2 Moral Character

Geoffrey P. Goodwin and Justin F. Landy

Edgar McGregor is a 20-year-old climate activist who regularly posts videos on Twitter showing him cleaning up trash in the park near where he lives. As of this writing, he has just posted his 955th trash pickup announcement. He does this work alone and, apparently, for no monetary reward. His actions evoke admiration and even awe, because he appears to be acting selflessly in service of a humanitarian and ecological goal that is greater than himself. His actions are those of a person possessing a notably high degree of moral character.

A person's *moral character* comprises the moral dimension of their personality (Kupperman, 1991). Although no definition of morality is universally endorsed, one feasible and ecumenical conception is that morality concerns a system of informal public norms pertaining to serious matters of right and wrong, or good and bad (see, e.g., Gert & Gert, 2020). Accordingly, a moral issue is one that is captured within such a system of norms, and moral character pertains to a person's propensity to enact moral behaviors, and to think, desire, and emote in ways that are relevant to such moral norms (see also Fleeson et al., 2014).

Within late twentieth-century moral philosophy, the revival of virtue ethics as a competitor to deontology and consequentialism spurred renewed interest in moral character. The core notion encapsulated within virtue ethics is that an act's moral goodness (or rightness) should be evaluated in terms of whether it accords with what a virtuous agent would do under the circumstances, rather than in terms of its consequences or its accordance with a moral rule. Such views come in several distinct forms, which differ in their conceptions of what virtue is and how it should guide our decision making (e.g., Adams, 2006; Annas, 2011; Anscombe, 1958; Foot, 1978; Hursthouse, 1999; McDowell, 1979; Swanton, 2003; for a critical perspective, see Louden, 1984). Nonetheless, at the center of each of these views is the idea that virtue and moral character are foundational concepts in ethics. This resurgence of interest in virtue ethics has refocused attention on moral character traits within philosophy, and has spurred targeted inquiry into specific moral character traits (for a review of virtue ethics, see Hursthouse & Pettigrove, 2018).

Alongside these philosophical developments, moral character has undergone a recent revival in the field of both personality and social psychology, having earlier been eschewed as a topic of inquiry. In the course of this review, we aim to summarize the role of moral character in four distinct areas of psychology.

First, we examine the question of whether moral character actually exists. Second, we review evidence demonstrating that moral character is a fundamental input to person perception and global impression formation. In this section, we will also consider limits to the hypothesis that morality is central to person perception, including whether there are clear cases in which moral people are not preferred or evaluated positively. Third, we review evidence that moral character is considered central to people's identity. Last, we examine how people draw inferences of moral character.

2.1 Does Moral Character Exist?

Although it seems commonsensical that there is such a thing as moral character, which people possess to varying degrees, there are noted skeptics of this idea (Doris, 1998, 2002; Harman, 1999). Indeed, some philosophers have argued that belief in moral character is not only empirically unsupported but also dangerous (Harman, 1999). Skepticism about broad moral character traits comprises at least two main themes (see Fleeson et al., 2015; Helzer et al., 2018). First, it is claimed that there are insufficient cross-situational correlations between purported moral character traits. This claim traces back to the seminal work of Hartshorne and May (1928), who observed that the correlation between children's honesty behaviors across any two situations was only around 0.20, which seemed too low to support the existence of broad character traits. Second, the success of situationist social psychology is taken to suggest that situational rather than personality variables overwhelmingly determine moral behavior. Milgram's (1963, 1974) obedience studies and Darley and Latane's (1968) bystander intervention study are among those held up as powerful illustrations of the role of situational forces.

These critiques were arguably exaggerated even at the time they were made (see, e.g., Miller, 2003), but recent evidence has made them even less tenable. The central question at issue is the degree of consistency present in people's moral behavior, across time and across situations.

One sort of recent study shows direct evidence for cross-situational stability in specific moral character traits. Meindl et al. (2015) tracked college students for nine days using an experience sampling method and found consistent differences between people in their (self-reported) levels of honesty, compassion, and fairness. Knowing how someone behaved at one point during a given day predicted how they would behave at a later point during the day (on the same trait), with correlations ranging from 0.35 to 0.74. More importantly, when behavior was aggregated, knowing how someone behaved during one half of the study predicted how well they would behave in the other half, with correlations ranging from 0.66 to 0.97 (see also Bleidorn & Denissen, 2015, for similar evidence). Bollich et al. (2016) used a similar longitudinal design but used coders to categorize snippets of speech rather than relying on self-reports. These researchers also observed significant and sizable within-person

correlations for moral behaviors such as empathy, gratitude, blame, and condescension, when examining both individual behaviors and behaviors in the aggregate ($rs > 0.47$).

Research within personality psychology over the past 20 years also provides broad support for the existence of stable moral character traits. During this time, an alternative to the "Big Five" model of personality was proposed, which added a sixth dimension designed specifically to capture honesty and humility (Ashton & Lee, 2007; Lee & Ashton, 2004). Many specific moral character traits have also been explored in depth, including gratitude (McCullough et al., 2002; Wood et al., 2008) and humility (Davis et al., 2011). Personality psychologists have also studied antisocial aspects of personality, including the recently proposed "dark factor" of personality, which purportedly extends beyond the narrower "dark triad" of personality traits (Moshagen et al., 2018).

An additional form of evidence examines consistency between ratings of a target's moral character. Cohen et al. (2013) found moderate correlations between a target person's own ratings of their honesty-humility and guilt proneness and the ratings of that person made by a well-acquainted other. Similarly, Helzer et al. (2014) also found moderate self–other agreement in ratings of moral character traits. Perhaps more important, they also observed correlations between several others' ratings of a target person's moral character – both in terms of their general moral character as well as in terms of discrete moral character traits (fairness, honesty, compassion, temperance, moral concern). Although this provides only indirect evidence, its most parsimonious explanation is that moral character is both visible to others and cross-situationally consistent.

In sum, there is accumulating evidence for the cross-situational and intertemporal stability of a variety of moral character traits, such that radical skepticism about their existence is not well supported.

2.2 Moral Character and Person Perception

With this backdrop in place, we turn now to the role of moral character in person perception and social cognition. Just as personality psychologists eschewed study of moral character for many years, considering it too "value-laden" to constitute a respectable part of personality science (see Nicholson, 1998, for a review), so too did social cognition researchers. Instead, the field emphasized evaluations of "warmth," rather than morality, as fundamental. Warmth overlaps with morality to some degree, but there are dissociations, such that a given behavior or trait can be warm but not moral (e.g., extroverted, funny), or moral but not warm (e.g., honest, principled).

In our view, this historical focus on warmth obscured the fundamental importance of moral character to the impressions we form of others. We defend this claim, beginning first with a review of research on warmth,

followed by a review of more recent evidence on the fundamental role of morality in person perception.

2.2.1 Early Studies of the Warmth–Coldness Dimension

An especially influential early finding showed that adding the single terms "warm" or "cold" to a list of traits describing a person generated wide-ranging inferences about other aspects of that person (Asch, 1946). For instance, warm individuals were inferred to be more humane, wise, altruistic, and even good-looking than cold individuals (see also Brambilla et al., 2021, for a review).

Warmth had a long legacy in social cognition following Asch's seminal studies. Rosenberg et al. (1968) conducted a highly influential study in which participants were asked to sort 64 trait terms into categories that were likely to be associated within the same people. Their results suggested that the traits could be organized in a two-dimensional space, with one dimension capturing good and bad social properties (warmth) and another dimension capturing good and bad intellectual properties (competence). This study is very well known and was a major precursor of later two-dimensional models of group and individual perception, which made a similar distinction between warmth and competence, or related terms (e.g., Abele & Wojciszke, 2013; Fiske et al., 2002; Judd et al., 2005). However, much less well known to researchers is that Rosenberg did not himself strongly defend a two-dimensional approach to person perception. In a later study using the same task, but with a different set of 60 traits, Rosenberg and Olshan (1970) found that a three-dimensional solution better fit the data (hard–soft, good–bad, active–passive). Using a somewhat different task, Peabody (1984) suggested that a minimum of four dimensions were needed to capture personality perception, though the exact dimensions depended on the method of factor extraction. Nonetheless, the two-dimensional approach established a firm foothold and came to dominate the study of person perception in the ensuing years.

2.2.2 The Rise of Morality

From our perspective, however, a notable shortcoming of such models is that they pay little attention to the moral dimension of human personality. On the one hand, many traits that are central to warmth have little to do with morality (e.g., sociability, extroversion, playfulness, and so on). And on the other hand, while there are some moral traits that overlap conceptually with warmth (e.g., kindness, benevolence), there are also many highly central moral traits (e.g., trustworthiness, courage, integrity, and so on) that can be enacted without warmth. For instance, Edgar McGregor's dedication to preserving his local natural park does not require warmth; indeed, it is not even an interpersonal behavior. Moreover, the operationalization of warmth in the literature has been quite inconsistent and has sometimes neglected morality altogether. For instance, Fiske et al. (2002) operationalized warmth as *warmth*,

good-naturedness, *tolerance*, and *sincerity*; the former two traits have minimal moral relevance but the latter two are quite important to moral character. However, Kervyn et al. (2012) operationalized warmth only as *warmth*, *friendliness*, *niceness*, and *sociability*, and Cuddy et al. (2007) operationalized it as simply *warmth* and *friendliness*. In these latter two cases, a substantive moral element is lacking. Thus, although warmth has some conceptual connection to aspects of morality, it has been operationalized inconsistently, often with no moral content at all.

An important paper by Wojciszke, Bazinska, and Jaworski (1998) refocused the field's attention on morality. The goal of this investigation was to compare the relative influence of morality and competence information on global impressions of others. Global impressions represent people's overall impressions of people on a simple, valenced, positive–negative response dimension. Using a sequence of varied and elegant study designs, these researchers established the overall dominance of morality over competence information. They first showed that when people are asked to state the traits that they "personally think are most important in others," the traits they generate tend to be related more to morality than to competence (Study 1). Next, they demonstrated that when people rated individuals they are acquainted with on both morality- and competence-related traits, the moral traits predicted their global impressions of these people appreciably better than did the competence traits (Study 3). Finally, using an experimental design, they showed that morality traits exerted a larger causal influence on global impressions than did competence traits, with an effect size more than double that for competence traits (Study 4). These findings strikingly highlight the importance of morality in person perception.

The focus of our own research was to establish a further separation between warmth and morality. In light of the tendency for research in the two-dimensional tradition to blur the distinction between warmth and morality, and sometimes to ignore morality altogether, we aimed to discover whether these dimensions are truly separate evaluations and, if so, which one exerts a greater influence on person perception. Our first step, following Wojciszke, Dowhyluk, and Jaworski (1998), was to conduct a "bottom-up" norming study, in which 170 traits were rated on their usefulness for evaluating a person on various higher-order dimensions, including morality, warmth, and abilities (i.e., competence). This study then set the stage for later studies, which separated traits into four categories: those useful for judging morality but less useful for judging warmth ("pure morality traits," e.g., honest, trustworthy, principled), those useful for judging warmth but less useful for judging morality ("pure warmth traits," e.g., sociable, warm, funny), those useful for judging both dimensions ("blended traits," e.g., kind, humble, cooperative), and those not useful for judging either dimension (e.g., athletic, musical, intelligent).

In one study, participants thought about individuals they were familiar with, and rated them on a variety of different traits (Goodwin et al., 2014, Study 3). The result of this study was that the pure morality traits predicted global impressions of the target individuals much better than did the pure warmth

traits. Another correlational study showed that people's valenced impressions of prominent deceased individuals described in the *New York Times* obituary page were better predicted by independent ratings of their morality than by ratings of their warmth (Study 7).

Experimental studies further showed that when morality and warmth were manipulated orthogonally in descriptions of hypothetical targets who varied in their interpersonal closeness to the self, the overall effect of morality was consistently larger than that of warmth (Studies 4–6). Furthermore, the dominance of morality over warmth was clearest for social roles rated as most important. Complementing this research, other research shows that evaluations of morality are central to both liking and respecting others (Hartley et al., 2016). And people tend to *search* preferentially for morality information when forming impressions of others as well. That is, they think that learning about moral traits would be more relevant than learning about either sociability or competence traits for forming a global impression of another person (Brambilla et al., 2011). Thus, overall, the evidence from these studies strongly indicates that morality is separable from warmth, and a stronger overall contributor to overall person impressions (for a review, see Goodwin, 2015).

2.2.3 Morality Influences Impressions Differently Than Sociability or Competence

One question that lingered after these studies, however, was how decisively morality can be distinguished from warmth or sociability.[1] Based on the studies we have described, our view was that this separation was quite pronounced. An alternative view is to try to absorb these results within a two-dimensional framework, by arguing that morality and sociability comprise separate facets of an overarching warmth dimension (see, e.g., Fiske et al., 2002, who note that "the warmth scale includes elements of both sociality [good-natured, warm, tolerant] and morality [sincere], but all are prosocial traits" [p. 889]; see also Bergsieker et al., 2012, who state that "success in navigating interpersonal interactions requires accurately inferring others' warmth (i.e., morality) and competence" [p. 1216]; and Fiske et al., 2007, who describe warmth as "the moral-social dimension" [pp. 77–78]).

We therefore attempted to adjudicate between these two theoretical views. We reasoned that, from a functional perspective, morality is important in global impressions because it indicates the nature of others' intentions toward us. Moral individuals are likely to be beneficial or benign, while immoral individuals are likely to intend harm. In contrast, both competence and sociability information serve as "amplifiers" of a person's intentions. If a moral person is

[1] In our original investigations on this topic, we contrasted morality and warmth (Goodwin et al., 2014), whereas in our later investigations we contrasted morality and sociability (Landy et al., 2016), which arguably represents a similar but slightly cleaner contrast. As we have already described, warmth incorporates some elements of morality whereas sociability does not.

competent, or sociable, this is a good thing, because they will be more effective in bringing about or promoting their desired ends. This is clear for competence, which involves effective goal pursuit by definition. Indeed, Wojciszke, Bazinska, and Jaworski (1998, Study 4) had previously found that a competent, immoral actor was viewed more negatively than an incompetent, immoral actor. We theorized that this logic should also apply to sociability, in that highly sociable individuals are better able to recruit allies, persuade others, and generally drum up support for their desired ends. The natural extension of this reasoning is that, similar to competence, sociability *is not* desirable in immoral persons. An immoral person who is also competent or sociable is more dangerous than one who lacks these traits. Thus, whereas moral traits are valuable regardless of the other traits that a person possesses, both competence and sociability traits are only valuable conditional on the moral traits that the person possesses. We dubbed this pair of ideas the *morality dominance hypothesis* and the *morality dependence hypothesis* (Landy et al., 2016).

To test these predictions, we employed a variety of study designs. One pair of studies examined global impressions by factorially manipulating a target's morality, competence, and sociability at the level of abstract traits (Landy et al., 2016, Study 2) or descriptions of behavior (Study 3). Morality dominated these impressions, such that moral individuals were always viewed either neutrally or positively, whereas immoral individuals were always viewed very negatively. Sociability and competence information had much less influence on people's impressions.

A subsequent study examined people's preferences for various traits in hypothetical targets (Landy et al., 2016, Study 4). Participants greatly preferred targets to be moral rather than immoral, regardless of whether the target was already known to be sociable or unsociable, competent or incompetent, thereby supporting the morality dominance hypothesis. Furthermore, when a target was known to be moral, participants preferred them also to be sociable and competent, but when targets were known to be immoral, participants preferred them to be unsociable and incompetent, thereby supporting the morality dependence hypothesis. Corroborating these results, when participants were asked to anticipate potential *changes* in their overall impressions, they indicated that their impression of a highly immoral target would become significantly *more negative* with the addition of positive information about his sociability or competence, whereas they indicated an anticipated positive shift when the target was only slightly immoral, or when he was moral (Landy et al., 2016, Study 5). Thus, in both studies, morality was valued unconditionally, whereas competence and sociability were valued conditional on a target person's prevailing morality.

In sum, morality and sociability function quite differently in how they contribute to global impressions. Morality contributes positively to impressions, regardless of a person's other traits, whereas sociability can be negative in the presence of immorality. If it were the case that morality and sociability are both subcomponents of a superordinate warmth dimension, one would not expect that they would be thought about in such divergent ways.

Thus, these findings pose a compelling challenge to two-dimensional models of social cognition.

2.2.4 Limitations to Morality Dominance

The unconditional valuation of morality traits provides an additional means of illustrating the considerable power of morality in shaping social evaluations, consistent with much other research. Yet, it would be premature to conclude from this work that all morality traits are valued at all times, in all contexts, and by all people. Several lines of research have sought to identify circumstances in which people respond negatively to morality in others.

Some research suggests that people dislike morality when others' righteous behavior seems implicitly to impugn their *own* morality (Monin et al., 2008). In one study, participants first completed a task that was plausibly construed as assessing racial bias (Monin et al., 2008, Study 2). The task induced White participants to choose an African American man as the culprit for a robbery (the evidence implicated him in particular), and most of them did indeed do so. These participants were then asked to make judgments of another (fictional) participant who had objected to the task from the outset and refused to participate further, claiming that the task itself reflected racial bias on the researchers' part. These "actor" participants judged this "moral rebel" negatively, ostensibly because the rebel's behavior cast their own morality in disrepute. In contrast, mere observers tended to judge the rebel quite positively.

However, it is not entirely clear that the actors themselves perceived the rebel's behavior as *genuinely* moral and thereby worthy of admiration or emulation. Indeed, a later study (Monin et al., 2008, Study 4) showed that the actors' ratings of the morality of the rebel were quite low. One likely reason is that actors perceived the rebel's behavior as reflecting moral grandstanding and self-righteousness, rather than genuine morality. As such, they may have thought that the rebel misinterpreted the meaning of the experimental task (e.g., as expressing rather than measuring prejudice). In contrast, since the observers did not actually participate in the task, they likely did not devote as much attention or cognitive resources to discerning its various possible meanings, and so the rebel's charge of racism may have seemed more plausible to them. In sum, while intriguing, this research falls short of showing that people sometimes react negatively toward individuals that they themselves consider *genuinely* moral.

Other recent research has shown that people sometimes dislike morality in others in a different way. Melnikoff and Bailey (2018) specifically challenged the morality dominance hypothesis proposed by Landy et al. (2016), arguing that rather than being valued unconditionally, moral traits are instead valued conditionally, depending on a person's current goals.

Four studies supported this notion. In an initial study, participants playing the role of a prosecuting attorney role preferred a merciless over a merciful juror and showed greater explicit and implicit liking toward the

merciless juror – thereby showing a "preference" for the less moral person. In a second study, participants envisaged themselves in an espionage-themed game, in which they and another party would spy on one another. Participants generally regarded an honest spy as a more moral person than a dishonest spy. Yet, when choosing a spy for their own side, participants preferred to have a dishonest spy and liked this spy more than an honest spy, showing a preference for the immoral person. Only participants who chose a spy for the other side preferred the honest spy (presumably because a habitually honest spy would be less effective). A third study showed that males in a committed relationship liked a fidelitous female more than an infidelitous female, whereas this difference was attenuated among males not in a committed relationship. In a fourth and final study, participants who had behaved selfishly in a dictator game and faced the possibility of reward or punishment by a third party, preferred a third party who expressed no concern with fairness. In contrast, participants who had made fair or generous offers in the dictator game preferred third parties who were concerned with fairness.

In sum, circumstances exist in which people seem to prefer and like immoral individuals more than moral individuals, with the reason being that the immoral individuals better suit their goals. Melnikoff and Bailey interpret this evidence as a significant challenge to the morality dominance hypothesis because it shows that the valuation of morality is fundamentally conditional rather than unconditional.

This evidence constitutes an important boundary condition to the valuation of morality, and warrants an amendment to the morality dominance hypothesis (see also Landy et al., 2018). However, perhaps the most important feature of these studies is the nature of the dependent variables. Melnikoff and Bailey (2018) focus their attention on participants' *preferences* as well as various explicit and implicit *liking* measures. Most research on the importance of morality in impression formation has focused on a different variable, namely *global impressions*. Of particular note in this regard is that Melnikoff and Bailey collected other measures, the results of which are much more consistent with our (and others') theorizing on morality dominance. For instance, in two studies, in addition to collecting preference ratings, Melnikoff and Bailey also asked participants which of the two individuals they would prefer to be friends with. This is arguably more similar to an overall impression measure than are the context-specific preference and liking measures. In Study 1, both the prosecutor and the defense attorney participants indicated a significantly greater desire to be friends with the merciful as opposed to the unmerciful juror. This difference was attenuated among the prosecutors, but still significant, and a complete reversal from who they preferred as a juror. Similarly, in Study 2, participants indicated a significantly greater preference to be friends with the honest spy, which was not moderated by which spy they were choosing – a result which again departs from the preference ratings. Desire for friendship was not measured in Studies 3–4, but one would assume that it would haved shown a similar pattern.

In essence, people sometimes do appear to have situation- and role-specific "preferences" for immoral individuals, in line with Melnikoff and Bailey's (2018) "goal-conditional" account. Nonetheless, even in these cases, they retain more positive overall impressions (as assessed by the friendship measure) of moral individuals, in line with the morality dominance hypothesis. Indeed, these results corroborate a speculation we had earlier made about one possible limit to morality dominance, namely that morality may not be valued by individuals who consider themselves immoral (Landy et al., 2016, pp. 1288–1289). Melnikoff and Bailey's research substantiates this idea by showing that among individuals who themselves are pursuing immoral (or at least amoral) goals, such as infidelitous relationships, unfair allocations of resources, or partisan prosecutorial goals, immorality may be valued in others who are instrumental in serving those goals. However, since most people – even prisoners incarcerated for violent crimes – consider themselves to be highly moral (Sedikides et al., 2014), circumstances in which people consider their own goals to be immoral may be fairly rare (see also Allison et al., 1989; Brown, 2012).

Melnikoff and Bailey's (2018) data also help refine the framing of the morality dominance hypothesis. We had originally written that "positive morality traits are always positive in person perception, and negative morality traits are always negative" (Landy et al., 2016, p. 1274). This framing is clearly too broad. Instead, the existing data support two more moderate claims. First, whereas the valuation of competence and sociability traits depends on the target person's morality, moral traits are valued unconditional of the other traits that a person possesses – in what might be described as a "trait-unconditional" manner. Second, moral traits exert an almost uniformly positive impact on *global* (rather than context-specific) evaluations of people and do so to a larger extent than do either competence or sociability traits.

Another perspective on whether moral traits are ever viewed negatively comes from other research we conducted, in which we drew a distinction between "core goodness" traits that are valued unconditionally (e.g., honesty, kindness, trustworthiness), and "value commitment" traits that are valued in a conditional way (e.g., dedication, commitment, discipline; Piazza et al., 2014; see also Slote, 1983, and Gert, 2004, 2005, for earlier statements of this idea). Like competence and sociability traits, value commitment traits amplify the badness of immoral actors (and the goodness of moral actors). For example, a "dedicated Nazi" should be seen as worse than simply a "Nazi," notwithstanding that dedication is typically a positive trait. However, unlike the amplification effects arising from competence or sociability traits, which result solely from actors' increased *effectiveness* in pursuing their goals, value commitment traits amplify actors' *valuation* of their goals (in addition to their effectiveness). A "dedicated Nazi" is not just a more competent Nazi, but also a more fervent one.

Several studies corroborated this idea, using evaluations of moral character, rather than global impressions, as the primary dependent variable. When value commitment traits were attached to bad actors like terrorists or Nazis, they

worsened participants' impressions of the actors' moral character (Piazza et al., 2014, Study 1). In contrast, the same traits improved moral impressions of positive or neutral agents (though not always significantly). Interestingly, value commitment traits such as dedication and commitment worsened evaluations of immoral agents to a greater extent than did a competence trait, intelligence (Piazza et al., 2014, Study 2). Whereas competence increases the effectiveness of a bad agent, it speaks less directly to the agent's values. Corroborating this idea, mediation analyses showed that perceptions of value endorsement (to the relevant cause) rather than effectiveness mediated the effect of commitment traits on moral character evaluations for bad agents (Piazza et al., 2014, Studies 3 and 4). In sum, core goodness traits such as honesty, kindness, and trustworthiness tend to be valued positively, unconditional on the other traits a person possesses, whereas value commitment traits are valued conditionally, in that they make someone with generally bad character seem even worse. This research therefore shows that certain *types* of moral traits in others are sometimes evaluated negatively.

A critique of this research was recently published by Royzman and Hagan (2017). They argue that even "core goodness" traits can be negative in some cases. On their analysis, the reason why people consider a "dedicated Nazi" worse than a "kind Nazi" is not because of a fundamental distinction between value commitment and core goodness traits, but because people interpret these traits as having differential scope. Essentially, people interpret a "dedicated Nazi" as being dedicated to Nazi causes specifically, whereas they interpret a "kind Nazi" as being kind generally, including to victims of the Nazi regime. Royzman and Hagan showed that once this difference is equated, by stipulating that someone is kind *to Nazis*, "core goodness" traits are valued conditionally – that is, they amplify rather than diminish the negativity of bad actors. The authors concluded that there is therefore no fundamental distinction between core goodness and value commitment traits, and that the supposed distinction is an experimental artifact.

This is a perceptive analysis, but it overlooks the fact that value commitment traits tend *inherently* to have a narrower scope of application than do core goodness traits. When you learn that a person is "dedicated," it makes sense to ask *to what* they are dedicated. A person could be dedicated to all of *their own* interests to a roughly equal extent, but they cannot be equally dedicated to *all possible* human interests, since some will conflict. This is not true to the same extent for core goodness traits. When you learn that a person is "kind," it is not so natural to wonder *to whom* this person is kind. Unlike dedication, which must attach to a limited set of interests, it is entirely possible for a person to be *generally* kind, including toward people to whom they had no prior connection or attachment, or toward two people whose goals are bitterly opposed to one another. Consequently, generalized kindness is logically coherent whereas generalized dedication is not. A person who is equally dedicated to both Nazi and anti-Nazi causes is inconsistent or even nonsensical, whereas this is not true of a person who is kind to both Nazis and their opponents. In sum, because kindness

and other core goodness traits are more general than value commitment traits,[2] the difference that Royzman and Hagan explicate is best characterized not as an experimental artifact, but as a naturally occurring difference in meaning.

Overall, then, while there are some circumstances under which some kinds of moral traits can sometimes be judged negatively (depending on what kind of judgment is elicited, and which kind of trait is studied), we maintain that, when forming overall impressions of a person, people generally consider moral traits to be unambiguously positive and desirable.

2.3 Morality and Identity

In Section 2.2, we focused our attention on the relation between moral character and global impressions. We turn now to the relation between moral character and identity.

People seem to treat a person's moral character as highly central to their identity. In one study, we asked participants to indicate how central each of 80 positively valenced traits was to a person's identity. Across traits, this judgment correlated significantly with those traits' judged relevance to morality, $r(78) = 0.64$, significantly more strongly than it did with their relevance to warmth, $r(78) = 0.46$ (Goodwin et al., 2014, Study 2).

Strohminger and Nichols (2014) investigated this relationship more comprehensively. They asked participants to consider various possible changes that a person might undergo and to consider how much of the original person was still present after these changes. Their focus was on comparing moral changes with other sorts of change, including changes in nonmoral personality, autobiographical memory, desires, basic cognition, somatic states, and perceptual capacities. These changes were instantiated using a variety of specific trait descriptions, and they included both positive and negative changes (e.g., a person becoming more honest or a person becoming more evil). Across all studies, changes to a person's moral traits led to the greatest perceived identity change. Similar results have also been observed among children (Heiphetz et al., 2017; Heiphetz et al., 2018).

The same conclusion was also supported in a naturalistic context (Strohminger & Nichols, 2015). Family members of individuals suffering from neurodegenerative diseases reported the degree of identity change that their relative had undergone, as well as the extent to which their relationship with this person had deteriorated. Three diseases were studied, frontotemporal dementia (FTD), Alzheimer's disease (AD), and amyotrophic lateral sclerosis (ALS). Of these three, FTD is typically reported to induce the greatest moral changes in people, including increased antisocial behavior, increased dishonesty, and reduced empathy. In contrast, ALS is associated with motor degeneration and tends

[2] This is just as clear for other core goodness traits such as honesty or trustworthiness, which were not studied by Royzman and Hagan (2017).

to produce the fewest moral changes, with AD somewhere in the middle in terms of moral change. Consistent with the idea that moral character is fundamental to identity, family members rated FTD sufferers as having undergone the largest changes in identity, closely followed by AD sufferers, and then ALS sufferers, though ratings of the overall impact on daily functioning did not differ across the three diseases. Moral changes were also most strongly associated with family members' reports of relationship deterioration, with this link being mediated by perceived identity change.

Thus, several lines of evidence support the notion that morality is fundamental to identity, but it is not yet known why this is. In our view, the most probable explanation is that morality is important to perceptions of identity for the same functional reason that it is important in impression formation: Knowing a person's moral traits is essential to predicting how harmful or helpful they are likely to be (Goodwin et al., 2014; Landy et al., 2016). It is therefore wise to pay special attention to moral characteristics in others. This in turn may lead to the sense that morality is essential to personal identity (Strohminger & Nichols, 2014). This is not the only possible explanation, however. For instance, Strohminger and Nichols (2014) hypothesize that morality may also be important to perceptions of identity because it is a uniquely human attribute (unlike, say, memory; see also Haslam, 2006). To our knowledge, there has not yet been targeted investigation of the underlying explanatory basis for the link between morality and identity.

2.4 Inferences of Moral Character

Thus far, this review has primarily focused on how people process and use moral character information once they are in possession of it. To close, we consider the question: How do people infer moral character in the first place?

One obvious answer is that people infer character from others' behavior. Indeed, perhaps because it is so obvious, no study that we know of has focused solely on establishing this point. Nonetheless, ample evidence exists to support it. For instance, in a study of impression updating, Reeder and Coovert (1986) showed that people readily draw global moral character inferences from single instances of moral or immoral behavior. Reeder and Spores (1983) made a similar demonstration in a study on the effect of situational demands on moral character inferences. Many other studies make a similar point, but we do not review them further here given how uncontroversial this point is.

The two studies also established an effect of valence, such that immoral behaviors tend to promote stronger inferences about moral character than do moral behaviors. For instance, Reeder and Spores (1983) showed that people were more inclined to take situational constraints into account when inferring positive moral character from moral behaviors than when inferring negative moral character from immoral behaviors. Reeder and Coovert (1986) further showed that initially negative moral impressions are updated less by the

addition of new positive information than initially positive impressions are updated by the addition of new negative information. In a similar vein, other studies have shown that negative moral information is more influential on overall impressions than is positive moral information (Riskey & Birnbaum, 1974; Skowronski & Carlston, 1989). These effects likely do not reflect general negativity dominance (e.g., Baumeister et al., 2001; Rozin & Royzman, 2001), as there tends to be a positivity bias in the ability domain (Skowronski & Carlston, 1989). Instead, they likely reflect trait diagnosticity – moral traits and behaviors are generally expected in others, and so negative information is more informative, whereas the reverse is true in the ability domain (Reeder & Brewer, 1979; Skowronski & Carlston, 1989).

More recent research has contested the accepted view of negativity dominance in the moral domain, arguing that impressions of negative moral character are inherently more uncertain and therefore more labile than impressions of positive moral character (Siegel et al., 2018). The difference may in part be methodological – whereas earlier studies provided descriptions of real-life behaviors that tended to be rather extreme and rare, Siegel et al.'s (2018) studies provided real-time evidence of more moderate behaviors enacted in a laboratory context, specifically, the decision to inflict mild electrical shocks on another person for money (see Crockett et al., 2021).

People also draw inferences of moral character from other sources, chiefly a person's mental states. Just as judgments of blame hinge on intentionality (Malle et al., 2014), so too do judgments of moral character. For instance, people judge others based simply on their intention to commit various actions, even when those intentions are thwarted or not acted upon (Hirozawa et al., 2020; for related evidence on the role of intentions, see Martin & Cushman, 2015 and Martin et al., 2022, who examine "partner choice" rather than character). This is especially true for immoral rather than moral intentions (Hirozawa et al., 2020). Similarly, immoral desires, which are necessary though not sufficient components of intentions (Malle & Knobe, 1997), are taken to be indicative of poor character, at least among American Protestants (Cohen & Rozin, 2001).

Character judgments are also heavily influenced by a person's reasons for acting, such that the same act performed for different reasons can lead to very different impressions of its author. For instance, an act of aggression that is performed for calculated, self-beneficial reasons leads to more negative person inferences than the very same act if is performed reactively in response to provocation (Reeder et al., 2002). Here too, the influence of a person's reasons for acting on character judgments parallels the effect of reasons on judgments of blame (Malle et al., 2014).

Other studies demonstrate the role of deliberative processes rather than reasons per se. For instance, Critcher et al. (2013) showed that the time a person spends processing a moral decision can influence judgments of their moral character. A person who quickly rejects an opportunity to do something immoral is evaluated more positively than a person who takes their time to

arrive at the same decision. Similarly, a person who quickly takes an opportunity to do something moral is evaluated more positively than a person who takes their time to make the same decision.

The preceding studies concern mental features that typically, though not inevitably, precipitate actions – intentions, desires, reasons, and deliberations. But even mental states that occur after an action is performed can influence judgments of moral character. Gromet et al. (2016) showed that an actor who feels pleasure or indifference following an immoral act they have performed is judged to have a more negative moral character, and to be more evil, than a person who is upset following the act or whose emotional reaction was not described. Taken as a whole, these studies on the role of mental states support the view that moral character is inferred from information about a person's "moral cognitive machinery" (Critcher et al., 2020; Helzer & Critcher, 2018).

Several other peripheral variables also influence judgments of moral character. For instance, people infer good moral character (specifically, trustworthiness) based on facial structure (Willis & Todorov, 2006), facial mimcry (Bocian et al., 2018), whether a person has endured incidental suffering (Schaumberg & Mullen, 2017), and whether a person makes choices that prioritize close others at the expense of the greater good (Hughes, 2017). Thus, while the role of actions and mental states suggests that, by and large, moral character judgments have a rational basis, other research suggests that moral character inferences may be tainted by normatively irrelevant factors.

Finally, research also addresses which particular moral traits contribute most strongly to moral character judgments. Evidence points to trustworthiness as being particularly central. Trustworthiness was rated by US students as the most important characteristic for an ideal person to possess (Cottrell et al., 2007), and as the most important trait in a close friend or work partner by German students (Abele & Brack, 2013). Honesty and trustworthiness were also rated as the two most prototypic traits in a person with "good character" (Lapsley & Lasky, 2001). Compassion is also seen as quite central; several traits related to compassion were rated just below honesty and trustworthiness in prototypicality (Lapsley & Lasky, 2001). A similar emphasis on trustworthiness and, to a somewhat lesser extent, compassion, emerges from other studies that have called for participants to rate the prototypicality or necessity of various character traits for being a moral person (e.g., Aquino & Reed, 2002; Walker & Hennig, 2004; Walker & Pitts, 1998; see Landy & Uhlmann, 2018, for a review).

Beyond these "core" traits, other research suggests that loyalty (Walker & Hennig, 2004) and fairness (Lapsley & Lasky, 2001) are also considered important. Similarly, traits like being hardworking (Amos et al., 2019; Celniker et al., 2022) and self-controlled (Berman & Small, 2018; Mooijman et al., 2018) also positively influence judgments of a person's character, as do traits such as bravery (Piazza et al., 2014). Thus, we know that judgments of character are multifaceted, with many traits contributing to them. However, as of now, little is known about *how* people integrate information about multiple relevant traits

to arrive at a holistic judgment of a person's moral character. Investigating this question is an important direction for future work.

2.5 Conclusion

We have reviewed research on moral character from several different disciplines. Research on this topic has accumulated rapidly over the past 10 years, and we now have a solid basis on which to draw the following conclusions.

Moral character exists, despite earlier skepticism. While moral behavior is not *entirely* consistent from one situation to another, there is enough cross-situational consistency that one's moral character can be reliably measured, and detected by others.

Moral character is a uniquely important aspect of impression formation and person perception. People weigh moral character information more heavily than they do either competence or sociability information, which appears to be because of the uniquely important functional information that another person's moral character provides about their likely behavior toward the self. Moreover, moral character information is evaluated differently from either competence or sociability information – moral character information is valued independently of the presence of other trait information, whereas sociability and competence information is evaluated conditional on a person's morality. This does not mean that moral character information is preferred by all people under all circumstances. Indeed, recent challenges have helped refine our understanding of when and by whom moral character information might not be valued. Even in light of these challenges, however, it still seems accurate to say that moral character information is dominant in person perception.

Moral character information is also of particular relevance to judgments of personal identity. As both hypothetical and real-world data show, when a person's morality changes, they are more likely to be seen as a "different person" than when they change in other ways. The postulated reason for the prominence of moral character in identity judgments parallels that for impression formation – moral character information provides uniquely functional information for navigating the social world.

Evaluations of moral character are multifaceted, and respond to numerous kinds of information about a person. Integrating some sources of information (e.g., morally relevant behaviors and the mental states that precipitate or follow them) seems rational and normatively defensible, whereas integrating other sources (e.g., facial structure, incidental suffering) seems to reflect bias.

These conclusions are all well established by research, but they do not represent a complete picture of the role of moral character in human psychology. For instance, other research that we have not reviewed here indicates that moral character information can also play an important role in moral judgments of transgressions. This research suggests that people are not solely

focused on evaluating transgressions in isolation from their wider context. They instead appear to use information about the transgression, its eliciting circumstances, and the mental states lying behind it, to construct a mental model of the person who committed it (see, e.g., Uhlmann et al., 2015, for a review). Recent research also indicates that despite its importance in social cognition, people generally do not possess strong desires to improve their own moral character (Sun & Goodwin, 2020). There are surely many other ways that our understanding of moral character will deepen with further research.

Moral character has always been present in the world, as the case of Edgar McGregor reminds us. But, for a long time, it has been curiously absent from psychological theorizing about personality and social cognition. Pizarro and Tannenbaum's (2011) influential chapter was titled "Bringing Character Back," and it served as a call to researchers to devote more attention to the role of moral character in social and moral evaluation. Eleven years later, their call has been answered: Character is back.

References

Abele, A. E., & Brack, S. (2013). Preference for other persons' traits is dependent on the kind of social relationship. *Social Psychology*, *44*(2), 84–94.

Abele, A. E., & Wojciszke, B. (2013). The Big Two in social judgment and behavior. *Social Psychology*, *44*(2), 61–62.

Adams, R. M. (2006). *A theory of virtue*. Oxford University Press.

Allison, S. T., Messick, D. M., & Goethals, G. R. (1989). On being better but not smarter than others: The Muhammad Ali effect. *Social Cognition*, *7*(3), 275–296.

Amos, C., Zhang, L., & Read, D. (2019). Hardworking as a heuristic for moral character: Why we attribute moral values to those who work hard and its implications. *Journal of Business Ethics*, *158*(4), 1047–1062.

Annas, J. (2011). *Intelligent virtue*. Oxford University Press.

Anscombe, G. E. M. (1958). Modern moral philosophy. *Philosophy*, *33*(124), 1–19.

Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, *83*(6), 1423–1440.

Asch, S. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, *41*(3), 1230–1240.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, *11*(2), 150–166.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370.

Bergsieker, H. B., Leslie, L. M., Constantine, V. S., & Fiske, S. T. (2012). Stereotyping by omission: Eliminate the negative, accentuate the positive. *Journal of Personality and Social Psychology*, *102*(6), 1214–1238.

Berman, J. Z., & Small, D. A. (2018). Discipline and desire: On the relative importance of willpower and purity in signaling virtue. *Journal of Experimental Social Psychology*, *76*, 220–230.

Bleidorn, W., & Denissen, J. J. A. (2015). Virtues in action – The new look of character traits. *British Journal of Psychology*, *106*(4), 700–723.

Bocian, K., Baryla, W., Kulesza, W. M., Schnall, S., & Wojciszke, B. (2018). The mere liking effect: Attitudinal influences on attributions of moral character. *Journal of Experimental Social Psychology*, *79*, 9–20.

Bollich, K. L., Doris, J. M., Vazire, S., Raison, C. L., Jackson, J. J., & Mehl, M. R. (2016). Eavesdropping on character: Assessing everyday moral behaviors. *Journal of Research in Personality*, *61*, 15–21.

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology*, *41*(2), 135–143.

Brambilla, M., Sacchi, S., Rusconi, P., & Goodwin, G. P. (2021). The primacy of morality in impression development: Theory, research, and future directions. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 64, pp. 187–262). Academic Press.

Brown, J. D. (2012). Understanding the better than average effect: Motives (still) matter. *Personality and Social Psychology Bulletin*, *38*(2), 209–219.

Celniker, J., Gregory, A., Koo, H., Piff, P. K., Ditto, P. H., & Shariff, A. (2022). *The moralization of unproductive effort. Journal of Experimental Psychology: General, 152*(1), 60–79.

Cohen, A. B., & Rozin, P. (2001). Religion and the morality of mentality. *Journal of Personality and Social Psychology*, *81*(4), 697–710.

Cohen, T. R., Panter, A. T., Turan, N., Morse, L., & Kim, Y. (2013). Agreement and similarity in self-other perceptions of moral character. *Journal of Research in Personality*, *47*(6), 816–830.

Cottrell, C. A., Neuberg, S. L., & Li, N. P. (2007). What do people desire in others? A sociofunctional perspective on the importance of different valued characteristics. *Journal of Personality and Social Psychology*, *92*(2), 208–231.

Critcher, C. R., Helzer, E. G., & Tannenbaum, D. (2020). Moral character evaluation: Testing another's moral-cognitive machinery. *Journal of Experimental Social Psychology*, *87*, Article 103906.

Critcher, C. R., Inbar, Y., & Pizarro, D. A. (2013). How quick decisions illuminate moral character. *Social Psychological and Personality Science*, *4*(3), 308–315.

Crockett, M. J., Everett, J. A. C., Gill, M., & Siegel, J. Z. (2021). The relational logic of moral inference. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 64, pp. 1–64). Academic Press.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631–648.

Darley, J. M., & Latane, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*(4), 377–383.

Davis, D. E., Hook, H. N., Worthington Jr. E. L., Van Tongeren, D. R., Gartner, A. L., Jennings II, D. J., & Emmons, R. A. (2011). Relational humility: Conceptualizing and measuring humility as a personality judgment. *Journal of Personality Assessment*, *93*(3), 225–234.

Doris, J. M. (1998). Persons, situations, and virtue ethics. *Noûs*, *32*(4), 504–530.

Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. Cambridge University Press.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902.

Fleeson, W., Furr, R. M., Jayawickreme, E., Helzer, E. G., Hartley, A. G., & Meidl, P. (2015). Personality science and the foundations of character. In C. B. Miller, R. M. Furr, A. Knobel, & W. Fleeson (Eds.), *Character: New directions from philosophy, psychology, and theology* (pp. 41–71). Oxford University Press.

Fleeson, W., Furr, R. M., Jayawickreme, E., Meindl, P., & Helzer, E. G. (2014). Character: The prospects for a personality-based perspective on morality. *Social and Personality Psychology Compass*, *8*(4), 178–191.

Foot, P. (1978). *Virtues and vices*. University of California Press.

Gert, B. (2004). *Common morality: Deciding what to do*. Oxford University Press.

Gert, B. (2005). *Morality: Its nature and justification* (revised ed.). Oxford University Press.

Gert, B., & Gert, J. (2020). The definition of morality. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. https://plato.stanford.edu/archives/fall2020/entries/morality-definition/

Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological Science*, *24*(1), 38–44.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148–168.

Gromet, D. M., Goodwin, G. P., & Goodman, R. A. (2016). Pleasure from another's pain: The influence of a target's hedonic states on attributions of immorality and evil. *Personality and Social Psychology Bulletin*, *42*(8), 1077–1091.

Harman, G. (1999). Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society*, *99*, 315–331.

Hartley, A. G., Furr, R. M., Helzer, E. G., Jayawickreme, E., Velasquez, K. R., & Fleeson, W. (2016). Morality's centrality to liking, respecting, and understanding others. *Social Psychological and Personality Science*, *7*(7), 648–657.

Hartshorne, H., & May, M. A. (1928). *Studies in the nature of character: Vol. 1. Studies in deceit*. Macmillan.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*(3), 252–264.

Heiphetz, L., Strohminger, N., Gelman, S., & Young, L. (2018). Who am I? The role of moral beliefs in children's and adults' understanding of identity. *Journal of Experimental Social Psychology*, *78*, 210–219.

Heiphetz, L., Strohminger, N., & Young, L. (2017). The role of moral beliefs, memories, and preferences in representations of identity. *Cognitive Science*, *41*(3), 744–767.

Helzer, E. G., & Critcher, C. R. (2018). What do we evaluate when we evaluate moral character? In K. Gray & J. Graham (Eds.), *Atlas of moral psychology* (pp. 99–107). Guilford Press.

Helzer, E. G., Furr, R. M., Hawkins, A., Barranti, M., Blackie, L. E., & Fleeson, W. (2014). Agreement on the perception of moral character. *Personality and Social Psychology Bulletin*, *40*(12), 1698–1710.

Helzer, E. G., Jayawickreme, E., & Furr, R. M. (2018). Moral character: Current insights and future directions. In V. Zeigler-Still & T. K. Shackelford (Eds.), *The SAGE handbook of personality and individual differences: Vol. 2. Origins of individual and personality differences* (pp. 278–300). SAGE Publications.

Hirozawa, P. Y., Karasawa, M., & Matsuo, A. (2020). Intention matters to make you (im)moral: Positive-negative asymmetry in moral character evaluations. *Journal of Social Psychology*, *160*(4), 401–415.

Hughes, J. S. (2017). In a moral dilemma, choose the one you love: Impartial actors are seen as less moral than partial ones. *British Journal of Social Psychology*, *56*(3), 561–577.

Hursthouse, R. (1999). *On virtue ethics*. Oxford University Press.

Hursthouse, R., & Pettigrove, G. (2018). Virtue ethics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. https://plato.stanford.edu/archives/win2018/entries/ethics-virtue/

Judd, C., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*(6), 899–913.

Kervyn, N., Bergsieker, H. B., & Fiske, S. T. (2012). The innuendo effect: Hearing the positive but inferring the negative. *Journal of Experimental Social Psychology*, *48*(1), 77–85.

Kupperman, J. J. (1991). *Character.* Oxford University Press.

Landy, J., Piazza, J., & Goodwin, G. P. (2016). When it's bad to be friendly and smart: The desirability of sociability and competence depends on morality. *Personality and Social Psychology Bulletin*, *42*(9), 1272–1290.

Landy, J. F., Piazza, J., & Goodwin, G. P. (2018). Morality traits still dominate in forming impressions of others. *Proceedings of the National Academy of Sciences*, *115*(25), Article E5636.

Landy, J. F., & Uhlmann, E. L. (2018). Morality is personal. In K. Gray & J. Graham (Eds.), *Atlas of moral psychology* (pp. 121–132). Guilford Press.

Lapsley, D. K., & Lasky, B. (2001). Prototypic moral character. *Identity*, *1*(4), 345–363.

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research*, *39*(2), 329–358.

Louden, R. B. (1984). On some vices of virtue ethics. *American Philosophical Quarterly*, *21*(3), 227–236.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, *25*(2), 147–186.

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, *33*(2), 101–121.

Martin, J. W., & Cushman, F. (2015). To punish or to leave: Distinct cognitive processes underlie partner control and partner choice behaviors. *PLoS ONE*, *10*(4), Article e0125193.

Martin, J. W., Leddy, K., Young, L., & McAuliffe, K. (2022). An earlier role for intent in children's partner choice versus punishment. *Journal of Experimental Psychology: General*, *151*(3), 597–612.

McCullough, M. E., Emmons, R. A., & Tsang, J.-A. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology*, *82*(1), 112–127.

McDowell, J. (1979). Virtue and reason. *The Monist*, *62*(3), 331–350.

Meindl, P., Jayawickreme, E., Furr, R. M., & Fleeson, W. (2015). A foundation beam for studying morality from a personological point of view: Are individual differences in moral behaviors and thoughts consistent? *Journal of Research in Personality*, *59*, 81–92.

Melnikoff, D. E., & Bailey, A. H. (2018). Preferences for moral vs. immoral traits are conditional. *Proceedings of the National Academy of Sciences*, *115*(4), E592–E600.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, *67*(4), 371–378.

Milgram, S. (1974). *Obedience to authority: An experimental view*. Harper & Row.

Miller, C. (2003). Social psychology and virtue ethics. *Journal of Ethics*, *7*(4), 365–392.

Mooijman, M., Meindl, P., Oyserman, D., Monterosso, J., Dehghani, M., Doris, J. M., & Graham, J. (2018). Resisting temptation for the good of the group: Binding moral values and the moralization of self-control. *Journal of Personality and Social Psychology*, *115*(3), 585–599.

Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology*, *95*(1), 76–93.

Moshagen, M., Hilbig, B. E., & Zettler, I. (2018). The dark core of personality. *Psychological Review*, *125*(5), 656–688.

Nicholson, I. (1998). Gordon Allport, character, and the "culture of personality," 1897–1937. *History of Psychology*, *1*(1), 52–68.

Peabody, D. (1984). Personality dimensions through trait inferences. *Journal of Personality and Social Psychology*, *46*(2), 384–403.

Piazza, J., Goodwin, G. P., & Rozin, P, & Royzman, E. B. (2014). When a virtue is not a virtue: Conditional virtues in moral evaluation. *Social Cognition*, *32*(6), 528–558.

Pizarro, D. A., & Tannebaum, D. (2011). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 91–108). American Psychological Association.

Reeder, G. D., & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review*, *86*(1), 61–79.

Reeder, G. D., & Coovert, M. D. (1986). Revising an impression of morality. *Social Cognition*, *4*(1), 1–17.

Reeder, G. D., Kumar, S., Hesson-McInnis, M. S., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology*, *83*(4), 789–803.

Reeder, G. D., & Spores, J. M. (1983). The attribution of morality. *Journal of Personality and Social Psychology*, *44*(4), 736–745.

Riskey, D. R., & Birnbaum, M. H. (1974). Compensatory effects in moral judgment: Two rights don't make up for a wrong. *Journal of Experimental Psychology*, *103*(1), 171–173.

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*(4), 283–294.

Rosenberg, S., & Olshan, K. (1970). Evaluative and descriptive aspects in personality perception. *Journal of Personality and Social Psychology*, *16*(4), 619–626.

Royzman, E. B. & Hagan, J. P. (2017). The shadow and the tree: Inference and transformation of content in psychology of moral judgment. In J.-F. Bonnefon & B. Tremoliere (Eds.), *Moral inferences* (pp. 56–74). Routledge.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320.

Schaumberg, R. L., & Mullen, E. (2017). From incidental harms to moral elevation: The positive effect of experiencing unintentional, uncontrollable, and unavoidable harms on perceived moral character. *Journal of Experimental Social Psychology*, *73*, 86–96.

Sedikides, C., Meek, R., Alicke, M. D., & Taylor, S. (2014). Behind bars but above the bar: Prisoners consider themselves more prosocial than non-prisoners. *British Journal of Social Psychology*, *53*(2), 396–403.

Siegel, J. Z., Mathys, C., Rutledge, R. B., & Crockett, M. J. (2018). Beliefs about bad people are volatile. *Nature Human Behaviour*, *2*(10), 750–756.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*(1), 131–142.

Slote, M. A. (1983). *Goods and virtues.* Clarendon.

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, *141*(1), 159–171.

Strohminger, N., & Nichols, S. (2015). Neurodegeneration and identity. *Psychological Science*, *26*(9), 1468–1479.

Sun, J., & Goodwin, G. P. (2020). Do people want to be more moral? *Psychological Science*, *31*(3), 243–257.

Swanton, C. (2003). *Virtue ethics: A pluralistic view*. Oxford University Press.

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, *10*(1), 72–81.

Walker, L. J., & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology*, *86*(4), 629–647.

Walker, L. J., & Pitts, R. C. (1998). Naturalistic conceptions of moral maturity. *Developmental Psychology*, *34*(3), 403–419.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598.

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin*, *24*(12), 1251–1263.

Wojciszke, B., Dowhyluk, M., & Jaworski, M. (1998). Moral and competence-related traits: How do they differ? *Polish Psychological Bulletin*, *29*(4), 283–294.

Wood, A. M., Maltby, J., Steward, N., & Joseph, S. (2008). Conceptualizing gratitude and appreciation as a unitary personality trait. *Personality and Individual Differences*, *44*(3), 621–632.