The Dynamic Moral Self:
A Social Psychological Perspective

Benoît Monin
Alexander H. Jordan
Stanford University

Chapter to appear in D. Narvaez & D. Lapsley (Eds.),
Moral Self, Identity and Character: Prospects for a New Field of Study,
Cambridge University Press.
The Dynamic Moral Self: A Social Psychological Perspective

When psychologists explore the role of the self in moral motivation and behavior, they typically take a personological approach. Some seek to describe a general personality structure shared by widely recognized moral exemplars, whereas others examine individual differences in the extent to which morality is central to one’s personal goals. A social-psychological approach to the moral self complements these personological perspectives by taking into account the situational malleability of moral self-regard, or one’s self-perceived moral standing at any given moment. Recent research reviewed in this chapter demonstrates the value added by this perspective: First, when people are made secure about their morality, they sometimes act less morally (moral credentials); second, moral exemplars are disliked rather than admired when their behavior is seen as an indictment of people’s own choices (moral resentment); and third, people sometimes boost their moral self-regard to compensate for failures in other domains (moral compensation). These phenomena underscore the importance of understanding moral self-regard as just one aspect of a highly dynamic self-concept.

The Self in Moral Psychology

For decades, moral psychology mostly left the self out of its analyses. It focused instead on moral reasoning and on the cognitive underpinnings of decisions about right and wrong (e.g., Kohlberg, 1969). The neglect of the self and emphasis on the mechanics of moral reasoning might have been a reaction against the perceived murkiness of psychodynamic theories influential at the time, and the dearth of empirical support for concepts such as “superego strength” to explain moral learning (see Kohlberg, 1963). In addition, psychologists in the moral reasoning tradition wanted to isolate that which was unique to moral thought and behavior, thus
excluding general psychological elements (e.g., “ego controls”) that were shared with non-moral behavior and cognition (see Kohlberg & Candee, 1984). Whatever its causes, the choice to ignore the self paid off at first, leading to a rigorously narrow focus making possible intensive research programs on the foundation of moral reasoning and its development.

Recently, however, moral psychology has begun to move beyond moral reasoning and to pay more attention to the moral self. Seminal work by Augusto Blasi (1983, 2004) suggested that the centrality of morality to one’s self-identity might be a major factor influencing the strength of the link between moral judgment and moral action, and thus a major influence on the production of moral behavior. Moral psychologists have eagerly embraced the project of investigating the moral self, but in the enthusiasm typical of nascent research areas, agreeing on a precise definition of the moral self has taken a back seat to establishing the value of this new approach. Exciting new discoveries have shown that moral psychology can be enriched by adding the moral self to its models (as this volume demonstrates), but a closer look reveals that investigators vary greatly in their definition of this construct. In the first half of this chapter we describe and distinguish among the different approaches to moral self, identity, and character prevalent in the existing literature, and argue that these approaches have emphasized stable personality structures; in the second half, we propose a new, more social psychological understanding of the moral self that further enhances our understanding of moral thought and behavior.

**Personological Approaches to the Moral Self**

The approaches to moral self and identity prevalent in the existing moral psychology literature are for the most part personological, focusing on stable individual differences between adults—even if these differences are assumed to result from differences in developmental
changes over the formative years. We group these approaches into two main categories: In the *moral personality* tradition, researchers aim to identify the general personality structure shared by individuals widely recognized as moral; in the *moral centrality* tradition, researchers focus on measures indexing how much one cares about being moral or the extent to which one moralizes life.

*The Moral Personality Approach*

One influential approach to the moral self seeks to map out the personality structure (the non-moral substrate) that gives rise to extraordinary moral commitments or achievements. Just as one may talk about the personality type associated with politicians or with artists, so there is, according to this approach, a moral personality type associated with morally accomplished individuals. Identifying the signature personality profile of the moral individual could, in turn, guide moral education by determining which traits should be fostered in the general population to promote moral behavior. Two research methods have dominated this approach to moral identity: the exemplar method and the prototype method. The exemplar method seeks to identify the personality traits shared by real-life moral leaders; the prototype method consists in asking individuals which traits they associate with morality. We discuss each of these in turn.

*The moral exemplar method.* The exemplar method begins by identifying individuals who are recognized by their community as moral, and then investigating whether these individuals differ from the rest of the population on personality variables, including non-moral ones. For example, Walker and Frimer (2007; this volume) found that moral exemplars (50 Canadians who had received medals for exceptional bravery or caring) were both more agentic and more communal and had more secure attachments compared to matched controls, supporting their
proposal that the moral self is associated with what they call (2007, p.845) a “personological core.”

Colby and Damon’s (1992) intensive study of 23 moral exemplars reported in *Some Do Care* shares elements of this approach. Through in-depth interviews, Colby and Damon identified distinctive non-moral personality features shared by individuals considered moral exemplars by their communities, such as a singular disregard of risk, positivity in the face of discouraging circumstances, open-mindedness to the ideas of those around them, and a desire for personal growth. Furthermore, some of the exemplars’ good deeds are described as habitual and non-reflective, as if they arose ineluctably from a personality structure rather than from self-conscious, deliberative moral reasoning. However, because Colby and Damon interpret these habitual moral choices as resulting from the merging of moral goals and personal goals, rather than as the non-moral expression of a given personality structure, we return to them below when we discuss another understanding of moral identity, to wit, moral self-importance.

A personality approach also characterizes Hart and colleagues’ studies of “care exemplars” among inner-city youths, selected through nomination by church groups, schools, or social agencies for their community work or civic engagement (Hart, Yates, Fegley, & Wilson, 1995). Hart and colleagues found that care exemplars described themselves using slightly more moral personality traits (e.g., honest, moral, trustworthy) than individuals in a matched control group, but sharper differences emerged in the way the exemplars structured their self-concept. They were more likely to incorporate parents and ideals in their self-concept, to include in it systematic beliefs and life plans, and to connect it to past and future selves, thus giving it a greater sense of continuity through time. Typical of the moral personality approach, Hart and
colleagues identified non-moral components of the self-concept associated with exemplary moral behavior. Thus the exemplar approach need not rely exclusively on a trait approach to personality, but can involve, as this example demonstrates, investigating more complex elements of the self-concept of recognized exemplars.

The moral prototype method. Rather than tapping into the personality structure of actual individuals, as in the exemplar method, the prototype method involves identifying which traits or attributes are included in people’s mental representations of a prototypically moral individual, or, in other words, their “naturalistic conceptions of moral maturity” (Walker & Pitts, 1998). Research using this method has identified a set of personality traits (e.g., caring, friendly, honest) generally seen as constitutive of moral individuals (Aquino & Reed, 2002, Pilot Studies 1 and 2; Lapsley & Lasky, 2001, Studies 1 and 2), and has shown that this prototype serves to organize new information, leading, for example, to false recognition of novel words associated with the activated prototype (Lapsley & Lasky, 2001, Studies 3 and 4). Walker and Hennig (2002) further documented the existence of multiple, distinct prototypes of a moral person (e.g., brave, just, caring), corresponding to different personality patterns. This plurality of prototypes resonates with the diversity of actual moral icons, be they historical, religious, or contemporary (e.g., George Washington, Martin Luther King Jr., or the Dalai Lama).

Whereas the exemplar method focuses on the unique personality features of actual exemplars, the prototype method documents the multifaceted representation of the moral ideal in a given culture. These two methods address related questions insofar as individuals are likely recognized by their community as moral leaders (and thus studied by students of the exemplar
approach) to the extent that they match the templates associated with the moral ideal in their culture (as documented by proponents of the prototype approach).

The Moral Centrality Approach

Besides the moral personality approach, a second school of research on the moral self focuses on the degree to which people think of themselves, or the world more generally, in moral terms. This moral centrality approach comes in two shades: moral self-importance, or the extent to which individuals care about being moral, and moral chronicity, or the extent to which individuals encode the social world in moral terms.

Moral self-importance. Moral self-importance refers to how much one wants to be moral. One can place great importance on being moral without necessarily feeling moral (which is instead the hallmark of the moral self-regard approach that we propose below) or being recognized by the community as a moral leader (as in the moral exemplar approach above). For individuals high in moral self-importance, morality is a critical life goal and an important yardstick for measuring how well they are doing in the world. According to Colby and Damon (1992), moral exemplars stand out in the extent to which they are committed to moral values, so much so that their personal and moral goals have become indistinguishable; satisfying moral goals has gained a primary status as a basic need, explaining why exemplars happily ignore other typical basic needs such as security or comfort if this gives them a chance to move toward their more central moral goals.

If moral exemplars’ exceptionally high level of moral self-importance is one factor helping to produce their moral behavior, might more ordinary individuals’ differing levels of moral self-importance be an important predictor of the translation of moral judgment into moral
action? Blasi (1983) argued exactly this, using the term “moral identity” (p. 201) to refer to the
extent to which being moral and acting morally are parts of one’s essential self, and explicitly
discussing this as a dimension of inter-individual difference.

More recently, Aquino and Reed (2002), following Blasi’s lead, developed a self-report
scale to measure the “self-importance of moral identity”—commonly abbreviated to “moral
identity” by the authors themselves (e.g., Reed & Aquino, 2003; Reed, Aquino & Levy, 2007).
This scale comprises two subscales, one that taps into whether it is personally important for the
individual to possess moral qualities (internalization), and one that taps into how much the
individual believes that his or her day-to-day activities communicate morality to others
(symbolization). Of note given our prior discussion of the moral personality, Aquino and Reed’s
scales do not ask people explicitly whether they want to be moral, but instead whether they want
to possess a set of characteristics associated with the moral prototype (caring, compassionate,
fair, friendly, generous, helpful, hardworking, honest, and kind).

Parallel to this work, the virtue component of the Contingencies of Self-Worth scale
developed by Crocker and Wolfe (2001) captures how much, for any given individual, virtuous
behavior contributes to a general sense of self-worth, alongside other non-moral sources such as
physical appearance, academic performance, and God’s love. Indeed, one way to think of the
self-importance approach to the moral self is that it conceives of morality as a contingency,
something that individuals aspire to and depend on to feel good about themselves.

Moral chronicity. An individual whose self-image depends on being moral must be
attuned to opportunities in the environment for achieving morality, or at least have a clear sense
that some actions make oneself more moral than others. Moral chronicity (Lapsley & Narvaez,
Monin & Jordan

A Social Psychological Approach to the Moral Self: Dynamic Moral Self-Regard

In contrast to the personological approaches presented so far, we propose a view of the self that is more reflective and more labile—one’s moment-to-moment answer to the question “How moral am I?” This dimension of the self-concept we call *moral self-regard*. This construct is not entirely unrelated to the personological approaches to the moral self reviewed above: Individuals feel secure in their moral self-regard, for example, when they feel confident that they
possess the traits identified in the moral prototype approach; furthermore, individuals high in moral self-importance likely attend more chronically to their moral self-regard, and this brand of self-regard may have a greater impact on their global self-worth than it does for individuals who place less general importance on being moral. Individuals may also show stable differences in their average moral self-regard. But we contend that as with the self-concept more generally (see Markus & Nurius, 1987), people’s thoughts and behavior are often guided by a “working” level of moral self-regard that fluctuates from moment to moment according to situational influences. Note that instead of just saying, in line with social-cognitive approaches, that the interaction between one’s self-concept and the situation needs to be taken into account to best predict behavior, we contend that situations actually can affect aspects of the self-concept and can therefore influence behavior through this mediator, rather than moderate the link between self and behavior.

Social psychology can complement the stable individual differences approaches documented above by examining precisely the ways in which one’s moral self-regard is sensitive to current circumstances, recent feedback, social comparison, or other situational factors. Whereas some authors have pitted situational forces against moral character (e.g., Doris, 2002), we see the two as intricately related: Moral self-regard is by definition sensitive to situational demands, and conversely when situations do influence behavior, they often do so by affecting moral self-regard. Moral self-regard is one part of a multifaceted and dynamic self-concept that is deeply motivational. People strive to preserve and enhance a positive self-concept; in the moral domain, people may do so behaviorally by actually being a good person, or cognitively by biasing their construal of the world, favoring self-flattering information, or putting down
threatening others. Moreover, since moral self-regard is just one aspect of a self that is artful in compensating for deficiencies, feelings of moral inadequacy may sometimes be remedied by dwelling on another valued dimension, and, conversely, feeling morally superior may sometimes serve to make up for more practical failings.

We now turn to three new lines of our own research that illustrate the value of this social psychological approach to the moral self. This research, we believe, documents the role of moral self-regard in everyday morality and demonstrates how adding this dynamic self-concept to the study of moral experience broadens the scope of phenomena that can be studied. The first line of research (moral credentials) shows that bolstering people’s moral self-regard can liberate them to act in more morally problematic ways in the future. The second (moral resentment) shows that when people’s moral self-regard is threatened by the behavior of heroes or saints, people may come to resent these superior others, even though by all accounts their behavior is exemplary. The third (moral compensation) goes beyond securing or threatening individuals’ moral self-regard and shows instead that individuals may boost their own moral self-regard (and put down others’ morality) to compensate for a threat to their self-concept in a non-moral domain.

Moral Credentials: The Licensing Effect of Secure Moral Self-Regard

One consequence of considering the role of moral self-regard in moral behavior is that it forces one to think of moral choices as a sequence rather than in temporal isolation. Moral and immoral actions occur in the context of prior moral and immoral actions, and moral self-regard provides a connecting thread between these instances. If what stops someone from engaging in immoral behavior is the fear of feeling immoral, then having engaged in unambiguously moral behavior beforehand could put that fear to rest, and, paradoxically, license later immoral
behavior. Along these lines, Mordecai Nisan’s (1991) moral balance model proposed that when individuals decide whether to engage in a morally relevant action, they do so in the context of previous such actions, keeping track of a moral balance that can incur credits (good deeds) and debits (bad deeds). Their goal is not to attain moral perfection, but merely to retain a reasonable level of moral self-regard, allowing for fluctuations as long as one remains above an unacceptable level that would clearly denote immorality. Individuals might thus feel that they can “afford” to engage in morally problematic behaviors given their accumulated moral credit—or they may feel less compelled to engage in moral behavior if they feel they have already done their good deed for the day. In support of this model, Nisan showed, for example, that participants encouraged to imagine that they had returned a found wallet in a first instance declared that they might be less likely to do so in a second instance.

Monin and Miller’s (2001) work on moral credentials documented a related phenomenon by actually providing participants with the opportunity to engage in moral behavior on one relevant dimension (prejudice), and then showing that this token demonstration liberated participants to display an otherwise problematic response later on. For example, participants given the opportunity to select a stellar African American applicant in a first job-selection task (establishing non-racist credentials) were more likely than participants in a control condition (who picked from an all-White applicant pool) to express that a second, unrelated job in a racist police force would be “better suited” for a White person. This second task was crafted to make it attractive for participants to favor a White person, but given discomfort about engaging in a behavior that feels unethical in a prejudice-conscious society, participants did not express this preference unless they had been secured in their non-racist self-image by the first, liberating
choice—a *moral credential*. Note that the effect obtains even when the experiment is set up so that its two parts allegedly go to two different experimenters (i.e., audiences), suggesting that the effect is driven by internal concerns about retaining an egalitarian moral self-regard and not just by a desire for effective self-presentation to others.

Thus, consistent with Nisan’s moral balance model, Monin and Miller’s social psychological experiments show that a prior moral act (even a token one) can license later ethically questionable behavior. An important theoretical point of contrast, however, is that whereas in the moral balance model the actor fully recognizes the immorality of the second act (but feels that he or she can afford the debit), the moral credentials model proposes that the first behavior serves as a cue with which to disambiguate the second behavior. In Monin and Miller’s situations, the second decision is ambiguous in that both legitimate and illegitimate motives could explain the same behavior (favoring a White police officer because a Black officer would be uncomfortable [vs. incompetent] in a racist police force), and the role of credentials is to disambiguate the behavior in a positive way. A resolution of the apparent tension between the two models may be that the moral balance model best applies in unambiguous cases, where the meaning of the behavior is unaffected by prior action but one’s moral self-image is debited, whereas moral credentials might be at work in more ambiguous cases, where the later behavior is interpreted as harmless because of one’s history, and moral self-image is unaffected. The predictions for behavior, however, are the same: In both cases, these two models support the importance of considering a dynamic self-concept by arguing that one’s moral self-regard based on recent past behavior affects future morally relevant actions.
Moral Resentment: When Others Threaten Moral Self-Regard

A counterintuitive prediction that arises from incorporating a dynamic view of the self-concept into the study of morality is that highly moral others may not always receive the respect and appreciation that they deserve. In particular, individuals whose superior moral choices call into question the morality of others are likely to draw hostility. Moral rebels—individuals who take a principled stand against a morally problematic situation—may, for example, be respected and liked by uninvolved observers, but resented by individuals who were in the same situation but did not rebel, because the rebel’s stance implies that it was wrong to remain passive. The reflexive defense of one’s fragile moral self-regard may, then, sometimes block people from learning from and being inspired by the behavior of morally exemplary peers (Monin, 2007).

In one experiment, a target individual who refused to complete a racist task was appreciated by neutral judges, but disliked by participants who had been asked to complete the racist task prior to the target—and who overwhelmingly did so without complaining (Monin, Sawyer, & Marquez, 2008). To participants who had willingly gone along with the problematic behavior, the otherwise exemplary stance of the rebel apparently represented a threat, which they addressed by putting him or her down, and reporting less respect for and attraction to the rebel. Demonstrating the role of the self-concept, participants whose self-concept had been secured before seeing the rebel (by reflecting on an important quality or value and how they recently demonstrated it) did not show the same backlash, even if they had done the racist task first. In fact, participants thus “self-affirmed” (Steele, 1988; Sherman & Cohen, 2006) were able to learn from the rebel’s gesture: They admitted having more freedom at the time of the task, and reported less comfort with their own choice.
This work directly supports the dynamic view of moral self-regard that we introduce in this chapter. First, a threat to individuals’ moral self-regard motivates them to derogate others. Second, the fact that this effect can be eliminated by affirming one’s general (typically non-moral) skills or values demonstrates the connection between moral self-regard and other aspects of the self-concept that contribute to self-worth, allowing for cross-domain compensatory processes. Being reminded that one is a great tennis player may offset the discomfort of having just acted in an ethically dubious way, and prevent a knee-jerk backlash against others who acted better, allowing personal growth.

*Moral Compensation: Moralization to Protect the Self*

The fungibility of sources of self-worth, with non-moral fixes addressing moral threats, raises the possibility that moral fixes can also serve to address non-moral threats. If moral self-regard is just one component of a multi-faceted self-concept, and one of the many sources that contribute to feeling like an adequate person, then it is possible that feeling moral can serve to compensate for feeling less adequate in other, non-moral domains. In particular, we propose that when a person demonstrates to individuals that their choices may not be as expedient as they could have been (e.g., when they have helped with a tedious task, followed a silly procedure, or respected an absurd rule, and then witness someone avoid this inconvenience without incurring a cost), they may retrospectively justify what they did by attributing it to their greater morality, and put down others as less moral for choosing shortcuts.

In a recent study of this phenomenon (Jordan & Monin, 2008), participants who completed a boring task for the experimenter—repetitive number-writing—and then saw a confederate quit the same task, without any negative consequence befalling him, elevated their
ratings of their own morality and castigated the morality of the confederate, compared to participants who simply completed the boring task (without seeing the confederate quit it) or who simply observed the confederate quit the task (without first having completed it themselves). Furthermore, self-affirming an important value or trait, as above, wiped out the phenomenon, suggesting once more that the effect was a response to self-threat. The conjunction of completing an exceedingly tedious task and seeing another person effortlessly avoid that same task apparently constituted a threat to participants’ general self-worth as rational, efficacious agents, and they compensated by boosting their moral self-regard and dimming their view of the other’s morality. In this case, feeling moral served mostly to feel a little less foolish. We see once more the considerable interplay between moral self-regard and other aspects of the dynamic self-concept.

Conclusion

This chapter had two main goals: In the first half, we outlined and provided a taxonomy of the existing personological approaches to the study of the moral self. This rapidly expanding area of research is richly textured, offering varied perspectives on moral self and identity. By sketching some broad distinctions between the moral personality and moral centrality approaches, our hope was not to further compartmentalize the field, but instead to dispel potential misunderstandings when similar terms are used to refer to different concepts, and thus to facilitate future discussion and collaborations by clarifying useful distinctions and complementarities. The second half of the chapter focused on developing a social psychological approach to the moral self—what we called moral self-regard—incorporating a view of the self-concept as dynamic, influenced by subtle situational cues and sometimes seemingly trivial.
behavior, and part of a global sense of self-worth (including non-moral sources) that individuals are strongly motivated to preserve. We presented three lines of research (moral credentials, moral resentment, and moral compensation) that we believe directly contribute to our understanding of the moral self, but do not fit within models that conceptualize the moral self as a personality structure or as moral centrality. All three lines of research speak to the moment-to-moment malleability (and behavior-generating power) of moral self-regard, even in response to minimal manipulations, supporting our claim that dynamic moral self-regard is an important construct to explore further if we wish to understand the moral self.

As the work in this volume demonstrates, the study of the moral self is rife with fascinating questions likely to keep researchers busy for years to come, and holds the promise of casting light on some of the thorniest issues in moral psychology. By carving out some of the emerging directions, and staking a claim for social psychologists in one of these promising avenues, we hope to foster continuing dialogue between the various areas of psychology eager to study moral phenomena. Rather than wanting to appropriate for social psychology one dimension of the endeavor, we hope that the proposal expressed in this chapter will be perceived for what it is, to wit, a commitment by social psychologists to do their share and contribute their expertise in the collective effort to better understand the role of the moral self in explaining everyday morality.


Author Note

Benoît Monin and Alexander H. Jordan, Department of Psychology, Stanford University. The authors wish to thank Elizabeth Mullen and the Stanford Moral Psychology Reading Group, as well as the editors of this volume, for their very helpful feedback on a previous version of this manuscript.

Correspondence concerning this chapter should be addressed to Benoît Monin, Department of Psychology, Jordan Hall, Stanford University, Stanford, CA 94305, or via electronic mail to monin@stanford.edu.