068	0.466	**0.281**	**0.020**	-0.091	0.186

Outcome Predictor	**EA** (R^2 = 0.38)		**XA** (R^2 = 0.32)		**Control** (R^2 = 0.29)	
RFG	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	0.143	0.243	0.040	0.822	0.126	0.180
MFQ-Fair	0.227	0.053	0.260	0.144	0.045	0.609
MFQ-Loyal	**-0.294**	**0.005**	**-0.427**	**0.014**	-0.095	0.268
MACQ-Family	-0.072	0.564	0.189	0.395	-0.132	0.146
MACQ-Group	**0.363**	**0.002**	0.077	0.644	**0.411**	**< .001**
MACQ-Fair	**-0.232**	**0.041**	-0.211	0.190	-0.073	0.374
MES	0.144	0.124	0.204	0.096	0.112	0.112
IB	0.215	0.059	0.222	0.076	**0.204**	**0.007**
IH	0.125	0.185	**0.263**	**0.026**	-0.069	0.312

Outcome Predictor	**EA** (R^2 = 0.41)		**XA** (R^2 = 0.30)		**Control** (R^2 = 0.38)	
EAIS-EX	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	0.156	0.196	0.164	0.360	0.115	0.188
MFQ-Fair	0.061	0.595	0.113	0.529	0.052	0.527
MFQ-Loyal	-0.092	0.355	-0.212	0.218	0.080	0.316
MACQ-Family	-0.233	0.059	0.055	0.808	**-0.192**	**0.024**
MACQ-Group	0.180	0.116	0.300	0.080	**0.257**	**0.004**
MACQ-Fair	0.073	0.508	-0.172	0.293	0.097	0.203
MES	**0.218**	**0.018**	0.157	0.204	0.084	0.200
IB	0.199	0.073	0.222	0.081	**0.339**	**< .001**
IH	0.126	0.174	0.053	0.654	-0.029	0.646

Outcome Predictor	**EA** (R^2 = 0.28)		**XA** (R^2 = 0.20)		**Control** (R^2 = 0.24)	
EAIS-EF	*β*	*p*	*β*	*p*	*β*	*p*

MFQ-Harm	0.021	0.873	-0.276	0.155	-0.161	0.097
MFQ-Fair	0.023	0.853	0.146	0.449	**0.192**	**0.035**
MFQ-Loyal	**-0.218**	**0.049**	-0.245	0.186	**0.218**	**0.015**
MACQ-Family	-0.060	0.657	0.374	0.127	-0.008	0.933
MACQ-Group	0.088	0.480	-0.117	0.521	0.005	0.956
MACQ-Fair	0.001	0.990	0.203	0.248	-0.054	0.524
MES	0.046	0.646	0.103	0.435	-0.030	0.678
IB	**0.277**	**0.024**	0.058	0.665	**0.216**	**0.006**
IH	0.199	0.052	**0.352**	**0.007**	**0.306**	**< .001**

Outcome Predictor	**EA** (R^2 = 0.32)		**XA** (R^2 = 0.28)		**Control** (R^2 = 0.17)	
BDT	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	0.163	0.207	-0.029	0.872	0.080	0.429
MFQ-Fair	0.158	0.198	0.127	0.484	0.172	0.071
MFQ-Loyal	**-0.471**	**< .001**	**-0.486**	**0.007**	-0.167	0.073
MACQ-Family	-0.207	0.117	0.054	0.815	**-0.230**	**0.020**
MACQ-Group	0.201	0.101	0.171	0.321	0.088	0.391
MACQ-Fair	-0.071	0.547	0.061	0.712	0.140	0.114
MES	-0.007	0.940	0.164	0.189	-0.057	0.455
IB	-0.016	0.895	0.202	0.116	0.067	0.409
IH	0.063	0.525	0.179	0.135	-0.013	0.857

Outcome Predictor	**EA** (R^2 = 0.09)		**XA** (R^2 = 0.10)		**Control** (R^2 = 0.15)	
RWCA ($)	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	-0.216	0.149	-0.035	0.861	0.023	0.825
MFQ-Fair	-0.102	0.470	-0.003	0.987	-0.113	0.238
MFQ-Loyal	0.032	0.795	-0.051	0.794	0.135	0.151
MACQ-Family	0.091	0.552	0.056	0.828	-0.001	0.990
MACQ-Group	-0.152	0.283	0.213	0.271	0.179	0.085
MACQ-Fair	0.019	0.891	-0.037	0.841	-0.005	0.953
MES	0.172	0.131	-0.213	0.131	0.100	0.196
IB	0.141	0.304	0.085	0.550	0.147	0.073
IH	-0.153	0.183	-0.171	0.201	-0.114	0.129

Outcome Predictor	**EA** (R^2 = 0.16)		**XA** (R^2 = 0.21)		**Control** (R^2 = 0.12)	
RWCA (Time)	*β*	*p*	*β*	*p*	*β*	*p*
MFQ-Harm	-0.163	0.254	-0.211	0.271	0.025	0.808
MFQ-Fair	0.204	0.135	-0.043	0.822	0.101	0.300
MFQ-Loyal	**0.359**	**0.003**	0.281	0.127	0.091	0.339

MACQ-Family	-0.158	0.279	-0.453	0.064	**0.200**	**0.048**
MACQ-Group	0.165	0.223	**0.431**	**0.020**	0.060	0.567
MACQ-Fair	-0.013	0.918	0.195	0.263	-0.104	0.255
MES	-0.022	0.840	-0.087	0.506	-0.007	0.933
IB	0.111	0.400	-0.134	0.317	0.009	0.913
IH	-0.134	0.221	-0.169	0.180	**-0.157**	**0.040**

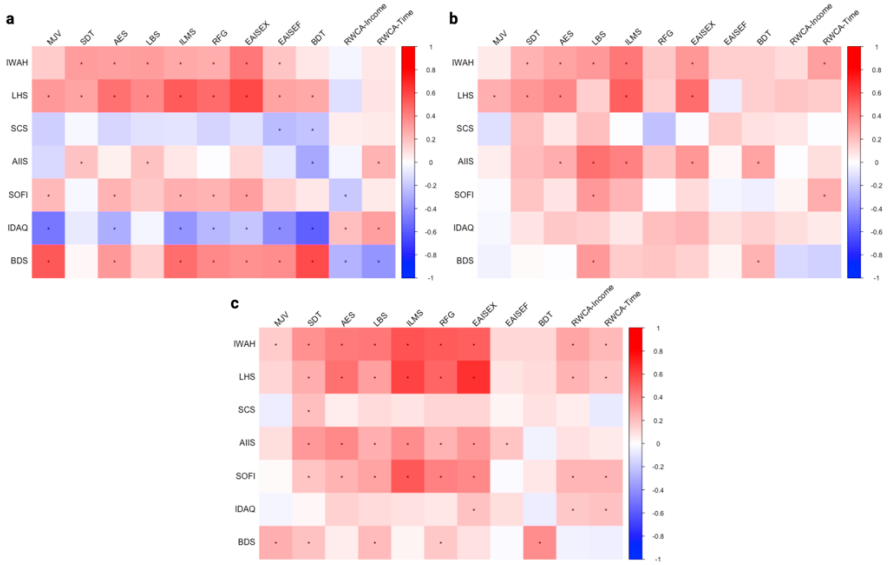
Among XAs, the value of loyalty was negatively associated with equitable and effective prosociality when all of the predictors were considered together, although group loyalty showed a moderate positive relationship with volunteerism among this subject group. Moreover, moral concern on the MES was associated with greater intergenerational concern. Perhaps most intriguing, the value of instrumental harm on the OUS was associated with greater intergenerational concern and expansiveness prioritization among XAs. No prior research has explored relationships between instrumental harm with prosocial outcomes in the population of organ donors, although prior research (Amormino, Ploe, et al., 2022) has found that this value is generally consistent across XAs and general population controls (an effect replicated here). Here, I find that XAs who do manifest this value at a higher level tend to show more equitable altruistic attitudes. These findings suggest tolerance for instrumental harm could in part be responsible for the altruism exhibited by XAs, whose modality of altruism can be construed as self-sacrificial harm for the benefit of others.

For controls, in line with my predictions: (1) valuing fairness was positively associated with moral judgments on the MJV, reported levels of altruistic enkrateia, and the prioritization of effectiveness on the EAIS; (2) moral concern on the MES was positively associated with intergenerational concern; and (3) impartial beneficence on the OUS was positively associated with generosity on the SDT, altruistic enkrateia, responsibility to future generations, and the prioritization of expansiveness and effectiveness on the EAIS. In other words, the basic moral value of fairness and the more specific ascription of moral value to distant others appeared to be

most consistently predictive of equitable and effective prosociality in ordinary adults, suggesting that expanding these values may be a viable avenue to boost prosociality among this population. However, the relationship between loyalty and prosocial behavior among controls was more nuanced, as it was in the bivariate context. Group loyalty was associated positively with altruistic equity and effectiveness, whereas familial loyalty was associated negatively, perhaps limiting the scope of beneficence. This pattern challenges the typical view that loyalty inherently narrows moral regard to closer relational targets (Graham et al., 2017), suggesting that loyalty to one's group can sometimes even predict *greater* prosociality at a distance. These findings underscore complexity in the moral underpinnings of prosocial behavior and highlight the need for further research to better understand these dynamics.

In examining other-inclusive identity and positive other-oriented attitudes (see Figure 15 for bivariate relationships, Table 13 for multiple regression results, and the "Analysis Script" on the OSF for full results), compassionate love for humanity (LHS) among EAs consistently correlated positively with prosocial behaviors, including intergenerational concern (ILMS, RFG), prioritization of effectiveness and expansiveness on the EAIS, and behavioral donations to effective causes. This suggests that EAs who harbor a stronger sense of compassionate love are more inclined to engage in prosocial behaviors aimed at achieving long-term and effective impact. Interestingly, these findings indicate that love for humanity is what distinguishes those within the EA movement who exhibit particularly expansive and effective prosocial mindsets, which contrasts the movement's emphasis on rational, deliberative processes in guiding altruism.

Figure 15. *Bivariate relationships between other-inclusive identity and positive other-oriented attitudes with equitable, effective and overall prosociality among EAs, XAs, and controls*



Note. *Heatmaps displaying correlation coefficients from -1 (blue) to +1 (red) in EAs (**a**), XAs (**b**), and controls (**c**). Asterisks correspond to statistically significant relationships.*

Table 13

Results from multiple linear regression models evaluating associations between individual variation in other-inclusive identity and positive other-oriented attitudes with individual variation in equitable, effective, and real-world prosociality within Samples 1, 2, and 3.

Outcome Predictor	**EA** ($R^2 = 0.40$)		**XA** ($R^2 = 0.10$)		**Control** ($R^2 = 0.10$)	
MJV	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	0.086	0.368	-0.071	0.708	0.074	0.490
LHS	0.183	0.098	**0.394**	**0.027**	0.136	0.254
SCS	-0.115	0.135	-0.122	0.377	-0.121	0.112
AIIS	-0.055	0.582	-0.037	0.832	0.056	0.556
SOFI	0.127	0.144	-0.118	0.451	-0.123	0.198
IDAQ	**-0.263**	**0.020**	-0.010	0.946	-0.057	0.471
BDS	**0.264**	**0.018**	-0.017	0.893	**0.244**	**0.002**
Outcome Predictor	**EA** ($R^2 = 0.19$)		**XA** ($R^2 = 0.14$)		**Control** ($R^2 = 0.21$)	

SDT	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	0.198	0.076	0.010	0.957	**0.229**	**0.025**
LHS	**0.277**	**0.031**	0.301	0.083	-0.052	0.644
SCS	-0.076	0.398	0.157	0.248	**0.175**	**0.016**
AIIS	0.129	0.272	0.026	0.880	**0.250**	**0.005**
SOFI	**-0.240**	**0.019**	-0.031	0.842	-0.009	0.919
IDAQ	-0.219	0.094	0.033	0.812	-0.050	0.505
BDS	-0.145	0.259	0.049	0.697	0.133	0.065

Outcome Predictor	**EA** (R^2 = 0.32)		**XA** (R^2 = 0.19)		**Control** (R^2 = 0.25)	
AES	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	0.128	0.208	0.048	0.790	0.175	0.076
LHS	**0.356**	**0.003**	**0.402**	**0.018**	**0.330**	**0.003**
SCS	**-0.168**	**0.042**	0.046	0.727	0.000	0.996
AIIS	-0.018	0.870	0.027	0.874	0.143	0.098
SOFI	0.059	0.526	-0.164	0.274	-0.104	0.233
IDAQ	-0.224	0.061	0.113	0.409	0.002	0.977
BDS	0.050	0.668	0.018	0.886	-0.010	0.881

Outcome Predictor	**EA** (R^2 = 0.19)		**XA** (R^2 = 0.38)		**Control** (R^2 = 0.23)	
LBS	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	0.141	0.205	0.030	0.850	**0.373**	**< .001**
LHS	**0.273**	**0.034**	-0.185	0.207	-0.096	0.385
SCS	-0.163	0.072	0.074	0.520	0.078	0.273
AIIS	0.075	0.521	**0.438**	**0.004**	0.017	0.844
SOFI	-0.022	0.829	**0.283**	**0.033**	0.147	0.099
IDAQ	0.031	0.813	-0.058	0.628	0.067	0.366
BDS	0.101	0.429	**0.317**	**0.004**	**0.155**	**0.030**

Outcome Predictor	**EA** (R^2 = 0.44)		**XA** (R^2 = 0.36)		**Control** (R^2 = 0.44)	
ILMS	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	0.013	0.887	0.126	0.433	**0.262**	**0.002**
LHS	**0.432**	**< .001**	**0.404**	**0.008**	**0.319**	**< .001**
SCS	-0.099	0.183	-0.129	0.270	-0.010	0.874
AIIS	0.095	0.332	0.167	0.259	-0.030	0.685
SOFI	0.005	0.951	0.019	0.883	**0.236**	**0.002**
IDAQ	**-0.246**	**0.024**	-0.085	0.482	-0.071	0.259
BDS	0.209	0.051	0.189	0.086	-0.071	0.241

Outcome Predictor	**EA** (R^2 = 0.34)		**XA** (R^2 = 0.16)		**Control** (R^2 = 0.34)	
RFG	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	0.062	0.536	0.047	0.798	**0.369**	**< .001**
LHS	**0.445**	**< .001**	0.162	0.339	0.190	0.066
SCS	**-0.165**	**0.042**	**-0.267**	**0.049**	0.049	0.457
AIIS	-0.115	0.276	0.039	0.816	-0.118	0.149
SOFI	0.037	0.682	-0.002	0.988	0.155	0.061
IDAQ	-0.020	0.864	0.158	0.258	-0.016	0.811
BDS	0.199	0.086	0.172	0.172	0.074	0.258

Outcome Predictor	**EA** (R^2 = 0.45)		**XA** (R^2 = 0.30)		**Control** (R^2 = 0.46)	
EAIS-EX	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	**0.202**	**0.029**	-0.014	0.932	0.139	0.097
LHS	**0.462**	**< .001**	**0.510**	**0.002**	**0.648**	**< .001**
SCS	-0.134	0.072	-0.079	0.518	0.016	0.780
AIIS	-0.077	0.426	0.054	0.724	-0.083	0.258
SOFI	0.029	0.728	-0.098	0.476	-0.074	0.315
IDAQ	0.010	0.928	0.165	0.194	0.050	0.415
BDS	**0.215**	**0.043**	0.180	0.118	0.010	0.860

Outcome Predictor	**EA** (R^2 = 0.30)		**XA** (R^2 = 0.10)		**Control** (R^2 = 0.06)	
EAIS-EF	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	0.096	0.350	0.351	0.068	0.079	0.478
LHS	**0.239**	**0.045**	-0.198	0.260	0.011	0.927
SCS	**-0.222**	**0.009**	0.198	0.155	0.048	0.540
AIIS	-0.026	0.813	-0.121	0.488	0.185	0.058
SOFI	0.042	0.652	-0.134	0.393	-0.150	0.128
IDAQ	**-0.351**	**0.004**	0.050	0.728	0.054	0.504
BDS	0.017	0.888	0.004	0.976	-0.019	0.807

Outcome Predictor	**EA** (R^2 = 0.45)		**XA** (R^2 = 0.18)		**Control** (R^2 = 0.15)	
BDT	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	0.042	0.644	-0.101	0.579	0.071	0.500
LHS	**0.318**	**0.003**	0.168	0.317	0.071	0.538
SCS	-0.124	0.090	0.131	0.323	0.032	0.668
AIIS	**-0.232**	**0.016**	0.286	0.091	-0.127	0.167
SOFI	-0.022	0.792	-0.199	0.186	0.033	0.725
IDAQ	**-0.283**	**0.008**	0.035	0.796	-0.007	0.927

BDS	**0.205**	**0.049**	0.243	0.052	**0.339**	**< .001**

Outcome Predictor	**EA** (R^2 = 0.13)		**XA** (R^2 = 0.09)		**Control** (R^2 = 0.13)	
RWCA ($)	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	0.028	0.805	0.084	0.664	**0.272**	**0.011**
LHS	0.098	0.458	0.249	0.160	0.042	0.720
SCS	0.036	0.699	0.074	0.594	0.048	0.526
AIIS	-0.170	0.163	-0.227	0.201	-0.174	0.062
SOFI	-0.199	0.060	-0.109	0.491	0.131	0.166
IDAQ	0.104	0.439	0.158	0.276	0.135	0.085
BDS	-0.255	0.057	-0.112	0.393	-0.088	0.241

Outcome Predictor	**EA** (R^2 = 0.18)		**XA** (R^2 = 0.18)		**Control** (R^2 = 0.13)	
RWCA (Time)	*β*	*p*	*β*	*p*	*β*	*p*
IWAH	-0.071	0.521	**0.392**	**0.035**	**0.243**	**0.023**
LHS	0.188	0.143	-0.120	0.472	-0.030	0.800
SCS	-0.054	0.551	-0.159	0.231	-0.083	0.270
AIIS	0.083	0.479	-0.110	0.513	**-0.198**	**0.035**
SOFI	-0.006	0.953	0.236	0.119	**0.207**	**0.029**
IDAQ	0.006	0.964	-0.015	0.914	**0.163**	**0.038**
BDS	**-0.400**	**0.002**	-0.177	0.155	-0.063	0.401

Additionally, outgroup humanization (BDS) and identification with all humanity were positively associated with prosocial outcomes, implying that recognizing the full humanity of stigmatized others and feeling connected to all people generally may be important ingredients for the prosocial actions of EAs. However, anthropomorphism (IDAQ), or attributing human-like characteristics to non-human entities, often correlated negatively with prosocial priorities, suggesting that for some EAs, prioritizing *animal* welfare might conflict with prioritizing *human* welfare. This finding aligns with previous research indicating that individuals who hold animals in higher moral regard oftentimes exhibit more pronounced biases against outgroup and stigmatized humans, an effect which manifests across most cultures and countries (Rottman et al., 2021; Syropoulos, Crimston, Markowitz, et al., 2024a). Moreover, the present finding related

to anthropomorphism is practically significant in light of the EA movement's tendency to place causes that aim to improve human and animal welfare alike among lists of top charities. Specifically, it's possible that in simultaneously championing humanitarian objectives and animal welfare initiatives, the EA movement may be deterring engagement from would-be newcomers who prioritize one of these goals but not the other.

Consistent with the findings among EAs, compassionate love for humanity (LHS) emerged again as the most consistent positive predictor of prosocial outcomes among XAs. Additionally, positive feelings towards others (SOFI), IWAH and humanization correlated positively with various prosocial measures, including those capturing expansive concern for future generations and real-world volunteerism. These findings suggest that among XAs, a broad and inclusive identity, coupled with love for humanity, may be key drivers of prosocial motivations. Similarly, for controls, compassionate love for humanity (LHS), IWAH, and tendencies to humanize stigmatized groups (BDS) were also linked to increased prosociality. However, interdependence on others (SCS) and allo-inclusive identity (AIIS) showed mixed effects across samples, sometimes predicting positive prosocial outcomes, sometimes negative, and other times showing no significant effect at all. Taken together, these findings emphasize that compassionate love, identification with humanity, and humanization are likely to promote prosocial behaviors across individuals spanning the spectrum of baseline levels of prosocial engagement. Nonetheless, the impact of factors like anthropomorphism and social connectedness is more nuanced and context-dependent, indicating that strategies to enhance prosociality may need to be tailored to specific psychological traits and group dynamics.

CHAPTER 4

DISCUSSION

For decades, social psychologists have sought a deeper understanding of the biases which engender prosocial favoritism for ingroup members, socially close and similar others, and singular identifiable targets (Bruneau et al., 2017; Dovidio et al., 2010; Patil et al., 2021; Sherif et al., 1961). Among the solutions put forth to overcome these biases, downregulating empathic emotion has risen to the forefront of contemporary debates within (Bloom, 2016; Caviola et al., 2021) and beyond (MacAskill, 2016, 2018; Singer, 2015) psychology. Indeed, seminal research revealed empathy's numerous prosocial benefits–experiencing the emotions of suffering victims as one's own, all else equal, makes people more likely to offer aid (Batson et al., 1983; Batson & Shaw, 1991; Coke et al., 1978; Eisenberg & Miller, 1987). But, like prosociality, empathy is often experienced in a biased manner (Amodio & Cikara, 2021; Bruneau et al., 2017; Cikara et al., 2011, 2014; Kteily & Bruneau, 2017; Vollberg et al., 2019). People more easily share in the emotions of ingroup members, relationally close targets, and similar others compared to those who are less-centrally categorized with the self. People also find it easier to share in the emotions of singular identifiable victims over statistical mass suffering (Kogut et al., 2015; Sharma & Morwitz, 2016; Västfjäll et al., 2014). In efforts to debias prosociality to be more equitable and effective, some prominent scholars have suggested combatting altruistic bias by removing empathy altogether from the prosocial calculus, replacing empathic emotion with deliberative reasoning and evidence to promote greater altruistic equality (Caviola et al., 2021; J. Greene, 2014; J. D. Greene et al., 2001, 2004, 2008; Patil et al., 2018).

The argument *against* empathy as a force for good and *for* reasoning as a solution to empathy is no better encapsulated than in the effective altruism (EA) philosophy and social movement, which canonizes empathy as an invariable impediment to high impact altruism and

reasoning as the clear solution (MacAskill, 2016, 2018; Singer, 2015). EAs strive to maximize the impact of their charitable contributions. They ostensibly prioritize reason and evidence in guiding their efforts, focusing on alleviating severe and preventable hardships faced by distant beneficiaries. The central tenet of their approach is to save as many lives as possible with each dollar donated, placing the highest value on achieving the greatest good efficiently. While EA is staunchly aligned with perspectives against empathy, until now, its exceptionally altruistic adherents have been woefully understudied in the current literature. On the other hand, emerging evidence on a separate population of exceptional altruists–non-directed organ donors (also known as extraordinary altruists or XAs)–challenges the perceived tension between empathy and reasoning in promoting equitable and effective altruism. This research suggests that XAs, instead of relying on heightened deliberative reasoning, exhibit enriched empathic capacities, enabling them to more easily share in the emotions of suffering strangers (Amormino, O'Connell, et al., 2022; Brethel-Haurwitz et al., 2018; Law, Amormino, et al., 2024; Marsh et al., 2014; K. M. Vekaria et al., 2020). These findings imply that empathy and reasoning may not be opposing forces as traditionally thought.

Importantly, the populations of EAs and XAs provide valuable test samples to explore whether empathy and reasoning are truly at odds within these populations and within ordinary adults. While EAs advocate for reasoning as the key to overcoming altruistic biases, this perspective lacks substantial empirical support. Conversely, the altruism of XAs appears to be driven by empathy rather than reasoning, though further research is needed to determine if this group also possesses enhanced reasoning abilities. Nonetheless, both groups engage in altruism that transcends the parochial biases typically observed in ordinary individuals. As such, comparing the two special populations to each other and to members of the general population allows for an examination of similarities and differences in the psychological features that

characterize the minds of each group. It also allows insight into the underlying architecture of prosociality–the features that best predict equitable and effective altruism within the minds of exceptional altruists and members of the general population alike.

Here, by bringing EAs, XAs, and demographically similar controls into the laboratory for the first time under a single, unified investigation, I aimed to explore whether there is a singular pathway to altruistic equity and effectiveness characterized solely by rational or empathic processes, or whether instead there are multiple routes to exceptional altruism that individuals could follow independently based on their unique abilities or concurrently to maximize altruistic outcomes. My overarching goal in conducting this research was to help reconcile debates over the roles of empathy and reasoning in altruism, shedding light on whether they are truly opposing or complementary forces in fostering prosocial behavior. Moreover, by identifying the features that predict equitable and effective altruism in each subject group, the present findings can serve as launching point for further research to test causal pathways between the central features identified here and prosocial outcomes in the general population, potentially harnessing these insights and future ones to guide interventions aimed at increasing altruistic equality in the real world for the benefit of society. To address these quandaries, I first compared EAs, XAs, and controls across various metrics, including equitable and effective altruistic attitudes and behaviors, empathic ability, reasoning ability, and other psychological traits known to influence prosocial behaviors. I then assed the extent to which capacities in empathy, reasoning and other prosocial traits associated with equitable and effective altruistic attitudes and behaviors within each subject group.

4.1. Insights on the Affective and Cognitive Architecture of Equitable and Effective Altruism Among EAs, XAs, and Ordinary Adults

Confirming the altruistic exceptionalism demonstrated by both EAs and XAs in the real world, both populations scored higher than controls on the battery of laboratory metrics included to capture equitable altruism (i.e., prosociality towards others expressed with impartial regard for the relational proximity of beneficiaries), effective altruism (i.e., prosociality which prioritizes greater over lesser impact), and altruism more broadly construed. With regard to overall prosociality, without taking into account the prioritization of equity or effectiveness, both special populations reported donating substantially more income to charities and volunteering more of their time in a given year to improve the lives of others. Beyond merely serving as a sanity check to validate the exceptionally altruistic samples, these findings suggest that the altruism displayed by XAs extends beyond singular and profound acts of altruism (e.g., donating an organ to a stranger) and manifests more comprehensively throughout their lives, in their routine donation behaviors and volunteerism.

Even more importantly, both EAs and XAs scored higher on measures that assessed prosociality in the context of real-world tradeoffs between prioritizing effective impact or beneficiary relational closeness, such as the Social Discounting Task (SDT) and the Moral Judgment Vignettes task (MJV). From the perspective of affluent individuals, the most effective impact is often achieved by allocating resources to socially distant beneficiaries suffering from hardships that have been largely eradicated in developed countries but are still prevalent in the developing world and can be addressed at scale for a relatively low cost (80,000 Hours, 2023; GiveWell, 2023; MacAskill, 2016, 2018; Singer, 2015). For example, a thousand dollars has more life-saving potential when used for malaria nets or vitamin A supplements in sub-Saharan

Africa than when allocated to cancer research in the U.S. or the U.K. Consequently, in the natural environments where philanthropic decisions are made, prioritizing altruistic effectiveness often requires prioritizing altruistic equity as well. The observed findings are significant because they challenge the possibility that XAs prioritize equitability but not effectiveness in their altruistic actions, given their focus on organ donation to strangers. This research suggests that, like EAs, XAs likely also prioritize effectiveness in their altruistic efforts.

In a similar vein, it is important to note that EAs and XAs both demonstrated a strong endorsement of equitable prosociality on tasks and measures that specifically evaluated this aspect of altruism in isolation, such as the LBS and the Expansive Altruism subscale of the EAIS. Moreover, membership to either of the special populations attenuated the effects of beneficiary distance on prosociality on the MJV task and the LBS, suggesting that both EAs and XAs are less sensitive to the parochial altruistic inequities that have been commonly observed among members of the general population (e.g., Everett et al., 2018; Fowler et al., 2021; Law et al., 2022; Law, Syropoulos, & Young, 2023; McManus et al., 2020, 2021; Syropoulos, Law, Amormino, et al., 2024a; Syropoulos, Law, & Young, 2024c). While the EA movement typically emphasizes equitable prosociality as a strategy for maximizing effectiveness–since these two features are often intertwined in real-world scenarios–the present findings indicate that EAs, like XAs, also exhibit a genuine inclination towards equitable altruism on its own, even when it does not directly contribute to enhancing altruistic impact.

Understanding that both effective altruists (EAs) and extraordinary altruists (XAs) similarly value and strive for equity and effectiveness broadens our understanding of altruistic motivations. EAs, who are guided by a philosophy that champions deliberative reasoning (Singer, 2015), and XAs, who typically act based on empathy (Marsh, 2019), share common prosocial goals, but potentially achieve them by way of distinct psychological bases. The present

insights into the largely distinct affective and cognitive profiles of EAs and XAs–discussed in greater detail later on in Chapter 4–deepen our comprehension not only of the underpinnings of altruism within these special populations, but of the psychological underpinnings of altruism more broadly. Namely, the present findings suggest that diverse altruistic populations can unite around shared principles, such as maximizing positive impact and extending equal regard to all suffering, irrespective of whether their primary motivations or aptitudes are rooted in empathy or reasoning. This highlights the likelihood for a variety of pathways toward achieving altruistic equity and effectiveness in the general population as well, emphasizing the potential importance of both cognitive *and* emotional capacities in driving prosocial behaviors.

Despite the findings above, on the Effectiveness Focus subscale of the Effective Altruism Interest Scale (EAIS), which measures explicit attitudes towards prioritizing altruistic effectiveness alone, XAs scored lower than both EAs and general population controls. Intriguingly, however, on the behavioral donation task (BDT), XAs and EAs alike elected to donate actual financial resources to a greater number of higher-impact causes (i.e., those with life-saving potential) instead of lower-impact causes (i.e., those that could merely improve lives) relative to ordinary adults. This suggests that, while XAs may not explicitly endorse the principle of effectiveness in their prosocial *attitudes* when it's disambiguated from the principle of altruistic equity, effectiveness prioritization is nonetheless evident in their prosocial *behaviors*. This finding may highlight a distinction between expressed attitudes and actual behaviors among the XA population, underscoring the potential complexity of their altruistic motivations and actions.

Future experimental research could help to elucidate the possibility raised above. For instance, a potential experiment could involve recruiting a larger sample of XAs to either undergo an effectiveness focus intervention or a sham procedure prior to completing the

effectiveness focus subscale of the EAIS and subsequently the BDT. Provided the intervention is successful at boosting *attitudes* in line with effectiveness on the EAIS, Bayesian inference could be used to determine if *behaviors* in line with effectiveness on the BDS remain equivalently high across experimental conditions. If so, these findings would suggest that XAs indeed prioritize effectiveness, not as a result of their explicit attitudes, but perhaps instead because they intuitively are drawn to causes where greater good can be done. Nonetheless, in general, the present findings offer ample support that both EAs *and* XAs, relative to controls, demonstrate remarkably heightened altruistic equity, effectiveness and overall prosociality on most metrics, validating both populations as not only exceptionally altruistic, but largely similar in the specific aspects of prosociality they prioritize.

Perhaps among the most intriguing insights from the present research–hinted at above–can be gleaned from the differences observed in empathic and reasoning abilities across the subject groups. Specifically, the present observations align with previous research conducted on members of the XA population (e.g., Marsh et al., 2019) and the untested sentiments expressed by EA advocates (e.g., Singer, 2015). Namely, XAs scored higher than controls on most measures of empathic ability while EAs scored higher than controls on most measures of reasoning ability. These findings, taken together with those above, provide the first evidence that the psychological aptitudes which distinguish the minds of EAs and XAs from those of average adults are largely distinct between the two subject groups, but that both of the special populations largely align in their exceptional altruistic preferences nonetheless.

Importantly, the effects above were observed when group differences were assessed individually at the level of each trait as well as when the affective, cognitive, and prosocial profiles of EAs, XAs and controls were assessed in a unified model employing profile analysis to account for shared variance among the outcomes. In other words, the present findings support

that there is not merely one monolithic psychological recipe for equitable and effective altruism, but a plurality. Two populations who both share robustly equitable and effective altruistic profiles are characterized by largely divergent affective and cognitive profiles. To the extent that empathy and reasoning support the altruism displayed by XAs and EAs respectively (or even within both populations), the findings suggest that perhaps by cultivating empathy, reasoning, or both among members of the general population, altruistic equity and effectiveness may be expanded.

Notwithstanding the predominant average tendencies described above, some notable nuances within these patterns were observed. For instance, neither population differed from controls in beliefs that empathy is malleable on the Theories of Empathy Scale (TES), in self-reported reasoning ability on the Rational Experiential Index (REI; perhaps because all subjects reported reasoning ability close to ceiling on this measure), or in Actively Open-Minded Thinking (AOT)–the belief in changing one's mind to accommodate new evidence. Furthermore, while XAs scored higher than controls on empathy towards national outgroups on the PES by 3.15 standard deviations, this difference only trended towards significance, perhaps owing to sample size constraints among the XA subject group. Moreover, XAs generally did not differ from controls on most measures of reasoning ability, although they did score higher than controls on the Heuristics and Bias Tasks (HBT), albeit to a lesser extent than EAs. This might suggest that baseline reasoning *abilities* between EAs and XAs may not be *vastly* different. However, more pronounced distinctions could lie in their *preferences* for effortful cognitive tasks (Cacioppo & Petty, 1982) and a greater tendency among EAs to *engage* in deliberative reasoning, especially when overriding intuitive responses in problem-solving contexts (Bialek & Pennycook, 2018; Frederick, 2005; Toplak et al., 2011).

Another noteworthy finding is that while XAs did not differ from controls on the Emotionally Evocative Statements Task (EEST), which measures the ability to accurately identify distinct emotions in written statements, EAs scored lower than both XAs and general population controls. The lack of observed differences between XAs and controls is not entirely surprising, as this metric has not been shown to be extremely diagnostic of *exceptional* empathic ability, as most healthy adults perform remarkably well on this task (Cardinale et al., 2018; Marsh & Cardinale, 2012). Yet, *low* performance on this task has been quite diagnostic of *impaired* empathic ability in prior research among psychopaths and individuals high in callous unemotional traits (CUTs; Marsh & Cardinale, 2012), suggesting that EAs may not merely be less empathic than XAs, but perhaps demonstrate certain limitations in their affective abilities relative to the average adult.

Future research is warranted to replicate the findings above and to further explore the extent of possible affective limitations among the EA population. However, if these findings are confirmed in future studies, they may help to explain why members of the EA community prioritize and advocate reasoning over empathy as a means for doing good. Specifically, if EAs generally have lower affective capacities, they might use their own experiences with empathy to assess the empathic abilities of others, leading them to conclude that empathy is not the most effective mechanism for promoting welfare. Indeed, considerable inquiry into Social Projection Theory suggests people often impute their own levels of knowledge, skills, and abilities onto others, which can influence their social judgments and expectations (Nickerson, 1999; Robbins & Krueger, 2005).

In sum, despite the nuances detailed above, when collectively taking into account the observed differences between the affective and cognitive profiles of EAs and XAs relative to controls, EAs generally possess heightened capacities in reasoning while XAs generally possess

heightened capacities in empathy. But how do capacities in reasoning and empathy associate with equitable and effective prosociality within each of the populations? In line with the findings above, the present findings indicate that both empathy *and* reasoning may be essential in fostering equitable and effective prosocial behavior among EAs, XAs, *and* ordinary adults. Namely, when assessed in the bivariate context and in the context of multiple regression, measures of empathic and reasoning ability were consistently positively related to equitable and effective prosociality among all three populations.

To start, these observations align with earlier research on altruistic kidney donors, which highlighted empathy as a key factor underlying their prosocial tendencies (Amormino, O'Connell, et al., 2022; Brethel-Haurwitz et al., 2018; Law, Amormino, et al., 2024; Marsh et al., 2014; K. M. Vekaria et al., 2020). The present findings extend this understanding by demonstrating that empathy among donors also predicts a prioritization of effectiveness in altruistic attitudes and behaviors. That is, the current investigation broadens the scope of previous ones, which primarily focused on altruism toward strangers and outgroups, by showing that an emphasis on maximizing impact is also associated with empathic abilities among XAs. Moreover, previous research on XAs has predominantly focused on the role of affect, particularly empathy, in guiding their prosocial behavior, often neglecting the potential influence of cognition and deliberative reasoning. However, the current findings suggest that reasoning may also play a crucial role in promoting altruistic equity and effectiveness among XAs, who already exhibit strong empathic abilities. Specifically, a greater capacity for reasoning is associated with higher levels of prosociality in XAs, beyond what can be explained by empathy alone. This indicates that both cognitive and affective processes are important in understanding the motivations behind the altruistic behaviors of organ donors.

Especially noteworthy is that associations between affective and cognitive capacities with prosociality, even for empathic predictors, were particularly strong among EAs. So, while the findings do support the perspective that reasoning predicts greater altruistic equity and effectiveness (Caviola et al., 2021; J. Greene, 2014; Huang et al., 2020; Lucius Caviola et al., 2022; Patil et al., 2021; Schubert & Caviola, 2024)–a perspective especially prominent among the EA community (Greaves & MacAskill, 2019; MacAskill, 2016, 2018; Singer, 2015, 2016)–they *do not* support that empathy and reasoning are at odds, nor that reasoning better-suited for guiding resource allocations to promote the greater good. Rather, the results largely suggest that empathy, even and perhaps especially among EAs, predicts greater altruism in line with the greater good, not less.

Although both empathy and reasoning were predictive of greater altruistic equity and effectiveness across subject groups, in some instances, deviations from this general tendency were observed. For one, when assessed in the context of multiple regression, the ability to correctly identify emotions in statements on the EEST associated negatively with generosity on the Social Discounting Task (SDT) among EAs, and with altruistic enkrateia (AES) and the prioritization of altruistic equity (i.e., expansiveness on the EAIS) among controls. While it's possible that this particular emotional capacity may in some ways limit certain aspects of prosociality, prior research has shown low scoring on this task to be predictive of *antisocial* traits, attitudes and behaviors (Cardinale et al., 2018; Marsh & Cardinale, 2012). Moreover, on other prosocial outcomes among EAs (e.g., behavioral donations to high-impact causes on the BDT, motivations to leave behind a positive legacy marked by prosocial impact on the ILMS), higher scoring on the EEST was associated with greater prosociality. As such, further empirical inquiry into the nuances observed with respect to the EEST are needed to first replicate and then better elucidate these findings.

Another intriguing and unexpected pattern emerged among XA subjects. Namely, when assessed in the context of multiple regression, greater outgroup empathy on the PES associated negatively with donors' sense of responsibility to future generations (RFG). This pattern is intriguing for two reasons. First, diverging with the pattern observed in XAs, outgroup empathy was associated positively with RFG among EAs and controls. Second, much of my prior research conducted among highly-powered samples of general population subjects dovetails with the present findings among the EA and control cohorts. That is, people who feel greater concern for the future of humanity often also tend to feel more expansive concern for humanity in the present, spanning targets from family members and friends, to outgroup members and stigmatized humans (Soutschek et al., 2016; Tamir & Mitchell, 2011; Tuen et al., 2023). Intergenerational concern and concern for others in the present are thought to hang together because a common cognitive and neural foundation is engaged when individuals transcend both the present (e.g., in the context of future thinking) and spans of social distance (e.g., in the context of perspective taking) in the mind's eye. Nonetheless, the present findings suggest that XAs may show a departure from these well-established tendencies, wherein, for organ donors in particular, feeling a greater sense of empathy for present-day outgroups may entail feeling lesser duty to protect those to come. Further research among this population is needed to replicate this finding and further explore its potential underpinnings, including focusing attention on whether the coupling between mental representations of temporal and social distances are less pronounced among the population of non-directed organ donors.

A final intriguing and unexpected pattern among associations with empathic capacities was observed in multiple regression with respect to general population controls. Specifically, members of the general population who reported viewing empathy as more malleable (on the TES) also tended to report a lower prioritization of effectiveness in altruism. This is intriguing in

light of prior research which has found that individuals with more flexible views on empathy tend to invest more empathic effort toward distant and stigmatized groups (Schumann et al., 2014). On the basis of this prior research, I had hypothesized a positive relationship between expansive lay theories of empathy and a focus on effective altruism. As noted in Chapter 3, this particular finding could potentially hint at an "empathic bystander effect" (Darley & Latane, 1968), where individuals believe that if empathy can be cultivated, others will take responsibility for addressing widespread suffering. Consequently, those who see empathy as malleable might overestimate others' willingness or capacity to prioritize effectiveness, thus diffusing their own sense of responsibility in contributing to these causes. Further research could investigate this hypothesis to better understand its impact on altruistic decision-making.

Some deviations from the general tendencies were observed among relationships between cognitive capacities with prosocial outcomes as well. Namely, tendencies to engage deliberative reasoning to correctly solve tricky word problems on the Cognitive Reflection Test (CRT) were negatively associated with RFG among EAs. Similarly, Actively Open-Minded Thinking (AOT), which refers to tendencies to change one's mind in the face of new evidence, negatively predicted real-world charitable donations and volunteerism among EAs and controls alike. Thus, in certain situations, reasoning abilities may have unintended consequences, potentially by causing individuals to focus more on the opportunity costs associated with giving, particularly when it involves equitable distributions. This effect can occur even among those who usually prioritize effectiveness in their altruistic actions.

It should be noted, however, that the unexpected findings discussed above were deviations from the general trends observed across subject groups, predictors, and outcomes. Generally speaking, where significant effects were observed, measures of empathic ability *and* reasoning ability were most often positively associated with measures of equitable and effective

prosociality. Thus, when taken collectively, it's likely that if either capacity is to be downregulated in the prosocial context, as has been suggested with respect to empathy in recent discourse (Bloom, 2016; Caviola et al., 2021; Prinz, 2011; Schubert & Caviola, 2024; Singer, 2016), doing so may significantly reduce levels of beneficence within already altruistic groups and the rest of us. This highlights the need to cultivate both empathy and reasoning to enhance altruistic behaviors and promote a balanced approach to prosociality. Further insight into how this may best be accomplished stems from observations made when empathy, reasoning, and their interactions were statistically accounted for to elucidate more complex patterns among the prosocial outcomes.

Namely, after reducing the dimensionality of empathic and reasoning ability and collapsing the battery of predictors into two distinct factors, the individual and combined associations of the two capacities were assessed with respect to each of the prosocial outcome metrics. Of note, and in line with the findings above, significant main effects of empathic ability were observed on many of the prosocial outcomes, and these associations were always in the positive direction. Significant main effects of reasoning ability were also observed on many of the prosocial outcomes, and these were also generally positive. Nonetheless, greater reasoning ability predicted lower real-world monetary contributions to charities among EAs and controls, and lower volunteerism among EAs in particular. Furthermore, empathic ability was a stronger predictor of expansive altruism across subject groups on the EAIS–which measures explicit prioritization of altruistic equity–while reasoning ability was a stronger predictor of effectiveness focus on the same scale. Yet, both reasoning and empathic ability yielded significant and positive main effects on behavioral donations to effective causes on the BDT in both of the special populations and tracked positively (though not significantly, warranting further replication with higher power) with behavioral donations among controls.

The findings above suggest that cognitive and emotional capacities may be uniquely suited to drive attitudes related to effectiveness and equity, respectively–both critical objectives of the EA movement. Specifically, empathic ability strongly predicts attitudes favoring altruistic equity, while reasoning ability is more aligned with explicit attitudes prioritizing effectiveness. However, in terms of actual behaviors prioritizing effectiveness, both empathy and reasoning contribute uniquely, indicating that both capacities are likely important for motivating effective altruistic actions. This underscores the need for a balanced development of both cognitive and emotional skills to fully realize the goals of effective altruism and to prioritize altruistic equity and impact more broadly.

Examining the interactions between empathy and reasoning reveals a more nuanced understanding, emphasizing the importance of developing both cognitive and emotional capacities. Specifically, the impact of empathic ability was more pronounced among EAs on the Behavioral Donation Task (BDT) and among XAs on the Moral Judgment Vignettes (MJV) task when reasoning ability was lower. This suggests that for exceptionally altruistic individuals, strengths in either reasoning or empathy can compensate for weaknesses in the other, helping to maintain altruistic equity and effectiveness even when one of these aptitudes is constrained. To maximize their positive impact, EAs and XAs could benefit from identifying their cognitive and affective strengths and weaknesses, and then working to enhance the area where they are less developed. This personalized approach to capacity building, though effort-intensive, could be particularly effective for members of these altruistic populations who are already highly motivated to contribute to global welfare, and thus may be poised to engage in such an undertaking.

However, members of the general population vastly outnumber the populations of EAs and XAs combined, and thus hold greater power over increasing altruistic equity and impact on a

global scale. For these ordinary adults, a more intriguing interaction was observed with respect to their attitudes in line with equity and effectiveness on the EAIS as well as their endorsement of altruistic enkrateia on the AES, which assesses their propensity to exert willpower to prioritize both values in altruism. Namely, the full benefits of empathy and reasoning for promoting altruistic equity and effectiveness were observed only when both capacities were strong. This suggests that the interaction between empathy and reasoning is crucial. Having high levels of both is necessary to maximize their positive associations with altruistic behaviors. These findings highlight the critical importance of developing both cognitive and emotional capacities to achieve the most significant prosocial outcomes among members of the general population. Intriguingly, this pattern was especially pronounced for attitudes in line with effectiveness in particular, where relationships with empathic ability and reasoning ability were negative at lower levels of the alternative capacity and positive at higher levels, with the crossover point being just above the average for reasoning ability and just below the average for empathic ability. Thus, it's possible that developing one capacity without developing the other may in fact have detrimental effects on levels of prosocial engagement among the vast majority people.

The findings above not only hold practical importance for how the goals of equity and effectiveness in altruism may be best promoted in the real-world, but also theoretical importance for our understanding of the relationship between empathy and reasoning in the context of prosocial behavior. Despite substantial evidence from seminal literature supporting that empathy is a force for good (Batson et al., 1983, 1997; Batson & Ahmad, 2009), one that can cultivate prosocial engagement across group boundaries, much of the recent theoretical discourse in psychology has taken to attributing the biases which are often observed in empathy and prosociality to empathy itself (Bloom, 2016; J. Greene, 2014; Prinz, 2011; Schubert & Caviola, 2024), suggesting that reasoning is a cure for the limitations of empathy. The present findings

challenge the perspective above and any that suggest prioritizing either empathy or reasoning alone is sufficient for fostering prosocial behavior. They indicate that both capacities are essential and that downregulating either one likely reduces prosocial engagement all else equal. Moreover, the findings suggest that prevailing views that favor reasoning over empathy, particularly as a driver of equity and effectiveness, may be based on a false dichotomy.

Instead of being an inherent aspect of empathy, parochial biases likely reflect broader patterns in people's attitudes that shape the specific groups for whom they feel greater empathy. This suggests that these biases can influence empathetic responses and might even play a role in reinforcing the development of such biases over time, creating a reciprocal relationship. Indeed, prior research has shown that people with more pronounced parochially biased attitudes also tend to exhibit more pronounced parochial biases in empathy (Bruneau et al., 2017). By the same logic, parochial biases may also cultivate downstream biases in reasoning, causing individuals to favor altruistic actions toward beneficiaries who are more closely aligned with their own identity. Extensive research on self-serving and close-other serving biases demonstrates how parochial biases influence the way people understand and interact with the world through deliberative reasoning and cognition. This includes findings on nationalism, self-serving bias, confirmation bias, and the ultimate attribution error, which all illustrate how parochiality can skew reasoning processes and decision-making, often reinforcing preferential treatment toward those perceived as similar or connected to oneself (Hewstone, 1990; Hutchinson & Smith, 1994; Klayman, 1995; Pettigrew, 1979; Sedikides et al., 1998; Snyder & Swann, 1978). These perspectives have been notably absent from the literature advancing the narrative against empathy but warrant further attention in developing a clearer picture of the nuances among the relationships between affect, cognition and altruism.

Namely, further experimental research is necessary to understand how biased attitudes might causally influence empathy and reasoning processes (and vice-versa) and how these, in turn, affect equitable and effective prosocial behavior. To explore relationships among these variables cross-sectionally, I examined the associations between inclusive identity traits and positive attitudes towards others–such as Identification with All Humanity (IWAH), Compassionate Love for Humanity (LHS), and tendencies to humanize outgroup members (BDS)–with empathic and reasoning abilities. The analyses revealed that higher capacities in empathy and reasoning are positively associated with less biased attitudes, particularly those most predictive of altruistic equity and effectiveness[7]. This suggests that parochial biases are linked to cognitive and affective capacities known to be associated with prosociality across exceptionally altruistic and general population cohorts. Moreover, greater empathic ability (and reasoning ability) is associated with less parochial bias rather than more. Future research should manipulate these facets to better understand the directionality of these relationships and how they impact prosocial behaviors. Such efforts going forward can help to build out our knowledge of how capacities in affect and cognition can be cultivated to reduce biases and how biases may be reduced to reciprocally enhance empathy and prosocial reasoning.

Nonetheless, what is clear from the present investigation is that the interplay between empathy and reasoning is critical, and that both capacities should be cultivated to maximize altruistic outcomes. Prior research has shown that both processes, and especially empathy, weigh heavily into people's altruistic decisions (Berman et al., 2018). What I find here is that *both*–not one or the other–not only factor into altruistic attitudes and actions but are important features

[7] These exploratory analyses revealed positive correlations between IWAH with empathy ($r = 0.34$, $p < .001$) and reasoning ($r = 0.13$, $p = .017$), LHS with empathy ($r = 0.52$, $p < .001$) and reasoning ($r = 0.19$, $p < .001$), and humanization on the BDS with empathy ($r = 0.33$, $p < .001$) and reasoning ($r = 0.43$, $p < .001$) across all three subject groups taken together. They suggest that *both* empathy *and* reasoning are associated with parochial biases such that that greater empathy and reasoning are associated with less bias, not more.

within the architecture of prosociality that often work together to yield the most robust impacts on altruistic equity and effectiveness. These effects are observed among EAs who explicitly prioritize reasoning (e.g., Singer, 2015), XAs who have primarily been studied in the context of their heightened empathic capacities (e.g., Marsh, 2019), and, importantly, members of the general population who make up the majority, and thus hold the greatest power to reshape prosociality for the sake of improving societal welfare. Therefore, efforts going forward that look to maximize equality in the distribution of resources or to ensure that resources are allocated towards the most impactful causes ought to develop strategies that harness cognition and affect together.

4.2. The Roles of Morality, Other-Inclusive Identity and Positive Other-Oriented Attitudes in Equitable and Effective Altruism

Beyond examining empathy and reasoning, this research also marks the first unified exploration of differences in moral beliefs and values, other-inclusive identity, and other-oriented attitudes among members of the effective altruism (EA) community, extraordinary altruists (XAs), and the general population. Namely, I explored how these groups differ in their moral frameworks and social attitudes, which have been shown to be crucial for shaping prosocial behaviors and altruistic actions among ordinary adults in past literature (Cikara et al., 2014; Crimston et al., 2016; Graham et al., 2017; Kteily & Bruneau, 2017; Syropoulos, Crimston, Markowitz, et al., 2024a; Waytz et al., 2019). Moreover, I explored how each of these values and attitudes associate with equitable and effective altruism in each of the three subject groups to chart their predictive power across each domain of prosociality. This special population approach is important because it helps to refine theories of prosocial behavior by providing empirical evidence of the specific moral values, beliefs, and social attitudes are that are amplified among and predict prosociality *within* populations of exceptional altruists. But the

broader power of this approach is that by identifying elevated values, beliefs and attitudes and their associations with prosociality in the special populations, a better understanding of those that support and show promise for amplifying prosociality among the general population can be developed.

Approaches like these have a long history of utility in psychological science. For instance, considerable insight has been gained on the cognitive and neural systems that support basic functions in average adults by studying the minds and brains of patient populations with pronounced *deficits* relative to normal levels of functioning. In this vein, much of our early understanding of memory was made possible by studying amnesiacs (Scoville & Milner, 1957). Likewise, pioneering insights on mental imagery is currently being developed through studying individuals with aphantasia–a condition marked by severe deficits in constructing visual representations in the mind's eye (Dawes et al., 2020). Alternatively, considerable knowledge has been gained from studying the minds of experts–those who display marked *enhancements* in various domains relative to normal levels of functioning. In this vein, seminal knowledge on spatial reasoning was developed through studying taxi drivers who show expert levels of ability in navigation (Maguire et al., 2000, 2016) while emerging insights on imagination are coming to light from studying the minds of artists–experts in divergent creativity (Meyer et al., 2019; Orwig et al., 2023).

Harnessing this methodological framework, my current approach involved studying a broad array of features that characterize the minds of two populations marked by largely distinct psychological profiles but who are aligned in their expertise in the domain of prosocial behavior. This approach allowed me to explore whether the same moral values (e.g., fairness, moral concern, impartial beneficence) and social attitudes (e.g., a sense of other-inclusive identity, tendencies to humanize and feel positive emotions towards others) that predict prosocial

behavior in ordinary adults are simply heightened in special populations like EAs and XAs, or whether instead these groups are characterized by distinct values and attitudes that uniquely predict their prosocial behavior. This distinction is crucial for understanding whether prosociality is driven by common factors across all groups or if different mechanisms are at play in these exceptionally altruistic populations that would otherwise go unnoticed when focusing on the general population alone.

I observed that certain moral values that have known associations with prosociality and equitability among ordinary adults were elevated among members of the special populations. Specifically, greater moral concern on the MES and endorsements of the prosocial utilitarian value of impartial beneficence (IB) on the OUS manifested at heightened levels in both EAs and XAs relative to controls. Prior research has shown that these values in the general population support prosociality when broadly construed (Crimston et al., 2016, 2018; Earp et al., 2023; Graham et al., 2017) and equitable and effective prosociality in particular (Everett, 2018; Law, Syropoulos, Coleman, et al., 2023; McManus et al., 2020; Syropoulos, Law, & Young, 2024d). Similarly, the moral value of loyalty on the MFQ–which is known to predict parochial biases and altruistic inequality (Graham et al., 2017)–was reduced in EAs and XAs relative to controls. Nonetheless, EAs and XAs showed unexpectedly lower endorsements of the moral value of fairness relative to controls and no differences on the moral value of harm, both of which have been shown to predict greater levels of care, compassion and equity among ordinary adults (Graham et al., 2011).

As discussed in Chapter 3, the findings related to harm are not entirely surprising, as most people score close to ceiling on this moral value, reasoning that it is morally wrong to intentionally harm others and morally right to alleviate their suffering (Graham et al., 2011). Moreover, the unexpected findings related to the moral value of fairness may stem from how

fairness was measured in the present research and in much of the existing literature. Both the Moral Foundations Questionnaire (MFQ; Graham et al., 2011) and the Morality as Cooperation Questionnaire (MAC-Q; Curry et al., 2019), despite being widely used, conflate two distinct dimensions of fairness: equitability and proportionality. Equitability refers to the principle that all individuals should receive an equivalent share of resources, while proportionality suggests that resources should be distributed based on merit or personal contributions. Emerging research by Atari et al. (2023) highlights that these two dimensions are markedly distinct and endorsed by different individuals across countries and cultures. Given that both EAs and XAs exhibit strong tendencies to allocate resources equitably, this raises the question of whether this conflation among the MFQ and MAC-Q is responsible for the unexpected group differences with respect the moral value of fairness on both scales. To better elucidate this possibility, further research is necessary using more sensitive metrics (e.g., the MFQ-2; Atari et al., 2023) that can tease apart equitability from proportionality.

Perhaps an even more interesting deviation from the predicted pattern of results pertains to the differences observed in the moral values of group loyalty and familial loyalty, as measured by the MAC-Q, between the special populations and controls. As mentioned above, both EAs and XAs scored lower than the general population on the moral foundation of loyalty as assessed by the MFQ. However, while EAs also scored lower than the general population on familial loyalty (MAC-Q), XAs did not differ from the general population on this measure. Furthermore, XAs scored even higher than the general population on group loyalty (MAC-Q), while EAs did not differ from the general population. The nuance in these patterns is especially intriguing given that EAs are bounded by a common ideology and often work together and share personal relationships with other EAs. There are EA communities that span the globe (see https://forum.effectivealtruism.org/groups for a map displaying all EA communities worldwide),

and EAs routinely interface with each other on forums, slack channels and social networks (*Groups — EA Forum*, 2024). Thus, one might predict that if either of the two special populations were to show greater group loyalty, it would be EAs, where loyalty to one's group may be synonymous with loyalty to the EA movement itself. Indeed, perhaps the tight-knit nature of the EA community is to account for why members of this population showed no difference in group loyalty relative to controls.

On the other hand, XAs are not bound by a common ideology, and thus loyalty to group members may not be as closely tied to a specific set of beliefs or causes, but rather other forms of social connection or personal bonds. Intriguingly, emerging research suggests that, despite engaging in sacrifice for strangers, altruistic organ donors maintain close bonds with family members and friends (Amormino et al., 2024), suggesting that close bonds may not inherently pose much of a barrier to equitable altruism (and vice versa), at least among the population of organ donors. Yet, EAs do show diminished tendencies to favor family members, while XAs do not. Thus, it's possible that while XAs do not suffer impoverished social relationships with close others and diminished signatures of close bonding, EAs may.

In support of the possibility raised above, members of the general population rate individuals who engage in welfare-maximizing altruism toward socially distant beneficiaries as morally worse and less trustworthy (Fowler et al., 2021; Law et al., 2022; Law, Syropoulos, Young, et al., 2024). Going forward, a new line of inquiry is needed to better elucidate the social consequences of membership to the effective altruism movement, as well as the social consequences that could come along with promoting the variety of cause-prioritization advocated within EA in the broader population. Research in this vein can look to address whether any social repercussions identified are endogenous to the psychology of individuals drawn to the movement, or whether they originate instead as an unintended consequence of engaging in

equitable and effective prosociality, by eroding the levels of trust that are often considered obligatory in close relationships (McManus et al., 2020). Ultimately, this research can help bridge the gap between descriptive theories of altruism (explaining what is) and normative theories (suggesting what ought to be) by providing actionable insights into how society might cultivate greater prosociality without incurring social costs to altruists themselves.

When assessing associations between moral values with the measures of prosociality across the subject groups, a few noteworthy patterns emerged. For one, prosocial moral beliefs and values, such as moral concern and impartial beneficence, predicted greater equitable and effective prosociality across the three samples. Across groups, the moral value of harm on the MFQ (i.e., harm reduction or care) also predicted greater prosociality on measures capturing (1) altruistic equity and (2) altruistic equity and effectiveness combined, but not those capturing effectiveness in isolation.

Yet, the value of fairness was a stronger predictor of equitable and effective prosociality in the general population compared to the two special populations, EAs and XAs. Related to the earlier discussion of fairness, this difference may be due to how fairness is measured, particularly the use of double-barreled items that do not clearly differentiate between equitability and proportionality (Atari et al., 2023). EAs and XAs, who strongly prioritize equitable resource distributions, might be more cautious about endorsing fairness when it is not clearly defined, leading to lower predictive strength among these subjects. In contrast, members of the general population, who the present findings suggest value equitable prosociality to a lesser degree than EAs and XAs, could be less sensitive to these nuances, resulting in greater predictive strength for fairness among controls. Nonetheless, this interpretation is speculative–further research employing a more sensitive measure of fairness is necessary to delineate whether differential effects are observed among its two dimensions.

The findings regarding loyalty were particularly intriguing. For EAs and XAs, loyalty (broadly defined, as measured by the MFQ) was negatively associated with measures of equitable and effective prosociality where significant effects were noted. In contrast, among the general population, loyalty was a *positive* predictor of various prosocial outcomes, including the prioritization of effectiveness on the EAIS, even after accounting for shared variance in multiple regression analyses. Furthermore, across all three groups, group loyalty was positively associated with several measures of altruistic equity and effectiveness. However, familial loyalty generally showed negative associations with these outcomes. This distinction between group and familial loyalty suggests that while a sense of loyalty to larger groups can enhance prosocial behaviors focused on equity and effectiveness, strong familial loyalty might detract from broader altruistic commitments.

The findings above diverge considerably from recent theoretical models of prosocial behavior which suggest that loyalty to one's group is invariably a contractionary force on people's levels of care and concern for distant others (e.g., the centripetal and centrfugal model of the moral circle; Graham et al., 2017; Waytz et al., 2019). Yet, no investigation prior has examined how this value tracks with such a broad array of prosocial outcomes that are uniquely equipped to capture the prosocial facets of equity and effectiveness. Here, I find that among exceptionally altruistic and ordinary individuals alike, loyalty can be predictive of less rather than more parochial bias in altruism. In other words, loyalty to those who are close does not inherently come at the expense of concern for those who are distant.

While this finding diverges with recent psychological theories of prosociality, it dovetails with more seminal conceptual frameworks of prosociality and moral concern from philosophy. Namely, the expanding moral circle model, originally advanced by Lecky (1869) and later built upon and popularized by Singer (1981), contends that concern for others begins with the self and

extends outward to include more distant others over time–both within individuals over the course of their lifespans and across people throughout history. By this logic, people's circles of regard start small with those closest to them, like family and friends, before growing to include ingroups, acquaintances, outgroup members, and eventually, all living beings within the natural environment. While experimental and longitudinal studies are needed to elucidate whether this model is truly supported within people across time, the present findings suggest that the same psychological features which predict close bonds within our groups can be predictive of more expansive attitudes towards more distant and dissimilar others. They suggest that concern can start close and radiate outward. Research going forward can investigate whether other features which support close relationships (e.g., attachment, trust) may also support concern at greater psychological distances. Such efforts could reveal avenues to broaden the scope of prosociality without risking social repercussions.

Other intriguing insights related to morality were observed in the present research which are elaborated on in greater depth in Chapter 3. One in particular warrants further elaboration here. Namely, EAs scored significantly higher than the other two subject groups on the instrumental harm (IH) subscale of the OUS, which captures endorsement of the utilitarian principle which posits that harm befalling the few is at times necessary to promote welfare among the many (Kahane et al., 2018). This was observed despite foundational texts within EA philosophy staunchly distinguishing the prescriptive ethics of EA from IH and aligning them more closely with the prosocial utilitarian principle of IB (Singer, 2015). This finding may have broader societal implications in light of concerns raised among popular discourse that the EA movement–in part owing to a high-profile financial scandal within the EA community (Cohen & Godoy, 2024)–may at times breed unethical behaviors by justifying them in the minds of transgressors through the lens of promoting instrumental harm.

I find here that EAs do endorse the value of IH by roughly a standard deviation more than controls and organ donors. Further research is needed to investigate the extent to which promoting the prosocial principles of EA may also cultivate values in line with IH harm as an unintended consequence. Such research can also help bridge descriptive and normative theories of prosociality by investigating whether this potential unintended consequence of cultivating equitable and effective altruism is aligned or misaligned with folk moral intuitions regarding what is right and wrong. That is, beyond identifying whether propensities to support instrumental harm may come along promoting the prosocial ideals of EA, future research should investigate whether the average adult sees this as a problem. This can help to inform whether and how efforts to promote altruism are implemented in the real world.

To summarize, I find that certain beliefs and values which are known to promote prosociality among the general population, such as moral concern and impartial beneficence, manifest at heightened levels among exceptionally altruistic groups and predict equitable and effective prosociality across populations. Thus, it stands to reason that by cultivating these values among the general population, altruistic equity and effectiveness may be expanded. Yet, exceptional altruists also manifested heightened levels of group loyalty, which predicted equitable and effective altruism across the populations. This provides initial support that this value, which has been traditionally associated with parochiality in current theories of prosocial behavior (e.g., Graham et al., 2017), may instead promote more expansive concern for the welfare of others and, in turn, greater altruistic equity and effectiveness. Thus, further research is warranted to clarify the complex role of loyalty in the prosocial context to help revise our current understanding of its potential for altruistic utility.

Meanwhile, fairness, which has been associated with equitability and prosociality in previous research and theory (e.g., Graham et al., 2011), was diminished among members of the

special populations, but predicted greater prosociality among controls. These findings ought to be replicated with more sensitive metrics that can distinguish between the various facets that comprise this value (i.e., equitability and proportionality; see Atari et al., 2023). Finally, diminished levels of familial loyalty among EAs and the findings related to instrumental harm warrant further exploration to help bridge descriptive and normative theories of prosociality. Such research can help us navigate working towards societal goals of promoting equality, fairness, and impact in ways that limit social consequences and best align with commonplace moral beliefs regarding right and wrong.

With respect to differences in social attitudes, comprising other-inclusive identity and positive other oriented attitudes, EAs and XAs felt a stronger sense of identity with all humans everywhere on the IWAH, greater compassionate love for all humans everywhere on the LHS, and reported feeling a greater number of positive emotions towards others on the SOFI. These findings align with the findings pertaining to empathy, suggesting that social emotions like love and the sense of connection that is fostered by identifying strongly with others, particularly when felt more universally for all, are key features within the psychological landscape of exceptional care. Scores on IWAH and the LHS were also among the most reliable predictors of equitable and effective prosociality across populations, suggesting that targeting these features may present a viable avenue for broadening prosociality among members of the general population. These findings align with earlier ones suggesting that among the general population, other-inclusive identity and love for all humans are predictive of prosociality when broadly construed (e.g., McFarland et al., 2012; Spreecher & Fehr, 2005) and attitudes in line with equitable and effective prosociality in particular (e.g., Law et al., 2022; Tuen et al., 2023). Further efforts down this line of inquiry may seek to experimentally manipulate these attitudes to (1) test causal

pathways and (2) examine whether developing these attitudes can usher enduring effects on prosocial attitudes and real-world action.

Humanization of outgroups on the BDS was also pronounced among XAs in particular, and widely predictive of prosociality across populations, suggesting that attributing capacities of mind to distant and stigmatized others may pose an additional viable avenue for promoting goals of altruistic equity and effectiveness. Yet, EAs also attributed greater capacities of mind to non-humans relative to controls and XAs. And, interestingly, among EAs in particular, greater anthropomorphism predicted reduced prioritization of intergenerational concern as well as equitable and effective altruistic attitudes on the MJV, the effectiveness focus subscale of the EAIS, and reduced behavioral donations to effective charities on the BDT. These findings may be attributable to the fact the effective altruism movement simultaneously advocates for human *and* animal welfare (MacAskill, 2018; Singer, 2015). They also resonate with prior findings which show that concerns for animals and nature are sometimes at odds with concerns for distant humans among members of the general population (Rottman et al., 2021; Syropoulos, Crimston, Markowitz, et al., 2024b). In a similar manner, scores on the AIIS, which capture a sense of identity not only with distant humans, but with members of the natural environment and physical world, at times predicted lower scores on measures capturing prosociality among controls. These findings could be indicative of tension between the goals promoted within the EA movement. Thus, it's possible that by advocating for animal and human welfare under the same umbrella of priorities, the EA movement may deter newcomers from adopting its prosocial principles, particularly for newcomers who endorse humanitarianism or environmentalism, but not both.

In sum, taken together, it seems that the most robust features which distinguish exceptional altruists from controls are also those that most reliably promote equitable and effective altruism across populations. Namely, developing a sense of identity and love that

extends beyond oneself and relationally close others to include all humanity is likely to yield the greatest benefits for prosocial attitudes and actions among ordinary individuals. Nonetheless, boosting identity with non-humans may pose a roadblock to humanitarian goals, warranting further research to investigate how both environmentalism and humanitarianism may be promoted. This is especially important given that both environmental and humanitarian challenges are likely to manifest and require mitigating in humanity's future (MacAskill et al., 2022). The literature on intergenerational beneficence may serve as a good starting point for future research seeking to explore how this balance may be best navigated (see Law, Syropoulos, & Young, 2023; Syropoulos, Law, & Young, 2024e, 2024d).

4.3. Limitations and Additional Future Directions

In addition to generating novel insights into key debates over the roles of affect and cognition in the context of promoting altruistic equity and effectiveness (e.g., Bloom 2016) and deepening theoretical and practical knowledge pertaining to descriptive and normative theories of prosocial behavior (e.g., Graham et al., 2017), morality (e.g., Crimston et al., 2016; Singer, 1981) and social attitudes (e.g., McFarland et al., 2012), this research boasts numerous methodological strengths. For one, it features two hard-to-reach samples comprising members of distinct populations of real-world altruists and a third sample of demographically similar ordinary adults under a single unified investigation–the first of its kind. These samples feature individuals from across the globe, extending far beyond the typical scope of student samples with WEIRD (Western, Educated, Industrialized, Rich, and Democratic) subjects or even the traditional crowd-sourced online samples commonly used in contemporary psychological research (Henrich et al., 2010). This diverse participant base combined with a focus on targeted special populations allows for a more comprehensive and inclusive understanding of the phenomena under investigation.

Furthermore, the research presented here emphasizes my steadfast adherence to open science practices. The sampling procedures, hypotheses, and analysis plan have been pre-registered on As Predicted and the deidentified data, analysis scripts, figure code, and output have been made available on the Open Science Framework (OSF) website. Moreover, the materials and hypotheses are described in detail and appended below, and all hypotheses conducted for exploratory purposes have been labeled accordingly. Finally, in the survey itself, I included a robust array of well-validated and emerging psychometric instruments to capture a scoping basket of psychological phenomena with remarkable breadth and depth, including prosociality (broadly construed, as well as equitable and effective prosociality in particular), empathic ability, reasoning ability, moral beliefs and values, differences in other-inclusive identity, and positive other-oriented attitudes. These tasks and measures included not only self-report methodologies but also behavioral metrics (e.g., donations of actual resources to real-world charities) and assessments of real-world actions (e.g., reported levels of income donated to charitable causes and time spent volunteering). This approach ensures a more comprehensive and accurate evaluation of the participants' altruistic behaviors and attitudes, moving beyond mere self-perceptions to capture concrete actions and decisions.

Despite these methodological strengths and the numerous insights made possible by conducting this research (which are discussed in detail throughout sections *4.1.* and *4.2.* as well as in Chapter 3), there are some limitations and unaddressed questions that warrant mention in order to open up further avenues for future research to address. The fist pertains to limitations of sample size, particularly among the extraordinary altruist (XA) sample of non-directed, living organ donors. Owing to the sheer rarity of members of the XA population–few people elect to donate their organs to others while they're living and even fewer elect to donate them to strangers (Marsh, 2019)–I was only able to recruit 65 XAs to participate in this research.

Although small, this sample size is consistent with much of the research that has been conducted among members of this population to date (Amormino, O'Connell, et al., 2022; Brethel-Haurwitz et al., 2018; Marsh, 2016a; O'Connell et al., 2019; Rhoads et al., 2023; K. Vekaria et al., 2017). Nonetheless, as a result of these constraints on statistical power, there were some instances where effect sizes that were greater than negligible remained non-significant, particularly when examining associations between predictors and prosocial outcomes within the XA sample. In order to adhere to best practices in frequentist statistics, I did not interpret these findings. Going forward, by employing additional means to sample additional members of the organ donor population (e.g., social media outreach, forging connections with hospitals and institutions with access to these populations), further research should be conducted to replicate the present findings among a larger sample of extraordinary altruists.

Relatedly, while I only focused on studying EAs and organ donors, these groups are not representative of all individuals who routinely engage in costly altruism for the benefit of strangers. Heroic rescuers (Lyons, 2006), humanitarian aid workers (Macpherson & Jr, 2021; Sparkman & Hamer, 2020), and people working to solve societal grand challenges in their careers (e.g., poverty, climate change, AI safety; George et al., 2016) also do tremendous good and warrant attention as well. Future research should look to assess the same questions addressed in the present investigation among members of these separate altruistic cohorts to explore whether additional features that support the architecture of prosociality may be brought to light.

Being the first rigorous investigation into the minds of effective altruists and the first to compare this population to extraordinary altruists and controls, I aimed to capture a broad range of psychological features using a large-scale correlational design. However, because of the correlational nature of this research, no causal inferences can be made at present. This investigation does, however, lay the groundwork for future studies to test causal pathways

among variables in a targeted manner. By identifying the most pronounced features among exceptional altruists that predict prosociality across populations, this study provides a foundation for more in-depth experimental research.

In addition to manipulating and measuring the most reliable predictors of equitable and effective prosociality identified in this research (such as empathic concern, outgroup empathy, scores on the CRT, IWAH, compassionate love for humanity, and moral concern), I propose specific experiments in sections *4.1.* and *4.2.* to clarify nuanced aspects of the findings. For instance, do attitudes precede, succeed, or relate reciprocally with biases in empathy and reasoning? Can cultivating loyalty to one's group can serve to broaden concern and prosocial behavior beyond group boundaries (e.g., Singer 1981)? Experimental follow-ups to the present investigation are warranted to provide deeper insights into questions like these and the mechanisms underlying equitable and effective prosocial behavior.

Provided it is a societal goal, further research could also seek to develop interventions that serve to cultivate greater altruistic equity and effectiveness. In my existing research, I have successfully employed philosophical appeals to cultivate greater concern for the welfare of distant future generations, which in turn has fostered downstream consequences on prosocial attitudes, donation behaviors, and proenvironmental engagement (Law, Syropoulos, & Young, 2023; Syropoulos, Law, Kraft-Todd, et al., 2024b; Syropoulos, Law, & Young, 2024e; Syropoulos, Law, Young, et al., 2024f). Similar interventions employed by other researchers have been utilized to garner endorsement of effective altruism principles and donations (Lindauer et al., 2020). Future research in this vein could leverage new technologies and more immersive delivery methods to enhance the strength and durability of interventions. For instance, developing Large Language Models (LLMs) or virtual reality experiences to simulate

conversations with or immersive scenes depicting distant others and their challenges might help individuals cognitively take the perspectives of these distant others and boost affective empathy.

It is important, however, that before interventions like those discussed above are developed and implemented, it is abundantly clear the extent to which they may engender unintended consequences. As discussed in section *4.2.*, the present research leaves open whether engaging in equitable and effective altruism has social repercussions. While recent findings suggest XAs do not suffer improvised relationships with those who are close, research among the general population finds that people view hypothetical effective altruists as less trustworthy than individuals who do lesser good but help those who are socially close (Kahane et al., 2018; Law et al., 2022; Law, Syropoulos, Young, et al., 2024; McManus et al., 2020). Research going forward ought to explore whether effective altruists have robust social lives like XAs, or whether their close bonds with others are constrained. Future research should also explore whether members of the general population update their social perceptions regarding effective altruists after being presented evidence related to the present findings. Namely, if members of the general public are informed that EAs do not value group loyalty to a lesser extent than the average adult, or if they are informed that EAs are not purely driven by a cold and rational calculus, perhaps this could restore their trustworthiness and mitigate negative social perceptions. Indeed, prior research has shown that beneficence driven by emotions is rated as morally better than beneficence driven by reasoning alone (Montealegre et al., 2020).

Similarly, beyond addressing social perceptions, developing interventions that focus on broadening empathy or fostering more expansive prosocial mindsets through enhancing group loyalty could help cultivate altruistic equity and effectiveness in a more benign manner. This approach could bridge the gap between descriptive and normative theories of prosociality, aiming to enhance societal welfare while mitigating potential negative downstream

consequences. For example, it could expand the scope of prosociality without sacrificing obligations to those who are close (see McManus et al., 2020) and build impartial beneficence in a manner that does not attract individuals with callous unemotional traits (see findings from the present investigation related to scores on the EEST among EAs and those related to the utilitarian principle of instrumental harm). This area of research warrants further exploration but holds promise for advancing both theoretical and practical understandings of prosocial behavior.

Finally, while the present research featured subjects from numerous countries and cultures, exploring the present quandaries in larger internationally-representative samples could yield additional benefits. For instance, it's possible that different features underlie equitable and effective altruism among subjects from cultures that put greater emphasis on long-term societal goals or that prioritize collectivism rather than individual achievement (Bearden et al., 2006; Hofstede, 2011). Moreover, there could be value in mapping the architecture of prosociality in well-powered representative samples from affluent countries, as these populations are particularly equipped to meaningfully contribute to global welfare through philanthropy. Research in this vein could yield targeted insights for developing interventions that effectively elicit philanthropic engagement among those with the greatest giving potential. And, by broadening the level of analysis from individuals to nations or societies, we can gain insights into how the prosocial architectures of a nation's citizens relate to its institutions, policies, and levels of human development. This approach can help understand the dynamics between individual prosocial psychological tendencies and broader societal factors, informing strategies to enhance global welfare both bottom-up and top-down.

CHAPTER 5

CONCLUSION

Psychologists and philosophers have long pitted deliberative reasoning against empathy in the quest for equality and effective impact in altruism. The present findings suggest this binary thinking might be misguided. While effective altruists (EAs) advocate for downregulating empathy to maximize impact, extraordinary altruists (XAs)–who perform acts like donating organs to strangers–demonstrate the profound power of empathy. This research reveals that exceptional altruism isn't about choosing between empathy and reasoning; it's about integrating both. EAs exhibit heightened reasoning skills and XAs exhibit heightened affective capacities. Yet, across diverse populations, both reasoning and empathy prove essential for promoting altruistic equity and effectiveness. Challenging the notion that empathy inherently hinders impartial and impactful altruism or that reasoning alone is sufficient, I find that the effects of reasoning and empathy are in fact likely to be most effective when developed together. That is, an active mind is not as useful without a bleeding heart, and a bleeding heart is not as useful without an active mind. Furthermore, expansive moral concern, utilitarian beliefs, compassionate love, and group loyalty emerge as additional critical predictors of altruistic behavior. The interplay of these factors suggests that fostering both cognitive and emotional capacities can enhance altruistic actions. So, altruism isn't a matter of choosing sides in the debate between empathy and reasoning. Exceptional altruism thrives on the collaboration between affect and cognition, coupled with broad moral values and social attitudes. This integrated approach paves the way for more equitable and effective prosocial behaviors, bridging philosophical ideals with real-world impact to hopefully usher a more equitable and thriving society going forward.

REFERENCES

80,000 Hours. (2023). *How to have a meaningful career with a large social impact*. 80,000 Hours. https://80000hours.org/career-guide/

Abrams, J. R., McGaughey, K. J., & Haghighat, H. (2018). Attitudes toward Muslims: A test of the parasocial contact hypothesis and contact theory. *Journal of Intercultural Communication Research*, *47*(4), 276–292. https://doi.org/10.1080/17475759.2018.1443968

Aitken, G. M. (1998). Extinction. *Biology and Philosophy*, *13*(3), 393–411. https://doi.org/10.1023/A:1006567916998

Amodio, D. M., & Cikara, M. (2021). The Social Neuroscience of Prejudice. *Annual Review of Psychology*, *72*(1), 439–469. https://doi.org/10.1146/annurev-psych-010419-050928

Amormino, P., Kagel, A. C., Li, J., & Marsh, A. (2024). *Does impartial altruism incur costs in close relationships?* OSF. https://doi.org/10.31234/osf.io/db8wv

Amormino, P., O'Connell, K., Vekaria, K. M., Robertson, E. L., Meena, L. B., & Marsh, A. A. (2022). Beliefs about humanity, not higher power, predict extraordinary altruism. *Journal of Research in Personality*, *101*, 104313. https://doi.org/10.1016/j.jrp.2022.104313

Amormino, P., Ploe, M. L., & Marsh, A. A. (2022). Moral foundations, values, and judgments in extraordinary altruists. *Scientific Reports*, *12*, 22111. https://doi.org/10.1038/s41598-022-26418-1

Anthony, N., & Emami, M. R. (2018). Asteroid engineering: The state-of-the-art of Near-Earth Asteroids science and technology. *Progress in Aerospace Sciences*, *100*, 1–17. https://doi.org/10.1016/j.paerosci.2018.05.001

Atari, M., Haidt, J., Graham, J., Koleva, S., Stevens, S. T., & Dehghani, M. (2023). Morality beyond the WEIRD: How the nomological network of morality varies across cultures.

Journal of Personality and Social Psychology, *125*(5), 1157–1188. https://doi.org/10.1037/pspp0000470

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. https://doi.org/10.1016/0022-3999(94)90006-X

Balliet, D., Thielmann, I., & Columbus, S. (2019). Situational Affordances for Prosocial Behaviour: On the Interaction between Honesty–Humility and (Perceived) Interdependence. *European Journal of Personality*, *33*(6), 655–673. https://doi.org/10.1002/per.2224

Barasch, A., Levine, E. E., Berman, J. Z., & Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *Journal of Personality and Social Psychology*, *107*(3), 393–413. https://doi.org/10.1037/a0037207

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C., Mersey, B., & Ferrer, E. A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature*, *471*(7336), Article 7336. https://doi.org/10.1038/nature09678

Baron, J., Isler, O., & Yilmaz, O. (2022). *Actively open-minded thinking and the political effects of its absence*. PsyArXiv. https://doi.org/10.31234/osf.io/g5jhp

Batson, C. D., & Ahmad, N. Y. (2009). Using empathy to improve intergroup attitudes and relations. *Social Issues and Policy Review*, *3*(1), 141–177. https://doi.org/10.1111/j.1751-2409.2009.01013.x

Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T., & Birch, K. (1983). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology*, *40*(2), 290–302. https://doi.org/10.1037/0022-3514.40.2.290

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective Taking: Imagining How Another Feels Versus Imaging How You Would Feel. *Personality and Social Psychology Bulletin*, *23*(7), 751–758. https://doi.org/10.1177/0146167297237008

Batson, C. D., Klein, T. R., Highberger, L., & Shaw, L. L. (1995). Immorality from empathy-induced altruism: When compassion and justice conflict. *Journal of Personality and Social Psychology*, *68*(6), 1042–1054. https://doi.org/10.1037/0022-3514.68.6.1042

Batson, C. D., & Powell, A. A. (2003). Altruism and prosocial behavior. In *Handbook of psychology: Personality and social psychology, Vol. 5.* (pp. 463–484). John Wiley & Sons, Inc. https://doi.org/10.1002/0471264385.wei0519

Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry*, *2*(2), 107–122. https://doi.org/10.1207/s15327965pli0202_1

Bayot, M., Vermeulen, N., Kever, A., & Mikolajczak, M. (2020). Mindfulness and Empathy: Differential Effects of Explicit and Implicit Buddhist Teachings. *Mindfulness*, *11*(1), 5–17. https://doi.org/10.1007/s12671-018-0966-4

Bearden, W. O., Money, R. B., & Nevins, J. L. (2006). A measure of long-term orientation: Development and validation. *Journal of the Academy of Marketing Science*, *34*(3), 456–467. https://doi.org/10.1177/0092070306286706

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, *10*(1), 122–142. https://doi.org/10.1006/game.1995.1027

Berman, J. Z., Barasch, A., Levine, E. E., & Small, D. A. (2018). Impediments to Effective Altruism: The Role of Subjective Preferences in Charitable Giving. *Psychological Science*, *29*(5), 834–844. https://doi.org/10.1177/0956797617747648

Bialek, M., & Pennycook, G. (2018). The cognitive reflection test is robust to multiple exposures. *Behavior Research Methods*, *50*(5), 1953–1959. https://doi.org/10.3758/s13428-017-0963-x

Bloom, P. (2016). *Against Empathy: The Case for Rational Compassion: Paul Bloom: 9780062339331: Amazon.com: Books*. https://www.amazon.com/Against-Empathy-Case-Rational-Compassion/dp/0062339338

Bobonich, C., & Destrée, P. (2007). *Akrasia in Greek Philosophy: From Socrates to Plotinus*. BRILL.

Boggio, P. S., Nezlek, J. B., Alfano, M., Azevedo, F., Capraro, V., Cichocka, A., Pärnamets, P., Rego, G. G., Sampaio, W. M., Sjåstad, H., & Van Bavel, J. J. (2023). A time for moral actions: Moral identity, morality-as-cooperation and moral circles predict support of collective action to fight the COVID-19 pandemic in an international sample. *Group Processes & Intergroup Relations*, 13684302231153800. https://doi.org/10.1177/13684302231153800

Bose, N., & Shepardson, D. (2023, May 4). Biden meets Microsoft, Google CEOs on AI dangers. *Reuters*. https://www.reuters.com/technology/white-house-meet-microsoft-google-ceos-ai-dangers-2023-05-04/

Bostyn, D. H., & Roets, A. (2016). The morality of action: The asymmetry between judgments of praise and blame in the action–omission effect. *Journal of Experimental Social Psychology*, *63*, 19–25. https://doi.org/10.1016/j.jesp.2015.11.005

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of Mice, Men, and Trolleys: Hypothetical Judgment Versus Real-Life Behavior in Trolley-Style Moral Dilemmas. *Psychological Science*, *29*(7), 1084–1093. https://doi.org/10.1177/0956797617752640

Brethel-Haurwitz, K. M., Cardinale, E. M., Vekaria, K. M., Robertson, E. L., Walitt, B., VanMeter, J. W., & Marsh, A. A. (2018). Extraordinary Altruists Exhibit Enhanced Self–Other Overlap in Neural Responses to Distress. *Psychological Science*, *29*(10), 1631–1641. https://doi.org/10.1177/0956797618779590

Brethel-Haurwitz, K. M., O'Connell, K., Cardinale, E. M., Stoianova, M., Stoycos, S. A., Lozier, L. M., VanMeter, J. W., & Marsh, A. A. (2017). Amygdala–midbrain connectivity indicates a role for the mammalian parental care system in human altruism. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1865), 20171731. https://doi.org/10.1098/rspb.2017.1731

Brethel-Haurwitz, K. M., Stoianova, M., & Marsh, A. A. (2020). Empathic emotion regulation in prosocial behaviour and altruism. *Cognition and Emotion*, *0*(0), 1–17. https://doi.org/10.1080/02699931.2020.1783517

Bruneau, E., Cikara, M., & Saxe, R. (2017). Parochial Empathy Predicts Reduced Altruism and the Endorsement of Passive Harm. *Social Psychological and Personality Science*, 934–942. https://doi.org/10.1177/1948550617693064

Burum, B., Nowak, M. A., & Hoffman, M. (2020). An evolutionary explanation for ineffective altruism. *Nature Human Behaviour*, *4*(12), 1245–1257. https://doi.org/10.1038/s41562-020-00950-4

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Cameron, C. D., & Payne, B. K. (2011). Escaping affect: How motivated emotion regulation creates insensitivity to mass suffering. *Journal of Personality and Social Psychology*, *100*(1), 1–15. https://doi.org/10.1037/a0021643

Cardinale, E. M., Breeden, A. L., Robertson, E. L., Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2018). Externalizing behavior severity in youths with callous–unemotional traits corresponds to patterns of amygdala activity and connectivity during judgments of causing fear. *Development and Psychopathology*, *30*(1), 191–201. https://doi.org/10.1017/S0954579417000566

Carlson, R. W., & Zaki, J. (2018). Good deeds gone bad: Lay theories of altruism and selfishness. *Journal of Experimental Social Psychology*, *75*, 36–40. https://doi.org/10.1016/j.jesp.2017.11.005

Caviola, L., David_Althaus, Stefan_Schubert, & Joshua_Lewis. (2022). What psychological traits predict interest in effective altruism? *Effective Altruism Forum*. https://forum.effectivealtruism.org/posts/7f3sq7ZHcRsaBBeMD/what-psychological-traits-predict-interest-in-effective

Caviola, L., Schubert, S., & Greene, J. D. (2021). The Psychology of (In)Effective Altruism. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2021.03.015

Chen, S., & Jordan, C. (2020). Incorporating Ethics Into Brief Mindfulness Practice: Effects on Well-Being and Prosocial Behavior. *Mindfulness*, *11*, 1–12. https://doi.org/10.1007/s12671-018-0915-2

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and Them: Intergroup Failures of Empathy. *Current Directions in Psychological Science*, *20*(3), 149–153. https://doi.org/10.1177/0963721411408713

Cikara, M., Bruneau, E., Van Bavel, J. J., & Saxe, R. (2014). Their pain gives us pleasure: How intergroup dynamics shape empathic failures and counter-empathic responses. *Journal of Experimental Social Psychology*, *55*, 110–125. https://doi.org/10.1016/j.jesp.2014.06.007

Cohen, L., & Godoy, J. (2024). Bankman-Fried sentenced to 25 years for multi-billion dollar FTX fraud | Reuters. *Reuters*. https://www.reuters.com/technology/sam-bankman-fried-be-sentenced-multi-billion-dollar-ftx-fraud-2024-03-28/

Coke, J. S., Batson, C. D., & McDavis, K. (1978). Empathic mediation of helping: A two-stage model. *Journal of Personality and Social Psychology*, *36*(7), 752–766. https://doi.org/10.1037/0022-3514.36.7.752

Crimston, C., Bain, P. G., Hornsey, M. J., & Bastian, B. (2016). Moral expansiveness: Examining variability in the extension of the moral world. *Journal of Personality and Social Psychology*, *111*(4), 636–653. https://doi.org/10.1037/pspp0000086

Crimston, C., Hornsey, M. J., Bain, P. G., & Bastian, B. (2018). Toward a Psychology of Moral Expansiveness. *Current Directions in Psychological Science*, *27*(1), 14–19. https://doi.org/10.1177/0963721417730888

Crockett, M. J., & Lockwood, P. L. (2018). Extraordinary Altruism and Transcending the Self. *Trends in Cognitive Sciences*, *22*(12), 1071–1073. https://doi.org/10.1016/j.tics.2018.09.003

Curry, O., Mullins, D., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, *60*(1). https://ora.ox.ac.uk/objects/uuid:8dd8d82d-3829-4857-bcf4-eebf196d11be

Darley, J. M., & Latane, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*(4, Pt.1), 377–383. https://doi.org/10.1037/h0025589

Davis, M. (1980). A Multidimensional Approach to Individual Differences in Empathy. *JSAS Catalog Sel. Doc. Psychol.*, *10*.

Dawes, A. J., Keogh, R., Andrillon, T., & Pearson, J. (2020). A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-65705-7

Day, M. V., Fiske, S. T., Downing, E. L., & Trail, T. E. (2014). Shifting liberal and conservative attitudes using moral foundations theory. *Personality & Social Psychology Bulletin*, *40*(12), 1559–1573. https://doi.org/10.1177/0146167214551152

Desbordes, G., Gard, T., Hoge, E. A., Hölzel, B. K., Kerr, C., Lazar, S. W., Olendzki, A., & Vago, D. R. (2015). Moving Beyond Mindfulness: Defining Equanimity as an Outcome Measure in Meditation and Contemplative Research. *Mindfulness*, *6*(2), 356–372. https://doi.org/10.1007/s12671-013-0269-8

Dickert, S., Sagara, N., & Slovic, P. (2011). Affective Motivations to Help Others: A Two-stage Model of Donation Decisions. *Journal of Behavioral Decision Making*. https://doi.org/10.1002/bdm.697

Dovidio, J., Johnson, J., Gaertner, S., Pearson, A., Saguy, T., & Ashburn-Nardo, L. (2010). Empathy and intergroup relations. *Prosocial Motives, Emotions, and Behavior*. https://doi.org/10.1037/12061-020

Dunbar, R. I. M., & Shultz, S. (2007). Evolution in the Social Brain. *Science*, *317*(5843), 1344–1347. https://doi.org/10.1126/science.1145463

Earp, B., McLoughlin, K. L., Caraccio, M., Clark, M. S., & Crockett, M. J. (2023). *Impartial Beneficence Predicts Greater Moral Concern for Socially Distant Others—But not at the Expense of Close Relationships* (SSRN Scholarly Paper 4489900). https://doi.org/10.2139/ssrn.4489900

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, *101*(1), 91–119.

Erlich, A., Garner, C., Pennycook, G., & Rand, D. G. (2023). Does Analytic Thinking Insulate Against Pro-Kremlin Disinformation? Evidence From Ukraine. *Political Psychology*, *44*(1), 79–94. https://doi.org/10.1111/pops.12819

Everett, J. A. C., Faber, N. S., Savulescu, J., & Crockett, M. J. (2018). The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence. *Journal of Experimental Social Psychology*, *79*, 200–216. https://doi.org/10.1016/j.jesp.2018.07.004

Everett, J. A. C., & Kahane, G. (2020). Switching Tracks? Towards a Multidimensional Model of Utilitarian Psychology. *Trends in Cognitive Sciences*, *24*(2), 124–134. https://doi.org/10.1016/j.tics.2019.11.012

Eyal, T., Liberman, N., & Trope, Y. (2008). Judging near and distant virtue and vice. *Journal of Experimental Social Psychology*, *44*(4), 1204–1209. https://doi.org/10.1016/j.jesp.2008.03.012

Fowler, Z., Law, K. F., & Gaesser, B. (2021). Against Empathy Bias: The Moral Value of Equitable Empathy. *Psychological Science*, *32*(5), 766–779. https://doi.org/10.1177/0956797620979965

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Galinsky, A. D., Ku, G., & Wang, C. S. (2005). Perspective-Taking and Self-Other Overlap: Fostering Social Bonds and Facilitating Social Coordination. *Group Processes & Intergroup Relations*, *8*(2), 109–124. https://doi.org/10.1177/1368430205051060

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708–724. https://doi.org/10.1037/0022-3514.78.4.708

Gates, B., & Gates, M. (2021). *The year global health went local | Bill Gates*. https://www.gatesnotes.com/2021-Annual-Letter

George, G., Howard-Grenville, J., Joshi, A., & Tihanyi, L. (2016). Understanding and Tackling Societal Grand Challenges through Management Research. *Academy of Management Journal*, *59*(6), 1880–1895. https://doi.org/10.5465/amj.2016.4007

GiveWell. (2023). *GiveWell | Charity Reviews and Research*. GiveWell. https://www.givewell.org/

Gourinat, J.-B. (2007). Akrasia And Enkrateia In Ancient Stoicism: Minor Vice And Minor Virtue? In Akrasia *in Greek Philosophy* (pp. 215–247). Brill. https://brill.com/display/book/edcoll/9789047420125/Bej.9789004156708.i-308_012.xml

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Chapter Two - Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism. In P. Devine & A. Plant (Eds.), *Advances in Experimental Social Psychology* (Vol. 47, pp. 55–130). Academic Press. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the Moral Domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385. https://doi.org/10.1037/a0021847

Graham, J., Waytz, A., Meindl, P., Iyer, R., & Young, L. (2017). Centripetal and centrifugal forces in the moral circle: Competing constraints on moral learning. *Cognition*, *167*, 58–65. https://doi.org/10.1016/j.cognition.2016.12.001

Greaves, H., & MacAskill, W. (2019). *The Case for Strong Longtermism*. https://philpapers.org/rec/GRETCF-4

Greene, J. (2014). *Moral tribes: Emotion, reason, and the gap between us and them*. Penguin Press.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive Load Selectively Interferes with Utilitarian Moral Judgment. *Cognition*, *107*(3), 1144–1154. https://doi.org/10.1016/j.cognition.2007.11.004

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron*, *44*(2), 389–400. https://doi.org/10.1016/j.neuron.2004.09.027

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science (New York, N.Y.)*, *293*(5537), 2105–2108. https://doi.org/10.1126/science.1062872

Groups—EA Forum. (2024). https://forum.effectivealtruism.org/groups

Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science*, *316*(5827), 998–1002. https://doi.org/10.1126/science.1137651

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, *7*(1), 1–16. https://doi.org/10.1016/0022-5193(64)90038-4

Haran, U., Ritov, I., & Mellers, B. A. (2013). The role of actively open-minded thinking in information acquisition, accuracy, and calibration. *Judgment and Decision Making*, *8*(3), 188–201. https://doi.org/10.1017/S1930297500005921

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*(1), 149–160. https://doi.org/10.1016/j.neuron.2010.09.003

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29–29. https://doi.org/10.1038/466029a

Hewstone, M. (1990). The 'ultimate attribution error'? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology*, *20*(4), 311–335. https://doi.org/10.1002/ejsp.2420200404

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture*, *2*(1). https://doi.org/10.9707/2307-0919.1014

Huang, K., Bernhard, R., Barak-Corren, N., bazerman, max, & Greene, J. D. (2020). *Veil-of-Ignorance Reasoning Favors Allocating Resources to Younger Patients During the COVID-19 Crisis* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/npm4v

Hutchinson, J., & Smith, A. D. (1994). *Nationalism*. Oxford University Press.

Jackson, J. E. (2005). Oblimin Rotation. In *Encyclopedia of Biostatistics*. John Wiley & Sons, Ltd. https://doi.org/10.1002/0470011815.b2a13060

Jones, B. A. (2022). A Review of Social Discounting: The Impact of Social Distance on Altruism. *The Psychological Record*, *72*(3), 511–515. https://doi.org/10.1007/s40732-021-00488-5

Jones, B. A., & Rachlin, H. (2009). Delay, Probability, and Social Discounting in a Public Goods Game. *Journal of the Experimental Analysis of Behavior*, *91*(1), 61–73. https://doi.org/10.1901/jeab.2009.91-61

Jones, B., & Rachlin, H. (2006). Social discounting. *Psychological Science*, *17*(4), 283–286. https://doi.org/10.1111/j.1467-9280.2006.01699.x

Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, *125*(2), 131–164. https://doi.org/10.1037/rev0000093

Kang, K. E., Vedlitz, A., Goldsmith, C. L., & Seavey, I. (2023). Optimism and pessimism toward science: A new way to look at the public's evaluations of science and technology

discoveries and recommendations. *Politics and the Life Sciences*, 1–20. https://doi.org/10.1017/pls.2023.9

Kant, I. (1873). *Theory of Ethics*. Longmans.

Kim, H., Ku, B., Kim, J. Y., Park, Y.-J., & Park, Y.-B. (2016). Confirmatory and Exploratory Factor Analysis for Validating the Phlegm Pattern Questionnaire for Healthy Subjects. *Evidence-Based Complementary and Alternative Medicine : eCAM*, *2016*, 2696019. https://doi.org/10.1155/2016/2696019

Klayman, J. (1995). Varieties of Confirmation Bias. In J. Busemeyer, R. Hastie, & D. L. Medin (Eds.), *Psychology of Learning and Motivation* (Vol. 32, pp. 385–418). Academic Press. https://doi.org/10.1016/S0079-7421(08)60315-1

Kogut, T., & Ritov, I. (2005a). The "identified victim" effect: An identified group, or just a single individual? *Journal of Behavioral Decision Making*, *18*(3), 157–167. https://doi.org/10.1002/bdm.492

Kogut, T., & Ritov, I. (2005b). The singularity effect of identified victims in separate and joint evaluations. *Organizational Behavior and Human Decision Processes*, *97*(2), 106–116. https://doi.org/10.1016/j.obhdp.2005.02.003

Kogut, T., Slovic, P., & Västfjäll, D. (2015). Scope insensitivity in helping decisions: Is it a matter of culture and values? *Journal of Experimental Psychology: General*, *144*(6), 1042–1052. https://doi.org/10.1037/a0039708

Kraus, S., & Sears, S. (2009). Measuring the Immeasurables: Development and Initial Validation of the Self-Other Four Immeasurables (SOFI) Scale Based on Buddhist Teachings on Loving Kindness, Compassion, Joy, and Equanimity. *Social Indicators Research*, *92*(1), 169–181. https://doi.org/10.1007/s11205-008-9300-1

Kteily, N. S., & Bruneau, E. (2017). Darker Demons of Our Nature: The Need to (Re)Focus Attention on Blatant Forms of Dehumanization. *Current Directions in Psychological Science*, *26*(6), 487–494. https://doi.org/10.1177/0963721417708230

Lama, Dalai. (2001). *Speech of His Holiness the Dalai Lama to the European Parliament, Strasbourg* [Audio recording].

Lamm, C., Batson, C. D., & Decety, J. (2007, January 10). *The Neural Substrate of Human Empathy: Effects of Perspective-taking and Cognitive Appraisal | Journal of Cognitive Neuroscience | MIT Press Journals*. https://www.mitpressjournals.org/doi/abs/10.1162/jocn.2007.19.1.42

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, *54*(3), 2492–2502. https://doi.org/10.1016/j.neuroimage.2010.10.014

Lasby, D., & Barr, C. (2018). *30 Years of Giving in Canada The Giving Behaviour of Canadians: Who gives, how, and why?*

Law, K. F., Amormino, P., Marsh, A. A., & O'Connor, B. B. (2024). Ethical reasoning versus empathic bias: A false dichotomy? *Trends in Cognitive Sciences*, *28*(1), 1–4. https://doi.org/10.1016/j.tics.2023.10.006

Law, K. F., Campbell, D., & Gaesser, B. (2022). Biased Benevolence: The Perceived Morality of Effective Altruism Across Social Distance. *Personality and Social Psychology Bulletin*, 01461672211002773. https://doi.org/10.1177/01461672211002773

Law, K. F., Syropoulos, S., Coleman, M., Gainsburg, I., & O'Connor, B. B. (2023). *Moral Future-Thinking: Does the Moral Circle Stand the Test of Time?* PsyArXiv. https://doi.org/10.31234/osf.io/c75ny

Law, K. F., Syropoulos, S., O'Connor, B. B., & Young, L. (2024). *The Probabilistic Price of Life Across Time: Generational and Probabilistic Distance Render a Life Today Worth More than Ten Tomorrow*. OSF. https://doi.org/10.31234/osf.io/s346z

Law, K. F., Syropoulos, S., & Young, L. (2023). Why Do Longtermists Care about Protecting the Environment? An Investigation on the Underlying Mechanisms of Pro-Climate Policy Support. *Sustainability*, *15*(24), Article 24. https://doi.org/10.3390/su152416732

Law, K. F., Syropoulos, S., Young, L., & O'Connor, B. B. (2024). *What We Owe the Present: The Perceived Morality of Longtermism and Helping Distant Future People*. OSF. https://doi.org/10.31234/osf.io/s5jkc

Leary, M. R., Tipsord, J. M., & Tate, E. B. (2008). Allo-inclusive identity: Incorporating the social and natural worlds into one's sense of self. In *Transcending self-interest: Psychological explorations of the quiet ego* (pp. 137–147). American Psychological Association. https://doi.org/10.1037/11771-013

Lehmann, L., & Rousset, F. (2020). When Do Individuals Maximize Their Inclusive Fitness? *The American Naturalist*, *195*(4), 717–732. https://doi.org/10.1086/707561

Leung, M.-K., Chan, C. C. H., Yin, J., Lee, C.-F., So, K.-F., & Lee, T. M. C. (2013). Increased gray matter volume in the right angular and posterior parahippocampal gyri in loving-kindness meditators. *Social Cognitive and Affective Neuroscience*, *8*(1), 34–39. https://doi.org/10.1093/scan/nss076

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*(1), 151–158. https://doi.org/10.1037//0022-3514.68.1.151

Lindauer, M., Mayorga, M., Greene, J. D., Slovic, P., Västfjäll, D., & Singer, P. (2020). *Comparing the Effect of Rational and Emotional Appeals on Donation Behavior* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/5wjuv

Lucius Caviola, David Althaus, Stefan Schubert, & Joshua Lewis. (2022). What psychological traits predict interest in effective altruism? *Effective Altruism Forum*. https://forum.effectivealtruism.org/posts/7f3sq7ZHcRsaBBeMD/what-psychological-traits-predict-interest-in-effective

Lutz, A., Brefczynski-Lewis, J., Johnstone, T., & Davidson, R. J. (2008). Regulation of the neural circuitry of emotion by compassion meditation: Effects of meditative expertise. *PloS One*, *3*(3), e1897. https://doi.org/10.1371/journal.pone.0001897

Lyons, M. T. (2006). *Who are the Heroes? Characteristics of People Who Rescue Others*. https://doi.org/10.1556/jcep.3.2005.3-4.2

MacAskill, W. (2016). *Doing Good Better: How Effective Altruism Can Help You Help Others, Do Work that Matters, and Make Smarter Choices about Giving Back* (Reprint edition). Avery.

MacAskill, W. (2018). Understanding Effective Altruism and Its Challenges. In D. Boonin (Ed.), *The Palgrave Handbook of Philosophy and Public Policy* (pp. 441–453). Springer International Publishing. https://doi.org/10.1007/978-3-319-93907-0_34

MacAskill, W. (2022). *What We Owe The Future*. https://whatweowethefuture.com/

Macpherson, R. I. S., & Jr, F. M. B. (2021). Humanitarian Aid Workers: The Forgotten First Responders. *Prehospital and Disaster Medicine*, *36*(1), 111–114. https://doi.org/10.1017/S1049023X20001326

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi

drivers. *Proceedings of the National Academy of Sciences*, *97*(8), 4398–4403. https://doi.org/10.1073/pnas.070039597

Maguire, E. A., Intraub, H., & Mullally, S. L. (2016). Scenes, Spaces, and Memory Traces: What Does the Hippocampus Do? *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *22*(5), 432–439. https://doi.org/10.1177/1073858415600389

Mandler, M. (2020). Distributive Justice for Behavioural Welfare Economics. *The Economic Journal*, *130*(628), 1114–1134. https://doi.org/10.1093/ej/ueaa008

Marsh, A. A. (2016a). Extraordinary Altruism. In J. D. Greene, I. Morrison, & M. E. P. Seligman (Eds.), *Positive Neuroscience* (pp. 143–156). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199977925.003.0010

Marsh, A. A. (2016b). Understanding amygdala responsiveness to fearful expressions through the lens of psychopathy and altruism: Amygdala Responses to Fearful Expressions. *Journal of Neuroscience Research*, *94*(6), 513–525. https://doi.org/10.1002/jnr.23668

Marsh, A. A. (2019). The Caring Continuum: Evolved Hormonal and Proximal Mechanisms Explain Prosocial and Antisocial Extremes. *Annual Review of Psychology*, *70*(1), 347–371. https://doi.org/10.1146/annurev-psych-010418-103010

Marsh, A. A., & Cardinale, E. M. (2012). Psychopathy and fear: Specific impairments in judging behaviors that frighten others. *Emotion*, *12*(5), 892–898. https://doi.org/10.1037/a0026260

Marsh, A. A., Stoycos, S. A., Brethel-Haurwitz, K. M., Robinson, P., VanMeter, J. W., & Cardinale, E. M. (2014). Neural and cognitive characteristics of extraordinary altruists. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(42), 15036–15041. https://doi.org/10.1073/pnas.1408440111

Marzilli Ericson, K. M., White, J. M., Laibson, D., & Cohen, J. D. (2015). Money Earlier or Later? Simple Heuristics Explain Intertemporal Choices Better Than Delay Discounting Does. *Psychological Science*, *26*(6), 826–833. https://doi.org/10.1177/0956797615572232

McFarland, S., Brown, D., & Webb, M. (2012). Identification With All Humanity as a Moral Concept and Psychological Construct. *Current Directions in Psychological Science*, *22*(3), 194–198. https://doi.org/10.1177/0963721412471346

McManus, R. M., Kleiman-Weiner, M., & Young, L. (2020). What We Owe to Family: The Impact of Special Obligations on Moral Judgment. *Psychological Science*, *31*(3), 227–242. https://doi.org/10.1177/0956797619900321

McManus, R. M., Mason, J. E., & Young, L. (2021). Re-examining the role of family relationships in structuring perceived helping obligations, and their impact on moral evaluation. *Journal of Experimental Social Psychology*, *96*, 104182. https://doi.org/10.1016/j.jesp.2021.104182

McRae, E. (2017). Empathy, compassion, and "exchanging self and other" in Indo-Tibetan Buddhism. In *The Routledge Handbook of Philosophy of Empathy*. Routledge.

Mentovich, A., Yudkin, D., Tyler, T., & Trope, Y. (2016). Justice Without Borders: The Influence of Psychological Distance and Construal Level on Moral Exclusion. *Personality and Social Psychology Bulletin*, *42*(10), 1349–1363. https://doi.org/10.1177/0146167216659477

Meyer, M. L., Hershfield, H. E., Waytz, A. G., Mildner, J. N., & Tamir, D. I. (2019). Creative expertise is associated with transcending the here and now. *Journal of Personality and Social Psychology*, *116*(4), 483–494. https://doi.org/10.1037/pspa0000148

Montealegre, A., Bush, L. S., Moss, D., Pizarro, D., & Jimenez-Leal, W. (2020). *Does Maximizing Good Make People Look Bad?* PsyArXiv. https://doi.org/10.31234/osf.io/2zbax

Mwalupaso, G. E., Korotoumou, M., Eshetie, A. M., Essiagnon Alavo, J.-P., & Tian, X. (2019). Recuperating dynamism in agriculture through adoption of sustainable agricultural technology—Implications for cleaner production. *Journal of Cleaner Production*, *232*, 639–647. https://doi.org/10.1016/j.jclepro.2019.05.366

Nations, U. (2022). Human Development Report 2021-22. In *Human Development Reports*. United Nations. https://hdr.undp.org/content/human-development-report-2021-22

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, *125*(6), 737–759. https://doi.org/10.1037/0033-2909.125.6.737

O'Connell, K., Brethel-Haurwitz, K. M., Rhoads, S. A., Cardinale, E. M., Vekaria, K. M., Robertson, E. L., Walitt, B., VanMeter, J. W., & Marsh, A. A. (2019). Increased similarity of neural responses to experienced and empathic distress in costly altruism. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-47196-3

Olivola, C. Y., & Shafir, E. (2013). The Martyrdom Effect: When Pain and Effort Increase Prosocial Contributions. *Journal of Behavioral Decision Making*, *26*(1), 91–105. https://doi.org/10.1002/bdm.767

Ord, T. (2021). *The Precipice: Existential Risk and the Future of Humanity a book by Toby Ord*. Hachette Books. https://bookshop.org/p/books/the-precipice-existential-risk-and-the-future-of-humanity-toby-ord/14906429

Orwig, W., Setton, R., Diez, I., Bueichekú, E., Meyer, M. L., Tamir, D. I., Sepulcre, J., & Schacter, D. L. (2023). Creativity at Rest: Exploring Functional Network Connectivity of Creative Experts. *Network Neuroscience*, 1–31. https://doi.org/10.1162/netn_a_00317

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, *76*(6), 972–987. https://doi.org/10.1037/0022-3514.76.6.972

Patil, I., Zanon, M., Novembre, G., Zangrando, N., Chittaro, L., & Silani, G. (2018). Neuroanatomical basis of concern-based altruism in virtual environment. *Neuropsychologia*, *116*, 34–43. https://doi.org/10.1016/j.neuropsychologia.2017.02.015

Patil, I., Zucchelli, M. M., Kool, W., Campbell, S., Fornasier, F., Calò, M., Silani, G., Cikara, M., & Cushman, F. (2021). Reasoning supports utilitarian resolutions to moral dilemmas across diverse measures. *Journal of Personality and Social Psychology*, *120*(2), 443–460. https://doi.org/10.1037/pspp0000281

Pettigrew, T. F. (1979). The Ultimate Attribution Error: Extending Allport's Cognitive Analysis of Prejudice. *Personality and Social Psychology Bulletin*, *5*(4), 461–476. https://doi.org/10.1177/014616727900500407

Pizarro, D., Uhlmann, E., & Salovey, P. (2003). Asymmetry in judgments of moral blame and praise: The role of perceived metadesires. *Psychological Science*, *14*(3), 267–272. https://doi.org/10.1111/1467-9280.03433

Prinz, J. (2011). *AGAINST EMPATHY - PRINZ - 2011—The Southern Journal of Philosophy—Wiley Online Library*. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2041-6962.2011.00069.x

Rachlin, H., & Jones, B. A. (2008). Social discounting and delay discounting. *Journal of Behavioral Decision Making*, *21*(1), 29–43. https://doi.org/10.1002/bdm.567

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, *17*(8), 413–425. https://doi.org/10.1016/j.tics.2013.06.003

Rhoads, S. A., Vekaria, K. M., O'Connell, K., Elizabeth, H. S., Rand, D. G., Kozak Williams, M. N., & Marsh, A. A. (2023). Unselfish traits and social decision-making patterns characterize six populations of real-world extraordinary altruists. *Nature Communications*, *14*(1), Article 1. https://doi.org/10.1038/s41467-023-37283-5

Riedener, S., Roser, D., & Huppenbauer, M. (2021). *Effective Altruism and Religion: Synergies, Tensions, Dialogue*. Baden-Baden, Germany: Nomos. https://philarchive.org/rec/RIEEAA-3

Robbins, J. M., & Krueger, J. I. (2005). Social Projection to Ingroups and Outgroups: A Review and Meta-Analysis. *Personality and Social Psychology Review*, *9*(1), 32–47. https://doi.org/10.1207/s15327957pspr0901_3

Roff, D. A., & Fairbairn, D. J. (2007). The evolution of trade-offs: Where are we? *Journal of Evolutionary Biology*, *20*(2), 433–447. https://doi.org/10.1111/j.1420-9101.2006.01255.x

Roser, M., & Ortiz-Ospina, E. (2019). Global Extreme Poverty. *Our World in Data*. https://ourworldindata.org/extreme-poverty

Rottman, J., Crimston, C., & Syropoulos, S. (2021). Tree-Huggers Versus Human-Lovers: Anthropomorphism and Dehumanization Predict Valuing Nature Over Outgroups. *Cognitive Science*, *45*(4), e12967. https://doi.org/10.1111/cogs.12967

Schubert, S., & Caviola, L. (2021). *Virtues for Real-World Utilitarians*. PsyArXiv. https://doi.org/10.31234/osf.io/w52zm

Schubert, S., & Caviola, L. (2024). *Effective Altruism and the Human Mind: The Clash Between Impact and Intuition*. Oxford University Press.

Schumann, K., Zaki, J., & Dweck, C. S. (2014). Addressing the empathy deficit: Beliefs about the malleability of empathy predict effortful responses when empathy is challenging. *Journal of Personality and Social Psychology*, *107*(3), 475–493. https://doi.org/10.1037/a0036738

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, *20*, 11–21. https://doi.org/10.1136/jnnp.20.1.11

Sedikides, C., Campbell, W. K., Reeder, G. D., & Elliot, A. J. (1998). The self-serving bias in relational context. *Journal of Personality and Social Psychology*, *74*(2), 378–386. https://doi.org/10.1037/0022-3514.74.2.378

Sharma, E., & Morwitz, V. G. (2016). Saving the masses: The impact of perceived efficacy on charitable giving to single vs. multiple beneficiaries. *Organizational Behavior and Human Decision Processes*, *135*, 45–54. https://doi.org/10.1016/j.obhdp.2016.06.001

Sherif, M., Harvey, O. J., White, B. J., & Hood, W. R. (1961). *Intergroup Conflict and Cooperation: The Robbers Cave Experiment*. 133.

Singer, P. (1981). The Expanding Circle. Ethics and Sociobiology. *Erkenntnis*, *20*(3), 377–381.

Singer, P. (2015). *The Most Good You Can Do: How Effective Altruism Is Changing Ideas About Living Ethically* (1st edition). Yale University Press.

Singer, P. (2016). *Famine, Affluence, and Morality*. Oxford University Press.

Small, D. A., Loewenstein, G., & Slovic, P. (2007). Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes*, *102*(2), 143–153. https://doi.org/10.1016/j.obhdp.2006.01.005

Snyder, M., & Swann, W. B. (1978). Behavioral confirmation in social interaction: From social perception to social reality. *Journal of Experimental Social Psychology*, *14*(2), 148–162. https://doi.org/10.1016/0022-1031(78)90021-5

Sodadee, B., Tongthipya, T., & Phrakhrupariyatvisutthikhun. (2018). *A Study of Form and Value of Logical Rationality in Dhamma Teaching of the Northeastern Buddhist Monks Influenced to the Public Faith* (SSRN Scholarly Paper ID 3250043). Social Science Research Network. https://doi.org/10.2139/ssrn.3250043

Soutschek, A., Ruff, C. C., Strombach, T., Kalenscher, T., & Tobler, P. N. (2016). Brain stimulation reveals crucial role of overcoming self-centeredness in self-control. *Science Advances*, *2*(10), e1600992. https://doi.org/10.1126/sciadv.1600992

Sparkman, D. J., & Hamer, K. (2020). Seeing the human in everyone: Multicultural experiences predict more positive intergroup attitudes and humanitarian helping through identification with all humanity. *International Journal of Intercultural Relations*, *79*, 121–134. https://doi.org/10.1016/j.ijintrel.2020.08.007

Sprecher, S., & Fehr, B. (2005). Compassionate love for close others and humanity. *Journal of Social and Personal Relationships*, *22*(5), 629–651. https://doi.org/10.1177/0265407505056439

Stagnaro, M. N., Pennycook, G., & Rand, D. G. (2018). Performance on the Cognitive Reflection Test is stable across time. *Judgment and Decision Making*, *13*(3), 260–267. https://doi.org/10.1017/S1930297500007695

Syropoulos, S., Crimston, C. R., Markowitz, E., Milfont, T. L., Law, K. F., Claessens, S., Kyritsis, T., Atkinson, Q., Bastian, B., & Rottman, J. (2024a). *Cross-Cultural Insights into Moral Expansiveness: Selective Valuation of Nature Versus Humans*. OSF. https://doi.org/10.31234/osf.io/3z5aj

Syropoulos, S., Crimston, C. R., Markowitz, E., Milfont, T. L., Law, K. F., Claessens, S., Kyritsis, T., Atkinson, Q., Bastian, B., & Rottman, J. (2024b). *Cross-Cultural Insights into Moral Expansiveness: Selective Valuation of Nature Versus Humans*. https://doi.org/10.31234/osf.io/3z5aj

Syropoulos, S., Law, K. F., Amormino, P., & Young, L. (2024a). Mapping the personality of (exceptional) intergenerational concern. *Personality and Individual Differences*, *231*, 112814. https://doi.org/10.1016/j.paid.2024.112814

Syropoulos, S., Law, K. F., Kraft-Todd, G., Mah, A., Markowitz, E., & Young, L. (2024b). Responsibility to future generations: A strategy for combatting climate change across political divides. *British Journal of Social Psychology*, *n/a*(n/a). https://doi.org/10.1111/bjso.12775

Syropoulos, S., Law, K. F., Kraft-Todd, G., & Young, L. (2023). *The Longtermism Beliefs Scale: Measuring Lay Beliefs for Protecting Humanity's Longterm Future*. PsyArXiv. https://doi.org/10.31234/osf.io/e34kv

Syropoulos, S., Law, K. F., Mah, A., & Young, L. (2024c). Intergenerational concern relates to constructive coping and emotional reactions to climate change via increased legacy concerns and environmental cognitive alternatives. *BMC Psychology*, *12*(1), 182. https://doi.org/10.1186/s40359-024-01690-0

Syropoulos, S., Law, K. F., & Young, L. (2024d). Caring for present and future generations alike: Longtermism and moral regard across temporal and social distance. *Group Processes & Intergroup Relations*, 13684302241242115. https://doi.org/10.1177/13684302241242115

Syropoulos, S., Law, K. F., & Young, L. (2024e). Longtermist Education Interventions Increase Concern for and Action to Protect Future Generations. *Social Psychological and Personality Science*, 19485506241228465. https://doi.org/10.1177/19485506241228465

Syropoulos, S., Law, K. F., Young, L., & O'Connor, B. B. (2024f). *Beyond Seven Generations: Mapping and Expanding the Temporal Horizon of Intergenerational Concern*. OSF. https://doi.org/10.31234/osf.io/4j8sn

Syropoulos, S., & Markowitz, E. (2023). Our responsibility to future generations: The case for intergenerational approaches to the study of climate change. *Journal of Environmental Psychology*, *87*, 102006. https://doi.org/10.1016/j.jenvp.2023.102006

Syropoulos, S., Watkins, H. M., Goodwin, G. P., & Markowitz, E. M. (2023). Disentangling the contributions of impact-oriented versus reputation-focused legacy motives on intergenerational concern and action. *Journal of Environmental Psychology*, *90*, 102092. https://doi.org/10.1016/j.jenvp.2023.102092

Tajfel, H. (1970). Experiments in Intergroup Discrimination. *Scientific American*, *223*(5), 96–102. https://doi.org/10.1038/scientificamerican1170-96

Tajfel, H. (1979). Individuals and groups in social psychology. *British Journal of Social and Clinical Psychology*, *18*(2), 183–190. https://doi.org/10.1111/j.2044-8260.1979.tb00324.x

Tamir, D. I., & Mitchell, J. P. (2011). The Default Network Distinguishes Construals of Proximal versus Distal Events. *Journal of Cognitive Neuroscience*, *23*(10), 2945–2955. Journal of Cognitive Neuroscience. https://doi.org/10.1162/jocn_a_00009

Tarrant, M., & Hadert, A. (2010). Empathic Experience and Attitudes Toward Stigmatized Groups: Evidence for Attitude Generalization. *Journal of Applied Social Psychology*, *40*(7), 1635–1656. https://doi.org/10.1111/j.1559-1816.2010.00633.x

Theron Pummer & William MacAskill. (2019). *International Encyclopedia of Ethics* (1st ed.). Wiley. https://doi.org/10.1002/9781444367072

Thompson, C. G., Kim, R. S., Aloe, A. M., & Becker, B. J. (2017). Extracting the Variance Inflation Factor and Other Multicollinearity Diagnostics from Typical Regression Results. *Basic and Applied Social Psychology*, *39*(2), 81–90. https://doi.org/10.1080/01973533.2016.1277529

Toi, M., & Batson, C. D. (1982). More evidence that empathy is a source of altruistic motivation. *Journal of Personality and Social Psychology*, *43*(2), 281–292. https://doi.org/10.1037/0022-3514.43.2.281

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*(7), 1275–1289. https://doi.org/10.3758/s13421-011-0104-1

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, *46*(1), 35–57. https://doi.org/10.1086/406755

Tuen, Y. J., Bulley, A., Palombo, D. J., & O'Connor, B. B. (2023). Social value at a distance: Higher identification with all of humanity is associated with reduced social discounting. *Cognition*, *230*, 105283. https://doi.org/10.1016/j.cognition.2022.105283

Turner, J. C., & Reynolds, K. J. (2003). The Social Identity Perspective in Intergroup Relations: Theories, Themes, and Controversies. In *Blackwell Handbook of Social Psychology: Intergroup Processes* (pp. 133–152). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470693421.ch7

Västfjäll, D., Slovic, P., Mayorga, M., & Peters, E. (2014). Compassion Fade: Affect and Charity Are Greatest for a Single Child in Need. *PLOS ONE*, *9*(6), e100115. https://doi.org/10.1371/journal.pone.0100115

Vekaria, K., Brethel-Haurwitz, K., Cardinale, E., Stoycos, S., & Marsh, A. (2017). Social discounting and distance perceptions in costly altruism. *Nature Human Behaviour*, *1*, 0100. https://doi.org/10.1038/s41562-017-0100

Vekaria, K. M., Hammell, A. E., Vincent, L., Smith, M., Rogers, T., Switzer, G. E., & Marsh, A. A. (2020). The role of prospection in altruistic bone marrow donation decisions. *Health Psychology*, *39*(4), 316–324. https://doi.org/10.1037/hea0000819

Vescio, Theresa K., Sechrist, Gretchen B., & Paolucci, Matthew P. (2003). Perspective taking and prejudice reduction: The mediational role of empathy arousal and situational attributions—Vescio—2003—European Journal of Social Psychology—Wiley Online Library. *European Journal of Social Psychology*, *33*(4). https://onlinelibrary.wiley.com/doi/abs/10.1002/ejsp.163

Vignoles, V. L., Owe, E., Becker, M., Smith, P. B., Easterbrook, M. J., Brown, R., González, R., Didier, N., Carrasco, D., Cadena, M. P., Lay, S., Schwartz, S. J., Des Rosiers, S. E., Villamar, J. A., Gavreliuc, A., Zinkeng, M., Kreuzbauer, R., Baguma, P., Martin, M., … Bond, M. H. (2016). Beyond the 'east–west' dichotomy: Global variation in cultural models of selfhood. *Journal of Experimental Psychology: General*, *145*(8), 966–1000. https://doi.org/10.1037/xge0000175

Vollberg, M. C., Gaesser, B., & Cikara, M. (2019). *Activating episodic simulation increases affective empathy*. PsyArXiv. https://doi.org/10.31234/osf.io/r6wmx

Wade-Benzoni, K. A. (2006). Legacies, Immortality, and the Future: The Psychology of Intergenerational Altruism. In A. E. Tenbrunsel (Ed.), *Ethics in Groups* (Vol. 8, pp. 247–270). Emerald Group Publishing Limited. https://doi.org/10.1016/S1534-0856(06)08012-1

Wade-Benzoni, K. A. (2017). A Golden Rule Over Time: Reciprocity in Intergenerational Allocation Decisions. *Academy of Management Journal*, *45*(5), 1011–1028. https://doi.org/10.5465/3069327

Waytz, A., Cacioppo, J., & Epley, N. (2010). Who Sees Human?: The Stability and Importance of Individual Differences in Anthropomorphism. *Perspectives on Psychological Science - PERSPECT PSYCHOL SCI*, *5*, 219–232. https://doi.org/10.1177/1745691610369336

Waytz, A., Dungan, J., & Young, L. (2013). The whistleblower's dilemma and the fairness–loyalty tradeoff. *Journal of Experimental Social Psychology*, *49*(6), 1027–1033. https://doi.org/10.1016/j.jesp.2013.07.002

Waytz, A., Iyer, R., Young, L., Haidt, J., & Graham, J. (2019). Ideological differences in the expanse of the moral circle. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-12227-0

Weisz, E., & Cikara, M. (2021). Strategic Regulation of Empathy. *Trends in Cognitive Sciences*, *25*(3), 213–227. https://doi.org/10.1016/j.tics.2020.12.002

Wilks, M., McCurdy, J., & Bloom, P. (2023). Who gives? Characteristics of those who have taken the Giving What We Can pledge. *Journal of Personality*, *n/a*(n/a). https://doi.org/10.1111/jopy.12842

Zaki, J. (2018). Empathy is a moral force. In *Atlas of moral psychology* (pp. 49–58). The Guilford Press.

Zaval, L., Markowitz, E. M., & Weber, E. U. (2015). How Will I Be Remembered? Conserving the Environment for the Sake of One's Legacy. *Psychological Science*, *26*(2), 231–236. https://doi.org/10.1177/0956797614561266

APPENDIX A1: EMPATHIC ABILITY MEASURES

Theories of Empathy Scale (Schumann et al., 2014)

Please read each statement below and indicate your agreement with each statement on the scale provided:

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

A person's level of empathy is something very basic about them, and it can't be changed much.
People can't really change how much empathy they tend to feel for others. Some people are very empathic, and some aren't, and they can't change that much.
No matter who somebody is, they can always change how empathic a person they are.
People can always change how much empathy they generally feel for others.
Anybody can change how empathic a person they are.

Parochial Empathy Scale (Bruneau et al., 2017)

Please indicate how much empathy you feel towards people from distant countries in the following situations:

1 (None at All)	2	3	4	5	6	7 (Very Much)

Parents whose children don't do well in school
Parents who cannot buy their children the gift they want for their birthday
Teenagers who get pregnant accidentally
Adults who are not able to find a job despite their qualifications

Interpersonal Reactivity Index Empathic Concern Subscale (Davis, 1980)

The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing the appropriate number on the scale at the top of the page: 1, 2, 3, 4, 5, 6, or 7. READ EACH ITEM CAREFULLY BEFORE RESPONDING. Answer as honestly as you can. Thank you.

Select the number you feel best characterizes the extent to which each item describes you on the scale provided:

1 (DOES NOT DESCRIBE ME WELL)	2	3	4	5	6	7 (DESCRIBES ME VERY WELL)

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

Emotionally Evocative Statements Task (Marsh, & Cardinale, 2012)
For each of the statements below, please indicate which emotion the statement is most likely to elicit if said to another person:

Anger Happiness Sadness Fear Disgust

Don't you have any real friends?
Everything you say is stupid.
I am better than you in so many ways.
I broke your phone on purpose.
I heard someone talking trash about you behind your back.
I stole 20$ from your wallet.
I think you cheated on the test.
I told someone your secret.
I told you to shut up.
I'm going to hurt your friend.
You always ruin everything.
You are a disgrace.
You are a huge idiot.
You are a waste of my time.
You are completely useless.
You are so ignorant sometimes.
You aren't very smart.
You really suck.
You're always wrong.
You're so stupid it's unbelievable.
I found hair in my food.
I haven't showered in days.
I haven't washed my hair in weeks.
I just blew my nose on my sleeve.
I just picked off my scab.
I never wash my hands.
I never wear deodorant.
I reuse my dirty Kleenex.
I saw our waiter sneeze on our food.
I saw someone spit into their hands.
I wear dirty underwear.
I'm chewing gum I found on the ground.

I'm going to spit into my soda.
It smelled like someone pooped in here.
It smells like vomit in the bathroom.
It smells like vomit.
It smells like you stepped in dog poop.
There is a fly in your food.
Your desk looks like a bird pooped on it.
Your lunch smells rotten.
I am going to make you bleed.
I could easily hurt you.
I could kill you if I wanted to.
I don't think you are safe here.
I hope something bad will happen to you.
I think something moved behind you.
I think you are being followed.
I want to hurt you.
I want to make you suffer.
I want to punch you.
I will hurt you if you tell anyone my secret.
I'll be stalking you.
I'll be watching everything you do.
I'll kill your family.
I'm about ready to hit you.
I'm going to beat you up.
I'm going to push you down the stairs.
If you don't leave now you'll be sorry.
You better watch your back.
You can't protect yourself from me.
I always feel better when I am with you.
I always love spending time with you.
I baked you cookies.
I bought you a present.
I found that thing you lost.
I got you tickets to the concert you wanted to go to.
I have an extra ticket to the movie–you should come.
I like you.
I love you.
I love your new idea.
I love your new shirt.
I'm so glad to see you.
You always make me smile.
You are in great shape.
You are really attractive.
You are really smart.
You are the nicest person I know.
You look really good.
You're a great friend.
You're amazing.

Everything you own is gone.
I can't stand you.
I don't like you anymore.
I don't trust you.
I don't want to be friends anymore.
I feel really distant from you.
I forgot your birthday.
I have no respect for you.
I heard your best friend is moving away.
I hope I never see your face again.
I saw your friend get hit by a car.
I used to think you were special.
I'm disappointed in you.
I'm not attracted to you.
Leave me alone.
None of your friends really like you.
Our friendship is over.
You abandoned me when I needed you most.
You really let me down.
You're not invited to my party.

Toronto Alexithymia Scale (Bagby et al., 1994)

Choose one response that best describes how each item applies to you:

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

I am often confused about what emotion I am feeling.
It is difficult for me to find the right words for my feelings.
I have physical sensations that even doctors don't understand.
I am able to describe my feelings easily.
I prefer to analyze problems rather than just describe them.
When I am upset, I don't know if I am sad, frightened, or angry.
I am often puzzled by sensations in my body.
I prefer to just let things happen rather than to understand why they turned out that way.
I have feelings I can't quite identify.
Being in touch with emotions is essential.
I find it hard to describe how I feel about people.
People tell me to describe my feelings more.
I don't know what's going on inside me.
I often don't know why I am angry.
I prefer talking to people about their daily activities rather than their feelings.
I prefer to watch "light" entertainment shows rather than psychological dramas.
It is difficult for me to reveal my innermost feelings, even to close friends.

I can feel close to someone, even in moments of silence.
I find examination of my feelings useful in solving personal problems.
I look for hidden messages in movies or plays.

Levenson Self-Report Psychopathy Scale (Levenson et al., 1995)

The test consists of twenty-six statements that could possibly apply to you. Please indicate your agreement with each statement on the scale provided:

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

Success is based on survival of the fittest; I am not concerned about the losers.
I find myself in the same kinds of trouble, time after time.
For me, what's right is whatever I can get away with.
I am often bored.
In today's world, I feel justified in doing anything I can get away with to succeed.
I find that I am able to pursue one goal for a long time.
My main purpose in life is getting as many goodies as I can.
I don't plan anything very far in advance.
Making a lot of money is my most important goal.
I quickly lose interest in tasks I start.
I let others worry about higher values; my main concern is with the bottom line.
Most of my problems are due to the fact that other people just don't understand me.
People who are stupid enough to get ripped off usually deserve it.
Before I do anything, I carefully consider the possible consequences.
Looking out for myself is my top priority.
I have been in a lot of shouting matches with other people.
I tell other people what they want to hear so that they will do what I want them to do.
When I get frustrated, I often "let off steam" by blowing my top.
I would be upset if my success came at someone else's expense.
Love is overrated.
I often admire a really clever scam.
I make a point of trying not to hurt others in pursuit of my goals.
I enjoy manipulating other people's feelings.
I feel bad if my words or actions cause someone else to feel emotional pain.
Even if I were trying very hard to sell something, I wouldn't lie about it.
Cheating is not justified because it is unfair to others.

APPENDIX A2: REASONING ABILITY MEASURES

7-Item Cognitive Reflection Test (Frederick, 2005; Erlich et al., 2022)

Please answer each of the following questions to the best of your ability:

A racquet and a ball cost $110 in total. The racquet costs $100 more dollars than the ball. How much does the ball cost? ($5)

If it takes 5 machines 5 minutes to make 5 components, how long would it take 100 machines to make 100 components? (5 minutes)

In a lake, there is a patch of lily pads. Everyday, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how many days would it take for the patch to cover half of the lake? (47 days)

If you're running a race and you pass the person in second place, what place are you in? (2^{nd})

A farmer had 15 sheep and all but 8 died. How many are left? (8 sheep)

Maria's father has 5 daughters but no sons - Nana, Nena, Nina, and Nona. What is the fifth daughter's name probably? (Maria)

How many cubic meters of dirt are there in a hole that is 3m deep x 3m wide x 3m long? (0 m^3)

Need for Cognition Scale (Cacioppo, & Petty, 1982)

For each of the statements below, please indicate to what extent the statement is characteristic of you. If the statement is extremely uncharacteristic of you (not at all like you) please select "1"; if the statement is extremely characteristic of you (very much like you) please select "7". Of course, a statement may be neither extremely uncharacteristic nor extremely characteristic of you; if so, please use a number in the middle of the scale that describes the best fit:

1 (Extremely Uncharacteristic)	2 (Uncharacteristic)	3 (Somewhat Uncharacteristic)	4 (Uncertain)	5 (Somewhat Characteristic)	6 (Characteristic)	7 (Extremely Characteristic)

I would prefer complex to simple problems.
I like to have the responsibility of handling a situation that requires a lot of thinking.
Thinking is not my idea of fun.
I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.
I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something.
I find satisfaction in deliberating hard and for long hours.
I only think as hard as I have to.
I prefer to think about small, daily projects to long-term ones.
I like tasks that require little thought once I've learned them.
The idea of relying on thought to make my way to the top appeals to me.

I really enjoy a task that involves coming up with new solutions to problems.
Learning new ways to think doesn't excite me very much.
I prefer my life to be filled with puzzles that I must solve.
The notion of thinking abstractly is appealing to me.
I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.
I feel relief rather than satisfaction after completing a task that required a lot of mental effort.
It's enough for me that something gets the job done; I don't care how or why it works.
I usually end up deliberating about issues even when they do not affect me personally.

Rational Experiential Index (Self-Reported Rational Ability; Pacini, & Epstein, 1999)
Using the following scale, please rate the extent that these statements refer to you:

1 (Not Very True of Myself)	2	3	4	5	6	7 (Very True of Myself)

I'm not that good at figuring out complicated problems
I am not very good at solving problems that require careful logical analysis
I am not a very analytical thinker
Reasoning things out carefully is not one of my strong points
I don't reason well under pressure
I am much better at figuring things out logically than most people
I have a logical mind
I have no problem thinking things through carefully
Using logic usually works well for me in figuring out problems in my life
I usually have clear, explainable reasons for my decisions

Actively Open-Minded Thinking Beliefs Scale (AOT; Baron, 2019; Haran et al., 2013)
Indicate your agreement with each item on the scale provided:

1 (strongly disagree)	2	3	4 (neither agree nor disagree)	5	6	7 (strongly agree)

People should take into consideration evidence that goes against conclusions they favor.
Being undecided or unsure is the result of muddled thinking.
People should revise their conclusions in response to relevant new information.
Changing your mind is a sign of weakness.
People should search actively for reasons why they might be wrong.
It is OK to ignore evidence against your established beliefs.
It is important to be loyal to your beliefs even when evidence is brought to bear against them.
There is nothing wrong with being undecided about many issues.

When faced with a puzzling question, we should try to consider more than one possible answer before reaching a conclusion.
It is best to be confident in a conclusion even when we have good reason to question it.

Heuristics and Bias Tasks (Toplak et al., 2011)

Regression to the Mean Task

After the first 2 weeks of the major league baseball season, newspapers begin to print the top 10 batting averages. Typically, after 2 weeks, the leading batter often has an average of about .450. However, no batter in major league history has ever averaged .450 at the end of the season. Why do you think this is?

1. When a batter is known to be hitting for a high average, pitchers bear down more when they pitch to him
2. Pitchers tend to get better over the course of a season, as they get more in shape. As pitchers improve, they are more likely to strike out batters, so batters' averages go down.
3. A player's high average at the beginning of the season may just be luck. The longer season provides a more realistic test of a batter's skill
4. A batter who has a hot streak at the beginning of the season is under a lot of stress to maintain his performance record. Such stress adversely affects his playing.
5. When a batter is known to be hitting for high average, he stops getting good pitches to hit. Instead, pitchers "play the corners" of the plate because they don't mind walking him.

Covariation Detection Task

A doctor had been working on a cure for a mysterious disease. Finally, he created a drug that he thinks will cure people of the disease. Before he can begin to use it regularly, he has to test the drug. He selected 300 people who had the disease and gave them the drug to see what happened. He selected 100 people who had the disease and did not give them the drug in order to see what happened. The table below indicates what the outcome of the experiment was:

	Cure	
	Yes	No
Treatment present	200	100
Treatment absent	75	25

Please use the slider to indicate whether the treatment is positively or negatively associated with the cure for this disease on a scale ranging from -10 (strong negative association) to +10 (strong positive association):

-10 = strong negative association 0 = no association +10 = strong positive association

-10 -8 -6 -4 -2 0 2 4 6 8 10

Probabilistic Reasoning Task

Assume that you are presented with two trays of black and white marbles: a large tray that contains 100 marbles and a small tray that contains 10 marbles. The marbles are spread in a single layer on each tray. You must draw out one marble (without peeking, of course) from either tray. If you draw a black marble, you win $2. Consider a condition in which the small tray contains 1 black marble and 9 white marbles, and the large tray contains 8 black marbles and 92 white marbles. From which tray would you prefer to select a marble in a real situation?

1. Small Tray
2. Large Tray

Gambler's Fallacy Task

Imagine that we are tossing a fair coin (a coin that has a 50/50 chance of coming up heads or tails) and it has just come up heads 5 times in a row. For the 6th toss do you think that:

1. It is more likely that tails will come up than heads.
2. It is more likely that heads will come up than tails.
3. Heads and tails are equally probable on the sixth toss.

Methodological Reasoning Task

The city of Middleopolis has had an unpopular police chief for a year and a half. He is a political appointee who is a crony of the mayor, and he had little previous experience in police administration when he was appointed. The mayor has recently defended the chief in public, announcing that in the time since he took office, crime rates decreased by 12%. Which of the following pieces of evidence would most deflate the mayor's claim that his chief is competent?

1. The crime rates of the two cities closest to Middleopolis in location and size have decreased by 18% in the same period.
2. An independent survey of the citizens of Middleopolis shows that 40% more crime is reported by respondents in the survey than is reported in police records.
3. Common sense indicates that there is little a police chief can do to lower crime rates. These are for the most part due to social and economic conditions beyond the control of officials.
4. The police chief has been discovered to have business contacts with people who are known to be involved in organized crime.

APPENDIX B: MORAL BELIEFS AND VALUES MEASURES

Moral Foundations Questionnaire (see Graham et al., 2014)

Part 1. When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking? Please rate each statement using this scale:

0- not at all relevant (This consideration has nothing to do with my judgments of right and wrong)	1- not very relevant	2- slightly relevant	3- somewhat relevant	4- very relevant	5- extremely relevant (This is one of the most important factors when I judge right and wrong)

Whether or not someone suffered emotionally
Whether or not some people were treated differently than others
Whether or not someone's action showed love for his or her country
Whether or not someone showed a lack of respect for authority
Whether or not someone violated standards of purity and decency
Whether or not someone was good at math
Whether or not someone cared for someone weak or vulnerable
Whether or not someone acted unfairly
Whether or not someone did something to betray his or her group
Whether or not someone conformed to the traditions of society
Whether or not someone did something disgusting
Whether or not someone was cruel
Whether or not someone was denied his or her rights
Whether or not someone showed a lack of loyalty
Whether or not an action caused chaos or disorder
Whether or not someone acted in a way that God would approve of

Part 2. Please read the following sentences and indicate your agreement or disagreement:

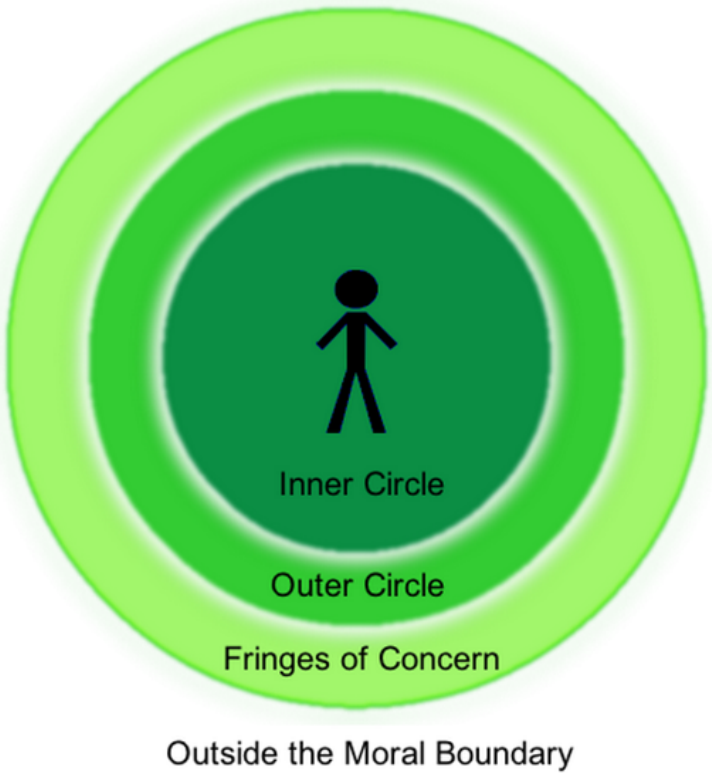
Strongly disagree	Moderately disagree	Slightly disagree	Slightly agree	Moderately agree	Strongly agree

Compassion for those who are suffering is the most crucial virtue.
When the government makes laws, the number one principle should be ensuring that everyone is treated fairly.
I am proud of my country's history.
Respect for authority is something all children need to learn.
People should not do things that are disgusting, even if no one is harmed.
It is better to do good than to do bad.
One of the worst things a person could do is hurt a defenseless animal.
Justice is the most important requirement for a society.
People should be loyal to their family members, even when they have done something wrong.
Men and women each have different roles to play in society.
I would call some acts wrong on the grounds that they are unnatural.
It can never be right to kill a human being.

I think it's morally wrong that rich children inherit a lot of money while poor children inherit nothing.
It is more important to be a team player than to express oneself.
If I were a soldier and disagreed with my commanding officer's orders, I would obey anyway because that is my duty.
Chastity is an important and valuable virtue.

Moral Expansiveness Scale (Crimston et al., 2016)

People sometimes talk about **'circles of moral concern'**. These circles are simple ways to make sense of the levels of moral consideration I have for different entities (e.g., people, animals, and the environment).
Where I place these entities within our moral circles is important as it reflects their moral worth, and has direct consequences for how I treat them.
On the following page you are given the opportunity to organise a range of entities and place them within your own moral circles that reflect your individual views and feelings.



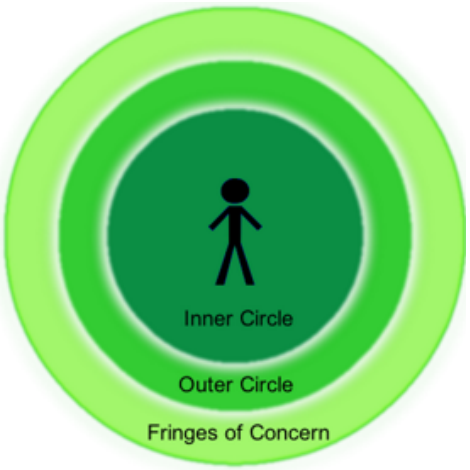
Please read the four boundary descriptions below carefully before completing the moral circle task.

Inner Circle of Moral Concern: These entities deserve the **highest level of moral concern and standing.** You have a moral obligation to ensure their welfare and feel a sense of personal responsibility for their treatment.

Outer Circle of Moral Concern: These entities deserve **moderate moral concern and standing.** You are concerned about their moral treatment; however, your sense of obligation and personal responsibility is greatly reduced.

Fringes of Moral Concern: These entities deserve **minimal moral concern and standing**, but you are not morally obligated or personally responsible for their moral treatment.

Outside the Moral Boundary: These entities deserve **no moral concern or standing.** Feeling concern or personal responsibility for their moral treatment is extreme or nonsensical.



Having carefully read these descriptions, please consider the **level of moral concern you personally have for each of the entities below** and drop each one into the appropriate moral circle box on the right.

Outside the Moral Boundary	Fringes of Moral Concern	Outer Circle of Moral Concern	Inner Circle of Moral Concern

Somebody from your neighborhood
Homosexual
Bee
Close friend
Partner/spouse
Foreign citizen
Head of State for Your Country (Position Not Specific Person)
Grand Canyon National Park
Chimpanzee
Co-worker
Murderer
Member of opposing political party
Coral reef
Family member
Old-growth forest
Apple tree
Terrorist
Dolphin
Somebody with different religious beliefs
Charity worker

Chicken
Soldier from Your Country
Redwood tree
Refugee
Rose bush
Fish
Mentally challenged
individual
Child molester
Citizen of Your Country
Cow

Morality as Cooperation Questionnaire (Curry et al., 2019)

Relevance Items

When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking? (*0–100; not at all relevant, not very relevant, slightly relevant, somewhat relevant, very relevant, extremely relevant*):

Family

Whether or not someone acted to protect their family.
Whether or not someone helped a member of their family.
Whether or not someone's action showed love for their family.

Group

Whether or not someone acted in a way that helped their community.
Whether or not someone helped a member of their community.
Whether or not someone worked to unite a community.

Reciprocity

Whether or not someone did what they had agreed to do.
Whether or not someone kept their promise.
Whether or not someone proved that they could be trusted.

Heroism

Whether or not someone acted heroically.
Whether or not someone showed courage in the face of adversity.
Whether or not someone was brave.

Deference

Whether or not someone deferred to those in authority.
Whether or not someone disobeyed orders.
Whether or not someone showed respect for authority.

Fairness

Whether or not someone kept the best part for themselves.
Whether or not someone showed favouritism.
Whether or not someone took more than others.

Property

Whether or not someone vandalised another person's property.
Whether or not someone kept something that didn't belong to them.
Whether or not someone's property was damaged.

Judgement Items

To what extent do you agree with the following statements? (*0–100; strongly disagree, disagree, neither agree or disagree, agree, strongly agree*):

Family

People should be willing to do anything to help a member of their family.
You should always be loyal to your family.
You should always put the interests of your family first.
Group
People have an obligation to help members of their community.
It's important for individuals to play an active role in their communities.
You should try to be a useful member of society.
Reciprocity
You have an obligation to help those who have helped you.
You should always make amends for the things you have done wrong.
You should always return a favour if you can.
Heroism
Courage in the face of adversity is the most admirable trait.
Society should do more to honour its heroes.
To be willing to lay down your life for your country is the height of bravery.
Deference
People should always defer to their superiors.
Society would be better if people were more obedient to authority.
You should respect people who are older than you.
Fairness
Everyone should be treated the same.
Everyone's rights are equally important.
The current levels of inequality in society are unfair.
Property
It's acceptable to steal food if you are starving. (R)
It's ok to keep valuable items that you find, rather than try to locate the rightful owner. (R)
Sometimes you are entitled to take things you need from other people. (R)

Oxford Utilitarianism Scale (Kahane et al., 2018)
Please rate how much you agree or disagree with each statement:
7-point Likert scale: 1 - Strongly Disagree, 7 - Strongly Agree
1. If the only way to save another person's life during an emergency is to sacrifice one's own leg, then one is morally required to make this sacrifice.
2. It is morally right to harm an innocent person if harming them is a necessary means to helping several other innocent people.
3. From a moral point of view, I should feel obliged to give one of our kidneys to a person with kidney failure since I don't need two kidneys to survive, but really only one to be healthy.
4. If the only way to ensure the overall well-being and happiness of the people is through the use of political oppression for a short, limited period, then political oppression should be used.
5. From a moral perspective, people should care about the well-being of all human beings on the planet equally; they should not favor the well-being of people who are especially close to them either physically or emotionally.
6. It is permissible to torture an innocent person if this would be necessary to provide information to prevent a bomb going off that would kill hundreds of people.
7. It is just as wrong to fail to help someone as it is to actively harm them yourself.
8. Sometimes it is morally necessary for innocent people to die as collateral damage—if more people are saved overall.

9. It is morally wrong to keep money that one doesn't really need if one can donate it to causes that provide effective help to those who will benefit a great deal.

Note. Mean scores on both subscales should be computed. Impartial Beneficence-items 1,3,5,7 and 9; Instrumental Harm-items 2,4,6 and 8.

APPENDIX C: OTHER-INCLUSIVE IDENTITY/POSITIVE OTHER-ORIENTED ATTITUDE MEASURES

Identification With All Humanity Scale (McFarland et al., 2012)

How close do you feel to people all over the world?

A = not at all close	B = not very close	C = just a little or somewhat close	D = pretty close	E = very close

How often do you use the word "we" to refer to people all over the world?

A = almost never	B = rarely	C = occasionally	D = often	E = very often

How much would you say you have in common with people all over the world?

A = almost nothing in common	B = little in common	C = some in common	D = quite a bit in common	E = very much in common

Please answer the following questions using the following choices:

A = not at all	B = just a little	C = somewhat	D = quite a bit	E = very much

Sometimes people think of those who are not a part of their immediate family as "family." To what degree do you think of all humans everywhere as "family?"

How much do you identify with (that is, feel a part of, feel love toward, have concern for) all humans everywhere?

How much would you say you care (feel upset, want to help) when bad things happen to people anywhere in the world?

How much do you want to be a responsible citizen of the world?

How much do you believe in being loyal to all mankind?

When they are in need, how much do you want to help people all over the world?

Compassionate Love for Humanity Scale (Sprecher, & Fehr, 2005)

Indicate the extent to which each statement is true of you on the scale provided:

1: Not at All True of Me	2	3	4	5	6	7: Very True of Me

When I see people I do not know feeling sad, I feel a need to reach out to them.

I spend a lot of time concerned about the well-being of humankind.

When I hear about someone (a stranger) going through a difficult time, I feel a great deal of compassion for him or her.

It is easy for me to feel the pain (and joy) experienced by others, even though I do not know them.

If I encounter a stranger who needs help, I would do almost anything I could to help him or her.

I feel considerable compassionate love for people from everywhere.

I would rather suffer myself than see someone else (a stranger) suffer.

If given the opportunity, I am willing to sacrifice in order to let the people from other places who are less fortunate achieve their goals.
I tend to feel compassion for people even though I do not know them.
One of the activities that provides me with the most meaning to my life is helping others in the world who need help.
I would rather engage in actions that help others, even thought they are strangers, than engage in actions that would help me.
I often have tender feelings toward people (strangers) when they seem to be in need.
I feel a selfless caring for most of mankind.
I accept others whom I do not know even when they do things I think are wrong.
If a person (a stranger) is troubled, I usually feel extreme tenderness and caring.
I try to understand rather than judge people who are strangers to me.
I try to put myself in a stranger's shoes when he or she is in trouble.
I feel happy when I see others (strangers) that are happy.
Those whom I encounter through work and public life can assume that I will be there for them if they need me.
I want to spend time with people I don't know well so that I can help enrich their lives.
I very much wish to be kind and good to fellow human beings.

The Self Construal Scale (Vingoles et al., 2017)

This is a questionnaire that measures a variety of feelings and behaviors in various situations. Listed below are a number of statements. Read each one as if it referred to you. Select the response that best matches:

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree

You like being different from other people.
You see yourself as unique and different from others.
You like it when people notice you in a group.
Being different from others makes you feel uncomfortable.
You try to avoid being noticeably different from others.
Being praised in front of others makes you feel uncomfortable.
Your happiness is unrelated to the happiness of your family.
When you talk about yourself, you don't say very much about your family.
If someone insults a friend, you rarely feel insulted yourself.
If someone in your family is sad, you feel the sadness as if it were your own.
When someone in your family achieves something, you feel proud as if you had achieved something yourself.
Your happiness depends on the happiness of your friends.
You prefer to do what you want without letting your family influence you.
You make decisions about your life on your own.
You always ask your family for advice before making a decision.
Other people have great influence over the choices you make.
You prefer to rely completely on yourself rather than depend on others.
You try to avoid being reliant on others.
You prefer to ask other people for help rather than rely only on yourself.
You feel uncomfortable in situations where you have to rely only on yourself.

You behave in the same way even when you are with different groups of people.
You always see yourself in the same way even when you are with different people.
You behave the same way at home and in public.
You act very differently at home compared to how you act in public.
You see yourself differently in different social environments.
You behave differently when you are with different groups of people.
You prefer to say what you are thinking, even if it is inappropriate for the situation.
You show your inner feelings even if it disturbs the harmony in your family.
You are comfortable expressing disagreement with friends.
You try to adapt to people around you, even if it means hiding your inner feelings.
You feel uncomfortable when you express disagreement with members of your family.
You try to maintain harmony among the people around you.
You value personal achievements more than good relations with the people close to you.
Your own success is very important to you, even if it disrupts your friendships.
You follow your personal goals even if they are very different from the goals of your family.
You value good relations with the people close to you more than your personal achievements.
You always put your family first, even if it means giving up your personal goals.
You are more concerned with your friends' happiness than your own success.

The Self-Other Four Immeasurables Scale (SOFI; Kraus, & Sears, 2008)
Indicate to what extent you have thought, felt, or acted this way toward others during the past week:

	Very slightly or not at all	A little	Moderately	Quite a bit	Extremely
Friendly toward others					
Joyful for others					
Accepting toward others					
Compassionate toward others					

Individual Differences in Anthropomorphism Questionnaire (Waytz et al., 2010)
We will now ask you to rate the extent to which you believe various stimuli (e.g. technological or mechanical items, wild and domestic animals, and natural things) possess certain capacities.

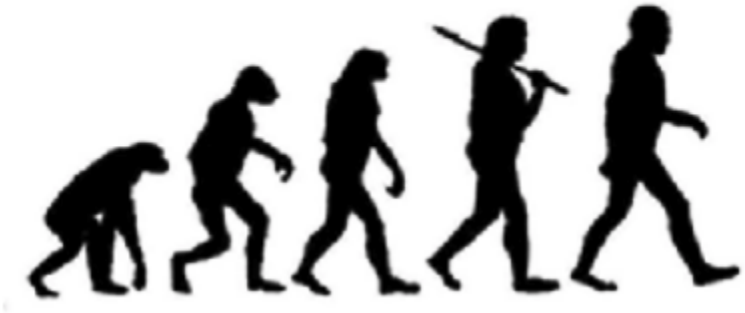
On a 0-10 scale (where 0 = “Not at All” and 10 = “Very much”), please rate the extent to which the stimulus possesses the capacity given.
We will ask you about the extent to which the stimulus has a mind of its own, has free will, has intentions, has consciousness, can experience emotions, is good-looking, is durable, is lethargic, is active, and is useful.
By “has a mind of its own” we mean able to do what it wants.
By “has free will” we mean able to choose and control its own actions.
By “has intentions” we mean has preferences and plans.
By “can experience emotion” we mean it has feelings.
By “has consciousness” we mean able to be aware of itself and its thoughts and feelings.
By “good-looking” we mean attractive.
By “durable” we mean able to withstand wear and damage.
By “lethargic” we mean moving slowly.
By “active” we mean moving frequently and quickly.
By “useful” we mean able to be used for something.

0 (Not at all)	1	2	3	4	5	6	7	8	9	10 (Very much)

To what extent is the desert lethargic?
To what extent is the average computer active?
To what extent does technology - devices and machines for manufacturing, entertainment, and productive processes (e.g., cars, computers, television sets) - have intentions?
To what extent does the average fish have free will?
To what extent is the average cloud good-looking?
To what extent are pets useful?
To what extent does the average mountain have free will?
To what extent is the average amphibian lethargic?
To what extent does a television set experience emotions?
To what extent is the average robot good-looking?
To what extent does the average robot have consciousness?
To what extent do cows have intentions?
To what extent does a car have free will?
To what extent does the ocean have consciousness?
To what extent is the average camera lethargic?
To what extent is a river useful?
To what extent does the average computer have a mind of its own?
To what extent is a tree active?
To what extent is the average kitchen appliance useful?
To what extent does a cheetah experience emotions?
To what extent does the environment experience emotions?
To what extent does the average insect have a mind of its own?
To what extent does a tree have a mind of its own?
To what extent is technology - devices and machines for manufacturing, entertainment, and productive processes (e.g., cars, computers, television sets) - durable?
To what extent is the average cat active?
To what extent does the wind have intentions?

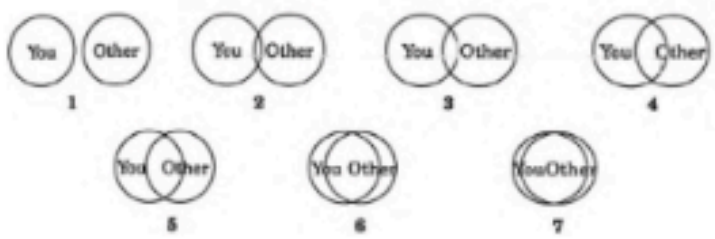
To what extent is the forest durable?
To what extent is a tortoise durable?
To what extent does the average reptile have consciousness?
To what extent is the average dog good looking?
Note. Items containing bolded adjectives above pertain to anthropomorphic qualities. All items were filler. The adjectives were not presented in bold to participants.

Blatant Dehumanization Scale (Kteily & Bruneau, 2017)
Pictured below is the popular 'Ascent of Man' diagram depicting evolutionary progress. Using the diagram as a reference and the slider provided, please indicate where you think each of the following groups belong on the scale, from ape-like human ancestors (0) to 'advanced' modern humans (100).



1. Europeans
2. Japanese
3. Australians
4. Muslims
5. Mexican Immigrants
6. ISIS Members

The Allo-Inclusive Identity Scale (Leary et al., 2008)
Below are seven diagrams that express varying degress of relatedness or connection with some other person or thing. For each of the people or things listed below, indicate which diagram best expresses your relationship with that person or thing. For example, Diagram 1 indicates no relationship or connectedness, Diagram 4 indicates a moderate degree of connectedness, and diagram 7 indicates complete connectedness:



The connection between you and the person with whom you feel closest.
The connection between you and your best friend of your own sex.
The connection between you and a wild animal (such as a squirrel, dear, or wolf).
The connection between you and the average citizen of your country.
The connection between you and the moon.
The connection between you and a homeless person on the street.
The connection between you and your best friend of the other sex.
The connection between you and a dog.
The connection between you and a tree.
The connection between you and a stranger on a bus.
The connection between you and all living creatures.
The connection between you and your family.
The connection between you and the Earth.
The connection between you and an eagle soaring in the sky.
The connection between you and the universe.
The connection between you and a person of another race.

APPENDIX D: EQUITABLE AND EFFECTIVE PROSOCIALITY

Moral Judgment Vignettes (Law et al., 2022)

This portion of the study looks at peoples' reactions to different stories about decision-making. You will read stories of situations where someone needs to decide between multiple options and winds up choosing one. After reading each story, you will be asked to report how morally acceptable or unacceptable you find the decision made by the actor in the passage and explain your reasoning for this judgment.

It's extremely important that you engage with each task as instructed on each trial, because you will need to answer questions about each scenario you read. Focus on engaging with each task as you are instructed, and you'll be fine. Also, please note that experimenters will monitor your responses at the conclusion of the study to ensure you complied with task instructions.

Please read the following stories carefully and answer the questions truthfully.

Please indicate the extent to which you find the action in the story morally acceptable:

1 (completely unacceptable)	2	3	4	5	6	7	8	9 (completely acceptable)

This person wishes to donate some money to a cause. One option is to donate money to a charitable organization which is able to provide (a fellow American, a person from the donor's town, a friend of the donor, a family member of the donor) who is experiencing conditions of famine with food for a month with a $300 donation. Another option is a charitable organization which is able to provide 3 African people living under conditions of famine with food for a month with a $300 donation. The potential donor is unsure of what to do with their money. Eventually, they decide to donate $300 to the charity that can feed 3 African people for $300. To what extent was it morally acceptable for the person in the story to donate money to 3 African people instead of 1 (friend, family member, person from their own country; person from their own town)?

This person recently won a $1,000,000 jackpot from the lottery and wants to donate $100,000 to a cause. One option is to give this money to a hospital, allowing (a person from the same country as the donor, a person from the same town as the donor, a friend of the donor, a family member of the donor) to have a special surgery. While one of these surgeries costs $100,000 to perform, it would save the life of the donor's (compartriot, community member, friend, family member) from a deadly disease. The other option is to give the money to the United Nations to help fight measles and rubella, diseases which plague much of the developing world. It costs $80 to save one life from measles or rubella, so a $100,000 donation would provide enough vaccines to save the lives of 1,250 people from the deadly diseases. The person contemplates where they should donate. Eventually, they decide to donate the $100,000 to the UN charity that can help many people in developing nations. To what extent was it morally acceptable for the person in the story to donate money to the UN charity for people in another country instead of 1 (friend, family member, person from their own country; person from their own town)?

Social Discounting Task (Tuen et al., 2023)
Imagine generating a list of the 100 people closest to you in the world. Number 1 would be your closest friend or relative and number 100 would be a distant acquaintance. Please abstain from listing financial benefactors (e.g., parents, grandparents, spouse).
Please provide the first name and last initial of the person occupying (spot 1, spot 2, spot 4, spot 10, spot 15):
Note. Participants then make a series of 27 choices between a smaller amount of money for themselves and a larger amount for others, using piped text to include the names of the people they listed in part one of the task described above. Please see below for an example:

Would you prefer $25 for yourself or $55 for (name of the person at spot 1)?

Longtermism Beliefs Scale (Syropoulos & Law et al., 2023)
1. I should act wisely because what I do today will influence an untold number of people.
2. It is important to consider the long-term consequences of our actions and decisions.
3. Intergenerational cooperation is important for addressing long-term challenges.
4. It is important that I reduce existential risks and promote sustainable development goals.
5. I should always have in view not only the present but also the coming generations.
6. There are things I can do to steer the long term future to a better course.
7. Positively influencing the long term future is a key moral priority of our time.

Note. Participants responded to each item with respect to people living 1,000, 10,000, 100,000, and 1,000,000 years in the future. Each item was presented on a separate page, and participants responded to each item for the four timeframes simultaneously on a 0 (strongly disagree) – 100 (strongly agree) slider scale.

Impact Legacy Motives Scale (Zaval et al., 2015; Syropoulos et al., 2023)

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

When thinking about the future…
- …It is important to me that my actions help future people.
- …I want to have an enduring positive impact on society.
- …It is important for me to leave a legacy of benefiting others.
- …I want my life to impact others in a positive way.

Behavioral Donation Task (Developed by Researcher)

You have been selected to help choose which charity to allocate a $1 donation towards. You will now be shown a series of choices where you will be asked to decide which of two different charities you would like the dollar to be donated to. The experimenter will donate the dollar to a charity you select on one of the following trials.

Note: Each charity from the lettered list (i.e., List I) is paired with each charity from the numbered list (i.e., List II) for a total of 16 trials. The number of decisions each participant makes to allocate the dollar towards charities in List II is subtracted from the number of decisions each participant makes to allocate the dollar towards charities in List I, yielding a composite "Effectiveness Index". The Effectiveness Index ranges from -16 (all ineffective charities) to +16 (all effective charities), with a score of '0' indicating 8 effective decisions and 8 ineffective decisions.

List I: Most Effective (GiveWell, 2022):
A
Malaria Consortium: Providing Medicine to Prevent Malaria in sub-Saharan Africa

Malaria is a deadly disease rampant in sub-Saharan Africa. Seasonal malaria chemoprevention is preventive medicine that saves children's lives. It is given during the four months of the year when malaria infection rates are especially high. Malaria Consortium saves lives in sub-Saharan Africa by funding the administration of this medication.
B
Against Malaria Foundation: Providing Nets to Prevent Malaria in sub-Saharan Africa

Malaria is a deadly disease rampant in sub-Saharan Africa. Bed nets save lives. Participants hang the nets and sleep under them so they are not bitten by malaria-carrying mosquitoes. The Against Malaria Foundation saves lives in sub-Saharan Africa by funding the provision of bed nets.
C
Hellen Keller International: Providing Supplements to Prevent Vitamin A Deficiency Internationally

Vitamin A deficiency leaves children vulnerable to infections and often leads to death. Vitamin A supplements can restore vitamin A to healthy levels. Hellen Keller International saves lives internationally by providing vitamin A supplements to children under 5 years old.
D
New Incentives: Providing Cash Incentives for Routine Childhood Vaccines in Nigeria

In Nigeria, many infants do not receive all of their recommended vaccines. Vaccines reduce the transmission of deadly, preventable illnesses. New Incentives saves lives in Nigeria by providing cash transfers to incentivize caregivers to bring babies to clinics for routine childhood vaccinations.
List II: Less Effective (Consumer reports, charity navigator):
1
Make-A-Wish America: Granting Wishes to Sick Children in America

Make-A-Wish America grants wishes to American children under the age of 18 with life threatening medical conditions to enrich the human experience with hope, strength and joy. Wishes typically fall into one of four categories: to go on a trip, to have something, to meet a celebrity or to be someone (a policeman, astronaut, actor, etc.).

2

Childhood Leukemia Foundation: Educating and Empowering American Childhood Cancer Patients

Childhood Leukemia Foundation's programs educate and empower childhood cancer patients in America. The organization primarily provides educational binders to parents of children with cancer, wigs to children suffering from cancer-treatment-related hair loss, and educational wish baskets containing toys, games, and iPads to American childhood cancer patients.

3

Help Heal Veterans: Enriching the Lives of American Veterans with Arts and Crafts

Help Heal Veterans offers a variety of therapeutic craft kits free of charge to America's veterans, both in-home and at community craft centers. The kits use recycled and sustainable materials, promote healing and show American veterans that they are remembered and cared about.

4

National Caregiving Foundation: Using Mailings to Educate the American Pubic About Alzheimer's Disease

The National Caregiving Foundation uses direct mail to communicate to the American caregiving community. They use their mailings to educate the American public about Alzheimer disease, including warning signs and symptoms and offer suggestions to caregivers, including care for wounded soldiers.

Effective Altruism Interest Scale (Caviola et al., 2022)

1 (strongly disagree)	2	3	4 (neither agree nor disagree)	5	6	7 (strongly agree)

Effectiveness Focus.

Imagine a situation where you intend to do good (e.g., to improve others' lives or the world) with a certain limited amount of resources available (e.g., your time or money). You can decide how to allocate your resources by choosing from different options that all do good. The stakes are high. In such a situation, when you can choose between different options of doing good...

It would be wrong to do something that only does some amount of good if there is an alternative course of action that would do much more good.
It would be the right choice to refrain from helping one person if that makes it possible to help a larger number of people.
Helping one person is less valuable than helping two people to the same extent.
You should follow evidence and reason to do what is most effective, even if you emotionally prefer another option.
The most important consideration is effectiveness - choosing the option that does the most good per resource invested.
You should usually help a large group of people over a smaller group, even if it seems unfair.

Expansive Altruism.
To what extent do you agree or disagree with the following statements?

As long as my and my family's basic material needs are covered, I want to use a significant amount of my resources (e.g., money or time) to improve the world.
I am willing to make significant sacrifices for people in need that I don't know and will never meet.
People in wealthy countries should donate a substantial proportion of their income to make the world a better place.
I would make a career change if it meant that I could improve the lives of people in need.
We should put a lot of emphasis on the well-being of people who live today.
From a moral perspective, the suffering of all beings matters roughly the same, no matter what species they belong to.
Note. The bolded items were used to capture altruistic enkrateia (AES).

Responsibility to Future Generations Scale (Syropoulos & Markowitz, 2023)
Please indicate your agreement with each statement on the scale provided:

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

When deciding how to live, I have a duty to consider the impact of my actions on future generations.
We need to reform to benefit future generations even if that means making some sacrifices now.
I feel personally responsible for protecting future generations.

Real-World Charitable Action (Income Donated and Time Volunteering)
In a given year, what percentage (out of 100) of your yearly income do you donate to charity?
In a given year, what percentage (out of 100) of your time do you devote towards volunteering to help others?

APPENDIX E: DEMOGRAPHICS

Demographics

INST

Thank you. Before concluding the study, we ask that you please provide us with the following information about yourself. We use this information to report the characteristics of people who helped with this research.

Age

What is your age in years?

Nat

What is your nationality (for example, United States)?

Lang

What is your native/first language (for example, English)?

Gender

What is your gender?

- female
- male
- other

Race

How would you describe your race/ethnicity? Please select all that apply:

- Caucasian/White
- African American/Black
- Hispanic/Latino/Latina
- Asian/Asian American
- Native American
- Pacific Islander
- Middle Eastern
- Multiracial
- Other (please specify):

Politics

What is your political orientation?

- very liberal
- liberal
- somewhat liberal
- moderate/middle of the road
- somewhat conservative
- conservative
- very conservative

Highest Level

What is the highest degree or level of school you have completed? If currently enrolled, highest degree received.

- No schooling completed
- Nursery school to 8th grade
- Some high school, no diploma
- High school graduate, diploma or the equivalent (for example: GED)
- Some college credit, no degree
- Trade/technical/vocational training
- Associate degree
- Bachelor's degree
- Master's degree
- Professional degree
- Doctorate degree

DonationMoney

In a given year, what percentage (out of 100) of your yearly income do you donate to charity?

DonationTime

In a given year, what percentage (out of 100) of your time do you devote towards volunteering to help others?

Income

What was your total household income before taxes during the past 12 months?

- Less than $25,000
- $25,000-$49,999
- $50,000-$74,999
- $75,000-$99,999
- $100,000-$149,999
- $150,000 or more
- Prefer not to say

Children

How many children under 18 live with you?

Employment

What best describes your employment status over the last three months?

- Working full-time
- Working part-time
- Unemployed and looking for work
- A homemaker or stay-at-home parent
- Student
- Retired
- Other

Marriage

What is your current marital status?

- Married
- Living with a partner
- Widowed
- Divorced/Separated
- Never been married

Household

How many people live or stay in this household at least half the time?

Religiosity

What is your religion?

- No Religion/Athiest
- Unsure/Agnostic
- Christian (all denominations)
- Buddhist
- Hindu
- Jewish
- Muslim
- Sikh
- Any other religion

Religiosity 2

How devout would you describe your religiosity?

- Far below average
- Somewhat below average
- Average
- Somewhat above average
- Far above average