# Moral Monism and Pluralism: Origins, Implications, and Debates

**Authors:** Zhang Haotian, Yu Feng, Xu Liying, Metaphysical Philosophy, Yu Feng, Xu Liying

**Date:** 2024-03-19T00:00:00+00:00

## Abstract

The debate between monism and pluralism constitutes one of the most active theoretical confrontations in the field of moral psychology in recent years. Moral monism contends that all external moral phenomena and internal moral structures can be explained by a single factor. Moral pluralism, conversely, maintains that morality cannot be explained by a single factor alone, but rather that there exist multiple qualitatively distinct moral dimensions that exhibit cultural sensitivity. Moral Dyad Theory and Moral Foundations Theory are typical representatives of this controversy. The two sides have engaged in debates concerning the explanatory power of harm, the phenomenon of moral dumbfounding, modular morality, and the independence of the purity dimension, among other aspects. Future research should further investigate the monism versus pluralism puzzle in morality from three specific aspects, thereby sustaining theoretical vitality in the field of moral psychology.

## Full Text

## Preamble

### Monism and Pluralism in Morality: Origins, Connotations and Debates

ZHANG Haotian[1], YU Feng[1], XU Liying[2], XUAN Zheli[1]
([1] Department of Psychology, School of Philosophy, Wuhan University, Wuhan 430072, China)
([2] School of Marxism, Tsinghua University, Beijing 100084, China)

**Abstract:** The debate between monism and pluralism represents one of the most active theoretical confrontations in moral psychology in recent years. Moral monism posits that all observable moral phenomena and underlying moral structures can be explained by a single factor. Moral pluralism,

conversely, argues that morality cannot be reduced to a single factor, but instead comprises multiple heterogeneous moral dimensions with cultural sensitivity. Moral Foundations Theory and Dyadic Morality Theory serve as typical representatives of this controversy. The two sides have engaged in debates concerning the explanatory power of harm, the phenomenon of moral dumbfounding, modular morality, and the independence of the purity dimension. Future research should further explore the monism-pluralism dilemma in morality through three specific avenues, thereby sustaining theoretical vitality in the field of moral psychology.

**Keywords:** monism, pluralism, moral foundations theory, dyadic morality theory, Haidt-Gray debates

## 1. Introduction

"Emotion or reason?" "Objective or relative?" "Nature or nurture?" The field of moral psychology is replete with controversies and competing theories. In recent years, the most active theoretical confrontation has been the debate between moral monism and moral pluralism (Graham et al., 2018). The divergence between monism and pluralism was already evident in early moral philosophical thought. Aristotle' s virtue ethics proposed that the highest human good is the "Mesotês" (the mean). The virtue of the mean unifies intellect and ethics, the relative and the absolute, with actions conforming to the mean being virtuous and deviations constituting vice. The ultimate goal of the mean is to achieve social justice and fairness (Aristotle, 2003). John Rawls similarly argued that the core of morality is justice, aimed at ensuring citizens' rights and the fairness of social cooperation (Rawls, 1971). This represents monistic thought, which views morality as a single dimension capable of explaining all external moral phenomena and internal moral essence. Contemporary representatives of monism in moral psychology include the stages theory of moral development (Kohlberg, 1969) and dyadic morality theory (Schein & Gray, 2018b), among others.

In contrast, pluralism maintains that morality possesses multiple dimensions that manifest across different sociocultural contexts. Sociological pioneer Émile Durkheim argued that morality comprises three fundamental elements: the spirit of discipline, the spirit of sacrifice, and the spirit of autonomy (Durkheim, 1925/1961). Contemporary representatives of pluralism in moral psychology include triadic moral discourse theory (Shweder et al., 1997), relational models theory (Rai & Fiske, 2011), and moral foundations theory (Graham et al., 2018), among others.

The debate between monism and pluralism involves theoretical discussions at the ontological level of moral psychology: Can morality be reduced to a single dimension, or does it encompass multiple qualitatively distinct moral values? This question holds significant importance for understanding the nature of morality. Moreover, from the perspective of theoretical construction, should good theoretical models aim to encompass as many moral elements as possible to explain

complex moral phenomena, or should they adhere to Occam' s Razor principle ( "entities should not be multiplied beyond necessity" ) in pursuit of theoretical parsimony? The discussion of "monism versus pluralism"also holds crucial value for this question.

Among the various theories mentioned above, Kurt Gray' s dyadic morality theory and Jonathan Haidt' s moral foundations theory represent typical representatives of this controversy. Over the past five years, both sides have published numerous theoretical and empirical articles exploring whether morality is monistic or pluralistic (e.g., Graham, 2015; Gray, Waytz, et al., 2012; Gray & Keeney, 2015a, 2015b; Schein & Gray, 2015), and their theoretical exchange continues to this day without resolution. This paper aims to provide a comprehensive and systematic summary of the debate between moral monism and pluralism, with particular focus on the recent active exchange between dyadic morality theory and moral foundations theory. Future research should further advance exploration of the nature of morality by clarifying moral structures, integrating the respective advantages of monistic and pluralistic research to construct more comprehensive integrated models, and strengthening attention to "non-WEIRD" cultural groups, particularly indigenous Chinese culture.

### 2.1.1 Definitions and Representative Theories

Moral monism posits that all surface-level moral phenomena and internal moral structures can be explained by a single factor (Beal, 2020). Moral psychology pioneer Lawrence Kohlberg argued that morality can be reduced to a single factor—reasoning about justice—and that moral development has a clear endpoint: becoming a moral thinker who advocates for fairness and equality (Kohlberg, 1969). Additionally, Gray and colleagues contend that the plurality of moral content is merely a superficial manifestation, with the core factor being singular: perceived harm (Gray, Waytz, et al., 2012).

Among theories of moral monism, the most active and representative in recent years is Gray et al.' s dyadic morality theory. This theory argues that the diversity of moral attitudes and beliefs can be completely explained by perceived harm (Schein & Gray, 2018b). In the process of moral judgment, there exists a "cognitive template" whereby immoral behavior is a process in which an intentional moral agent inflicts harm upon a vulnerable moral patient, with causation linking the three elements—intentional agent, suffering patient, and causation—all of which are essential for perceiving harm (Schein & Gray, 2018b). A moral agent refers to a moral actor who forms subjective judgments through cognition and reasoning, can implement moral behaviors according to their intentions, and can be held responsible—commonly understood as the "perpetrator" in social terms. The moral patient, conversely, is the passive recipient of the moral action, commonly understood as the "victim" (Gray & Wegner, 2009). Importantly, the prerequisite for perceiving harm and making immoral judgments is that a causal relationship must exist between the intentional moral agent who commits the harmful act and the victimized moral patient. The clearer this

causal relationship, the more apparent the harm becomes. When causal relationships are ambiguous, as in the "side-effect effect" commonly seen in moral dilemmas, the severity of moral judgments diminishes (Schein & Gray, 2018b).

Furthermore, "perceived harm" requires clarification due to its rich connotations. First, Gray and colleagues argue that "harm" is not objective or external. For instance, some behaviors considered immoral (such as copulating with a frozen chicken or consensual incest between siblings with protection) may not cause direct, objective physical harm or property damage, yet are still perceived as posing potential harm to the body and social norms (e.g., Royzman et al., 2015; Schein & Gray, 2016). Second, research indicates that people can perceive and categorize harm within extremely short timeframes, suggesting that harm perception is intuitive rather than rational (Schein & Gray, 2018b). Finally, harm manifests in diverse forms, including not only common physical or emotional harm but also deception-based harm, group-based harm, disrespect-based harm, and spiritual harm (Hartman et al., 2022). Although these manifestations appear to support "pluralism,"dyadic morality theory argues that all moral judgments can ultimately be reduced to a single factor: perceived harm (Schein & Gray, 2018a).

How is perceived harm realized through cognitive mechanisms? Gray, Young, et al. (2012) propose that it relies on mind perception—the perception of whether entities such as humans and objects possess mental capacities—which also constitutes the essence of morality. Using a "bottom-up approach" driven by data, Gray et al. (2007) employed factor analysis to divide dimensions of mind perception into two components: agency and experience. Agency concerns action and thought, corresponding to moral responsibility, while experience concerns perception and emotion, corresponding to moral rights. Moral agents are typically perceived as having greater agency than experience, whereas moral patients show the opposite pattern (Gray & Wegner, 2009). For example, mentally ill individuals who violate the law are often not held criminally responsible, partly because they lack mental capacity and thus cannot serve as moral agents. It should be noted that these two dimensions are not moral dimensions themselves but rather individual perceptions of mind used to explain the "cognitive template" in moral judgment. Dyadic morality theory posits only one monistic dimension of morality: perceived harm.

### 2.1.2 Supporting Evidence

The important role of harm in moral judgment is undeniable. Gray and colleagues have attempted to demonstrate that perceived harm is the primary—indeed, the sole—influencing factor in moral judgment, rather than other factors previously considered important in moral psychology. Taking disgust as an example, Gray argues that it cannot easily become a cause of moralization because many disgusting behaviors are not immoral, such as a fly landing in rice (Gray & Schein, 2016). In Schein, Ritter, et al.' s (2016) study, researchers asked participants to rate the moral wrongness, harmfulness, and disgusting-

ness of 24 disgust-related behaviors, including disgusting sexual acts (e.g., oral sex), pathogen-related disgusting behaviors (e.g., seeing cockroaches scurrying on the floor), and moral violations (e.g., deceiving a friend). Results showed that perceived harm best distinguished between disgusting immoral behaviors and merely disgusting behaviors. In two additional studies, researchers asked participants to evaluate the immorality of same-sex marriage and blasphemy. Although disgust directly predicted immorality, perceived harm fully mediated this relationship, indicating that perceived harm is the most powerful factor explaining immoral judgments, with no direct causal relationship between disgust and immoral judgments when harm is present (Schein, Ritter, et al., 2016).

Beyond directly arguing from "perceived harm" itself, Gray and colleagues have conducted a series of experiments demonstrating that "purity" morality is not substantially different from perceived harm. Researchers asked participants to rate the immorality and harmfulness of three types of behaviors: harm violations (e.g., murder), harmless acts (e.g., riding a bus), and purity violations (e.g., smearing feces on a Bible). Results showed that when forced to respond quickly, perceptions of harm were higher, and even in purity moral scenarios without objective harm, participants still intuitively believed these behaviors caused varying degrees of harm. These findings indicate that perceiving immorality as harmful occurs intuitively, without deliberate rationalization (Gray et al., 2014).

Furthermore, in the political domain, liberals and conservatives have different moral psychological foundations. For instance, Graham et al. (2009) found that liberals value individualizing moral foundations such as "care"and "fairness"more highly, while conservatives value binding foundations such as "loyalty," "authority," and "purity" more highly. However, Gray and colleagues argue that no insurmountable moral gulf exists between different political orientations. Instead, their research shows that both liberals and conservatives share the same essential moral concept: concern for harm, with differences only in how they perceive harm (Landy, 2016; Schein & Gray, 2015). Across seven studies, researchers found that both conservative- and liberal-leaning participants first thought of harm-related events when recalling immoral incidents (Experiment 1) and considered harm the most important factor in immoral events (Experiments 2 and 3). Additionally, through implicit testing measures (reaction times), they found that harm was closely linked to immoral events (Experiment 4). Researchers also asked participants to evaluate ten moral scenarios commonly used in moral foundations theory, finding high correlations among different moral foundations (Experiment 5) and that harm was the strongest predictor of immoral events (Experiment 6). Even previously identified purity violation events (e.g., partner infidelity) were more closely linked to harm (Experiment 7) (Schein & Gray, 2015).

Based on this evidence, Gray and colleagues argue that harm is the most powerful factor explaining moral judgments, and that other scholars' proposed moral foundations such as purity, fairness, and loyalty can all be explained through perceived harm (Schein & Gray, 2015). Different political groups' preferences

for moral foundations essentially represent different interpretations of harm.

### 2.2.1 Definitions and Representative Theories

Moral pluralism holds that morality cannot be explained by a single factor but instead encompasses numerous qualitatively distinct dimensions that manifest across different cultures and social strata, exhibiting cultural sensitivity (Graham et al., 2013, 2018). In response to Kohlberg' s monistic moral view, early pluralism proponents Carol Gilligan and Elliot Turiel argued that the core of morality lies not only in pursuing justice but also in caring for others (Gilligan, 1988) and avoiding harm (Turiel, 1983). Cultural anthropologist Richard Shweder was among the first scholars to systematically propose moral pluralism. Based on fieldwork in India, Shweder and colleagues criticized previous monistic moral views as merely reflecting individualistic moral perspectives prevalent in Western secular culture, while neglecting numerous moral domains that might exist in other cultures. Grounded in a global cultural context, Shweder proposed the more pluralistic "triadic moral discourse theory," which suggests that observed moral domains can be primarily divided into three types: first, "the ethic of autonomy," where people rely on concepts such as harm, rights, and justice to protect individual autonomy, with anger as the corresponding moral emotion; second, "the ethic of community," where people rely on concepts such as duty, respect, and loyalty to maintain institutional and social order, with contempt as the corresponding moral emotion; and third, "the ethic of divinity," where people rely on concepts such as sanctity and sin to protect the inherent divinity within each person and maintain personal purity and sacred morality, with disgust as the corresponding moral emotion (Shweder et al., 1997; Rozin et al., 1999).

Additionally, cultural anthropologist Alan Fiske critiques monistic morality. Based on his earlier "relational models theory," Fiske views morality as "relationship regulation" that maintains cooperative social relationships with others, identifying four distinct foundational moral motivations: "unity," "hierarchy," "equality," and "proportionality." Unity involves avoiding or eliminating threats from negative events by maintaining ingroup integrity; hierarchy is the motivation to respect superiors in social groups; equality refers to maintaining equal welfare and voice within communities; and proportionality means that rewards and punishments must be proportional to merit, and benefits proportional to contribution (Rai & Fiske, 2011).

Beyond these moral theories, the most representative pluralistic theory in recent years is moral foundations theory (Graham et al., 2013, 2018), which aims to explain the diversity and relativity of moral judgments. Moral foundations theory advocates nativism, cultural learning, intuitionism, and pluralism. Nativism holds that humans have evolved innate moral foundations that are continuously refined through postnatal experience to form mature moral values. Cultural learning means that innate moral foundations develop through cultural evolution. Intuitionism means that moral judgments are based on intuition rather

than deliberation. Finally, moral foundations theory advocates pluralism because humans have faced numerous adaptive social challenges throughout evolutionary history, leading to the formation of different moral foundations when addressing these challenges.

To date, most research has focused on five moral foundations. The first two are individualizing foundations that emphasize protection of individual rights. The first is "Care/harm" : during evolution, caregivers who were more sensitive to children' s needs and suffering had offspring more likely to survive. This moral foundation entails compassion and care for oneself and others, and the avoidance of pain and harm. The second is "Fairness/cheating" : all social animals have sociality, and to avoid lose-lose situations, fairness and reciprocal altruism are needed to achieve common interests, requiring monitoring of others' behavior and reputation to form social norms that punish cheaters.

The latter three are binding foundations that emphasize values constraining and building social groups. The first is "Loyalty/betrayal" : in competition for social resources, the most cohesive alliances tend to defeat less cohesive rival alliances, so loyalty to the ingroup (e.g., patriotism) is also a socially prescribed virtue. The second is "Authority/subversion" : groups with power are respected and seen as legitimate, typically functioning better than leaderless or normless groups. People who disrespect authority or tradition often trigger anxiety about disorder and contempt for disruptors. The third is "Purity/degradation" : this moral foundation evolved in humans to avoid disease infection, accompanied by disgust emotions specific to pathogens. In contemporary times, the purity foundation has also generalized to social life, including rejecting outgroups and excluding those who do not adhere to group sacred customs.

Although most research has concentrated on these five moral foundations, many others may exist, such as liberty/oppression, honesty/lying, etc. (Graham et al., 2018). However, despite multiple revisions to moral foundations theory, the five foundations mentioned above remain the most widely recognized.

### 2.2.2 Supporting Evidence

Moral pluralism is supported by rich empirical research. Based on fieldwork in India, Shweder and colleagues first identified basic moral elements (e.g., male superiority) through interview studies, then used hierarchical cluster analysis to extract three factors from 25 basic elements: the ethics of autonomy, divinity, and community (Menon & Shweder, 1994). Rozin et al. (1999) subsequently proposed the CAD (contempt, anger, disgust) model to identify moral emotions corresponding to the three ethics, asking participants to match moral scenarios from each ethic with emotional images or words and make corresponding emotional expressions. Results showed that the ethic of autonomy corresponded more to anger, the ethic of community to contempt, and the ethic of divinity to disgust, once again validating the plurality of moral values from an emotional perspective.

Furthermore, fellow cultural anthropologist Alan Fiske conducted fieldwork in traditional African tribal cultures, studying human cooperation and planning from a social relations perspective, and proposed four basic relationship types: communal sharing, authority ranking, equality matching, and market pricing (Fiske, 1992). These types were confirmed through confirmatory factor analysis (Haslam & Fiske, 1999) and validated for cross-cultural universality across four countries including China and Korea (Fiske, 1993). Building on this, Rai and Fiske (2011) further proposed the relational norms model, arguing that the purpose of morality should be maintaining effective social relationships. Empirical research has found that people rely on these four different motivations to make moral judgments across domains including consumer behavior (McGraw et al., 2012), organizational behavior (Giessner & Van Quaquebeke, 2010), and justice (Copes et al., 2013), demonstrating the plurality of moral structure and causation.

Haidt and colleagues constructed moral foundations theory based on previous research including moral discourse theory and relational models theory, drawing on wisdom from cultural anthropology, evolutionary psychology, and sociology, and adopting a top-down theoretical construction approach. Initially, this work was primarily theoretical speculation without empirical research exploring why morality could be specified into five dimensions (Haidt, 2012). Only later did researchers use questionnaire development and qualitative interviews to validate the foundations. Graham et al. (2011) developed the Moral Foundations Questionnaire, with results showing good internal, external, and practical validity. Confirmatory factor analysis demonstrated good fit for the five-factor structure model, providing empirical support for the foundations (Graham et al., 2011). Cross-cultural research showed that Eastern groups paid more attention to loyalty and purity foundations than Western groups, with no significant differences in attention to the other three foundations, and that the relationship between political orientation and moral foundations did not vary across cultures, providing empirical support for the cross-cultural universality of moral pluralism (Graham et al., 2011). Subsequent studies have also demonstrated good cross-cultural stability of the Moral Foundations Questionnaire (e.g., Doğruyol et al., 2019; Yalçındağ et al., 2019). All these studies indicate that different moral foundations emerge in groups with different characteristics, and that this moral diversity possesses cross-cultural stability, thereby supporting moral pluralism.

Additionally, Haidt and colleagues have found that people with different political orientations emphasize different moral foundations. McAdams et al. (2008) conducted qualitative research, conducting life narrative interviews with 128 adults deeply involved in religious and political activities and coding their responses according to the five moral foundations. Results showed that conservatives emphasized respect for social hierarchy, ingroup loyalty, and purity or sanctity foundations more strongly, while liberals placed greater emphasis on care and fairness foundations. Graham et al. (2009) used four different methods to find that liberals valued individualizing foundations more than conservatives, while conservatives attended to all five moral foundations and valued binding

foundations more than liberals. Recent meta-analytic research supports these fundamental differences between conservatives and liberals (Kivikangas et al., 2021), demonstrating that not only do multiple moral foundations exist, but their distribution also varies by social class and political orientation, further proving the plurality of moral foundations.

## 3. The Debate Between Moral Monism and Pluralism

In the early development of moral psychology, both monistic theories represented by Kohlberg and pluralistic theories represented by Shweder, Fiske, and others existed in parallel without theoretical dialogue. Only with the later emergence of moral foundations theory and dyadic morality theory did monism and pluralism truly engage in theoretical confrontation. In 2012, Gray and colleagues published a target article in the prominent theoretical review journal *Psychological Inquiry* titled "Mind Perception is the Essence of Morality" (Gray, Young, et al., 2012), formally proposing the monistic claim that all external moral phenomena and internal concepts can be explained by perceived harm. Haidt and colleagues published commentary articles in the same issue, acknowledging the theoretical contributions of dyadic morality theory while refuting the monistic claim (Graham & Iyer, 2012; Koleva & Haidt, 2012). In subsequent years, both sides have engaged in extensive and profound discussions on whether morality is monistic or pluralistic. The following sections summarize the main points of contention and commentary from both sides regarding moral dumbfounding, modularity versus constructivism, and the independence of the purity moral dimension.

### 3.1 Monism' s Rebuttal to Pluralism

#### 3.1.1 The Phenomenon of Moral Dumbfounding is Not Harmless

Although some moral scenarios have no connection to objective harm—such as copulating with a frozen chicken or using a national flag to clean a toilet—people still intuitively judge these behaviors as immoral. When asked for specific reasons, people cannot provide adequate justification for their judgments, a phenomenon known as "moral dumbfounding" (Haidt et al., 2000), which serves as primary evidence for Haidt' s intuitionist moral view. The core explanation involves mechanisms of "purity violations" or "loyalty violations" that are unrelated to harm (Scott et al., 2016), and these moral foundations represent important concepts driving the development from monism to pluralism (Gray et al., 2021). However, Gray and colleagues deny that these behaviors are completely harmless, arguing that Haidt narrowly equates perceived harm with objective harm (Guglielmo, 2018). So-called independent mechanisms such as "purity" and "loyalty" can all be explained through perceived harm. Additionally, Gray argues that the moral scenarios selected by Haidt and colleagues are too "weird," limiting the explanatory power of moral dumbfounding to some extent.

Several empirical studies have examined the "harmlessness" of moral dumbfound-

ing scenarios, suggesting they are not harmless. Royzman and colleagues (2015) replicated the traditional moral dumbfounding paradigm. Unlike previous studies, they measured whether participants truly believed these behaviors were harmless, asking questions such as: "How credible do you think it is that Julie and Mark (who are siblings) could have sexual relations without damaging their relationship? How credible is it that there would be no negative consequences?" Results showed that participants insisted incest would produce negative consequences and real-world harm. In subsequent interviews, even when researchers repeatedly explained that incest was harmless, participants remained skeptical, with 57.1% believing incest caused emotional harm to Julie and Mark. Another study also demonstrated that harm fully predicted participants' immoral judgments of moral dumbfounding phenomena, even if only "potential harm" that had not occurred (Stanley et al., 2019). Based on this evidence, Gray argues that so-called "harmless wrongs" actually involve varying degrees of harm.

Moreover, regarding the "weirdness"of moral dumbfounding scenarios, Gray and colleagues criticize that the scenarios involved are too bizarre. Behaviors such as copulating with frozen chickens, eating dead dogs, selling one's soul, or rolling in urine clearly differ from everyday moral dilemmas, and these strange scenarios do not adequately reveal the nature of moral cognition. Moral development is meant to guide our daily interpersonal interactions and promote cooperation, a point that Haidt et al.' s research clearly deviates from (Gray & Keeney, 2015a). In one study (Gray & Keeney, 2015b), participants evaluated moral scenarios commonly used in moral foundations theory, assessing them on five dimensions: severity, weirdness, wrongness, harmfulness, and impurity. Participants rated purity violation scenarios as weirder and less severe than harm scenarios. Interestingly, in harm scenarios, participants not only rated harmfulness higher but also rated impurity higher than in purity violation scenarios. Researchers then asked participants to list their own examples of purity and harm scenarios, finding that naturally nominated scenarios were more realistic, such as intentional homicide and marital infidelity. Comparing natural scenarios with moral foundations scenarios revealed that participants considered moral foundations scenarios weirder and less harmful and severe than natural scenarios. In summary, natural scenarios better predicted harm and purity morality than moral foundations scenarios, and harm had greater explanatory power for purity violation scenarios than purity itself.

### 3.1.2 Rebuttal to Moral Modularity

Another point of monists' rebuttal to pluralists concerns the modularity hypothesis. Gray argues that moral foundations theory explicitly endorses modularity, whereby different psychological characteristics correspond to different cognitive modules with distinctness and specificity (Schein, Hester, et al., 2016). Moral foundation modules are "switches" in animal brains: when a specific stimulus (e.g., a moral violation) activates a specific module (e.g., the purity foundation), it leads to certain behaviors or experiences (e.g., purity-related moral judgments)

(Haidt, 2012, p. 123) and elicits specific emotions corresponding to the module.

However, Gray contends that the modular explanation of moral foundations theory lacks persuasiveness because a single behavior may activate multiple modules (e.g., marital infidelity may simultaneously activate loyalty and harm foundations), and different moral foundations show similarly high overlap. Schein and Gray (2015) asked participants to rate ten moral scenarios according to moral foundations, finding that all moral foundations had an internal consistency coefficient as high as 0.89. Additionally, data from the Moral Foundations Questionnaire showed high correlations between different foundations, such as 0.88 between authority and loyalty, 0.72 between harm and fairness, and 0.80 between purity and authority (Graham et al., 2011). These high correlations themselves cannot demonstrate that the sub-dimensions of moral foundations theory have high discriminant validity, especially when correlations between different foundations exceed within-foundation consistency (e.g., the correlation between loyalty and authority is higher than loyalty' s consistency with itself).

Furthermore, Gray opposes Haidt et al.' s claim of moral emotion modularity (Haidt, 2012; Rozin et al., 1999), arguing that a moral foundation does not elicit a specific emotion module but rather simultaneously triggers multiple emotional experiences. For instance, harm violations may simultaneously evoke anger and disgust (Gray et al., 2017), and researchers have found that purity violations elicit more smiling responses and angry expressions rather than the disgust emotion proposed by Haidt, Shweder, and others (Franchin et al., 2019). In response, Gray and colleagues maintain a "constructivism of moral emotions" (Cameron et al., 2015), arguing that moral emotions are not elicited through different moral modules but are generated through combinations of three basic components—norm violation, negative affect, and perceived harm—possessing domain-generality. Certain moral behaviors produce different moral emotions, such as anger or disgust, because they trigger these three factors (Gray et al., 2017).

### 3.1.3 Purity is Not Sufficient to be an Independent Moral Foundation

Whether purity can serve as an independent moral dimension is currently the most contentious issue between monists and pluralists (Gray et al., 2021). Gray et al. (2021) systematically reviewed 158 studies on purity morality, finding almost no evidence supporting purity as an independent psychological construct or as a unique mechanism of moral judgment. The definition of purity morality itself is overly complex, with at least nine different definitions involving pathogen avoidance, maintaining natural order, etc. Bibliometric analysis indicates considerable heterogeneity between purity' s definition and actual measurement. Among 135 articles measuring purity morality, only 19 (14%) achieved consistency between definition and measurement, and this disconnect is detrimental to theoretical development (Gray, 2017; Gray et al., 2021). In contrast, the concept of harm is more concise: damage to body or mind (Haslam, 2016), and actual measurement is typically consistent with definition.

Moreover, in multiple empirical studies, purity and harm moral foundations show no clear discriminant validity (Gray et al., 2014; Gray et al., in press; Gray & Keeney, 2015b; Schein & Gray, 2015). Whether evaluating moral judgments of moral foundations theory scenarios or scenarios nominated by experimental participants, purity judgments and harm judgments show significant high correlations: rs $\geq$ .87 (Gray & Keeney, 2015b, study 1). Even some scenarios previously identified as purity violations can elicit intuitive perceptions of harm (Gray et al., 2014; Gray & Keeney, 2015b). Other research shows that impurity scenarios are more likely to be perceived as harmful, whereas harm scenarios are rarely perceived as impure (Chakroff et al., 2017), and harm is more closely linked to moral judgments than purity (Rosenfeld & Tomiyama, 2021).

Finally, Gray et al. (2021) argue that behaviors defined as pure by moral psychologists may also represent abstract moral directives about harm—harm that may not involve direct physical damage but constitutes higher-level harm. Specifically, "assaulting others" represents first-order harms that are immoral, while more abstract higher-order moral principles, such as certain dietary habits or religious rituals (e.g., not eating pork), relate to a group' s cultural presuppositions. Research on child socialization shows that children may initially use the ethic of autonomy (e.g., stealing toys harms other children) to explain theft, but through processes of cultural-religious socialization may later use the ethic of divinity (e.g., theft defiles God' s will) to explain immoral behavior (DiBianca Fasoli, 2018). Purity norms emerge later in childhood and are not as universal as first-order harms, and this socialization process is built upon a foundation of harm.

## 3.2 Pluralism' s Rebuttal to Monism

### 3.2.1 Monism Oversimplifies Moral Connotations

Haidt et al. (2015) argue that dyadic morality theory oversimplifies moral phenomena and concepts, and that complex moral phenomena cannot be reduced to harm alone. They sarcastically suggest that Gray has wielded "Occam' s Razor" as "Occam' s Chainsaw" (Koleva & Haidt, 2012), cutting away important moral dimensions such as purity and fairness. Haidt also argues that although Gray emphasizes moral monism, he simultaneously insists on "harm pluralism" in his definition of harm, with different types of harm—physical, interpersonal, psychological, etc.—merely representing different discursive expressions of other moral dimensions. Thus, monistic research has inevitably moved toward moral pluralism (Haidt et al., 2015).

Second, Haidt contends that monists overemphasize and exaggerate harm' s role in moral judgment, rendering their theory unfalsifiable (Koleva & Haidt, 2012). For example, when people make immoral judgments about clearly harmful events (e.g., murder, assault), harm can be considered the primary basis for moral judgment. However, for harmless scenarios (e.g., consensual sibling sex with protection), monists explain them as involving "perceived harm" or "po-

tential harm," making the concept of "harm" overly subjective and broad. This allows all immoral phenomena to be explained in a self-consistent manner, making the theory unfalsifiable and insufficient as a rigorous scientific proposition (Haidt et al., 2015).

Furthermore, Haidt et al. (2015) identify numerous methodological problems in monism research. Regarding the relationship between political orientation and moral foundations, Schein and Gray (2015) claimed that liberals and conservatives show no differences in moral cognition. However, their measurement of political orientation had significant problems: although they used a 7-point Likert scale (1 = very liberal to 7 = very conservative), they incorrectly dichotomized the scale, categorizing individuals with neutral political orientations (those selecting 4 on the scale) as conservatives, leading to unreliable results. Therefore, their failure to find differences in moral foundations between groups lacks sufficient persuasiveness. Additionally, Schein and Gray (2015) only examined the connection between harm-related vocabulary and moral judgments without measuring other moral dimensions such as fairness or loyalty, so they cannot claim that harm is the sole factor explaining moral judgments.

### 3.2.2 Defense of Moral Modularity

In response to monists' rebuttals of moral foundations theory' s modularity hypothesis, Haidt et al. (2015) argue that this represents a misunderstanding of the theory. Moral foundations theory does not propose five "Fodorian modules" (i.e., independent functional modules responsible for specific psychological processes) (Fodor, 1983) but instead employs the more flexible and overlapping concept of "massive modularity" with "functional specialization." These modules process specific stimuli and information inputs to adapt to the environment. The modules are not completely sealed but continuously developing: humans first possess a "first draft" of the moral mind that then develops through specific cultural learning.

Graham et al. (2018) further explain that moral foundations theory has never claimed that moral judgments consist of five completely distinct modules. Moral development begins with some innate knowledge— "original triggers" (e.g., care, fairness)—that exists in learning modules, making it easy to learn certain things and difficult to learn others. Through postnatal cultural learning, individuals develop many new modules under the influence of "current triggers"(e.g., specific cultures), forming different moral foundations. These modules are then linked to specific emotional and motivational responses. Therefore, complex moral structures are constructed and continuously developing, and cannot be explained by five completely independent modules, let alone by a single module (e.g., perceived harm).

### 3.2.3 Evidence that Purity Can Be an Independent Moral Dimension

Confronting challenges from moral monists, Haidt argues that purity as an independent moral dimension has broad interdisciplinary evidence and deep theoretical connotations, integrating wisdom from religious studies, cultural anthropology, evolutionary biology, and evolutionary psychology. Years ago, Haidt et al. (1993) used self-constructed non-harm scenarios (using a national flag to clean a toilet, copulating with a dead chicken, etc.) to find that higher-social-class college students considered these stories matters of social convention or personal preference, while most lower-social-class students made harsher moral judgments about these behaviors, with emotional responses predicting moral judgments better than harm assessments. Some experimental studies on embodied morality similarly found that physical disgust from impurity increased the severity of moral judgments (Wheatley & Haidt, 2005). Based on this, Haidt argues that impurity disgust can explain people' s moral condemnation of non-harm scenarios. Since then, empirical research on purity morality has grown continuously. Graham et al. (2018) summarized evidence for purity as an independent moral foundation in a recent theoretical article, elaborating on its essential differences from harm.

First, harm and purity morality differ essentially at the cognitive level. For example, frequency and importance influence harm judgments more than purity judgments (Rottman & Young, 2019). The thinking styles involved in purity and harm also differ: activating or training analytic thinking can increase individuals' attention to purity foundations, while activating or training holistic thinking increases attention to harm foundations (Alper & Yilmaz, 2020). People consider purity violations more morally wrong when they themselves are victims, but consider harm violations more morally wrong when others are victims (Dungan et al., 2017). Compared to harm judgments, impurity judgments condemn the act itself less and the actor more (Uhlmann & Zhu, 2014). Additionally, intention influences purity less than harm judgments (Young & Saxe, 2011); accidental versus intentional harm produces different activation in the right temporoparietal junction, whereas accidental versus intentional purity violations show no such difference (Chakroff et al., 2017). These different types of judgments are also associated with different facial micro-expressions and neural systems (Wasserman et al., 2017).

Second, the purity moral foundation can predict important psychological phenomena and behaviors that harm cannot. For example, purity uniquely predicts issues such as same-sex marriage and euthanasia (Koleva et al., 2012), whereas other moral foundations and political ideologies cannot. Purity also predicts vaccine hesitancy debates (Amin et al., 2017), pro-environmental attitudes (Rottman et al., 2015), and social distance in real-world social networks (Dehghani et al., 2016). Attention to impurity predicts moral judgments of suicide better than harm judgments (Rottman et al., 2014). Purity judgments predict disgust sensitivity better than judgments related to any other moral foundation (Wagemans et al., 2018), even in harm scenarios. Additionally, research

on moralization processes shows that the process of moralization is mediated by disgust rather than anger or harm when using "moral shock" methods (e.g., increasing moral conviction about abortion through images of aborted fetuses), thereby refuting dyadic morality theory (Wisneski & Skitka, 2017).

## 4. Summary and Outlook

The divergence between monism and pluralism was already evident in early moral philosophical systems, traceable to the disagreement between "moral relativism" (the view that no universal moral principles exist, and any moral value applies only within specific cultural boundaries and personal value systems) and "moral universalism" (a monistic ethical stance holding that universal ethics applicable to all people exist). In recent years, however, this theoretical confrontation has entered the purview of moral psychology (Beal, 2020). Table 1 briefly summarizes the consensus and points of contention in previous research on moral monism and pluralism. Overall, moral foundations theory strives to depict the full picture of human morality across cultures, whereas dyadic morality theory is more parsimonious, exemplifying Occam' s Razor and possessing theoretical elegance. Yet the two perspectives are not fundamentally different. As Haidt acknowledges that "Care/harm" is the most central dimension in moral foundations theory (Haidt et al., 2015), Gray also acknowledges three of the four basic tenets of moral foundations theory (nativism, cultural learning, intuitionism) (Gray & Keeney, 2015a). Their primary disagreement concerns whether "harm" can explain all external moral phenomena and internal moral attitudes, cognitions, and emotional processes. We propose that future research on the monism-pluralism debate in morality can proceed along three specific lines.

First, morality requires a clearer definition, which perhaps even ethics cannot achieve (MacIntyre, 1957). Whether morality is monistic or pluralistic involves the specification of morality' s essence, and many moral philosophers even argue that morality may have no essence (Taylor, 1978). If morality has an essence but it cannot be identified, the problem is unsolvable. If morality has no essence, it may only have meaning in linguistic usage or cognition, making psychological operational definitions or average layperson opinions the essence of morality. This creates a gap between "expert insights" and "folk theory" : researchers' and participants' understanding of concepts may differ, such that what researchers consider "harmless" may not be considered harmless by participants (De Villiers-Botha, 2020; Royzman et al., 2015). More problematically, in actual research, the theoretical definitions researchers adopt are often inconsistent with their operational definitions. For example, as mentioned above, only 14% of studies on purity morality achieved consistency between theoretical and operational definitions (Gray et al., 2021). Therefore, to ensure the validity of moral research, future studies should define moral concepts more clearly and ensure that only one construct is tested at a time. Because moral phenomena are complex and multifaceted, some commonly used moral scenarios in research do not correspond to specific moral foundations. For example, the scenario "publicly scolding one'

s parents" (Graham et al., 2011) simultaneously involves both "Care/harm" and "Authority/subversion" dimensions, and adultery simultaneously involves harm, purity, and loyalty morality. If scenarios do not clearly reflect a specific moral dimension, researchers cannot precisely identify which moral dimension influences people' s moral judgments.

Second, beyond focusing on the monism-pluralism controversy itself, we must also step back from the dispute and objectively examine the different characteristics of these two moral orientations, learning from each other' s strengths to advance the overall development of moral psychological research. Specifically, monism emphasizes the simplicity of moral constructs and measurement precision, while pluralism highlights how different cultural groups understand the nature of morality. These represent two different approaches to theoretical construction and perspectives on moral essence. However, monism cannot explain all moral scenarios, such as harms from omissions and fair competition scenarios involving harm (e.g., boxing matches), which are difficult to explain through dyadic morality theory (Piazza et al., 2019; Royzman & Borislow, 2022). Similarly, pluralism research has difficulty replicating stable and consistent five moral foundations across different social classes and cultural groups (e.g., Curry et al., 2019; Harper & Rhodes, 2021). Future research should combine the advantages of monistic and pluralistic approaches, particularly considering how different situations and cultures moderate understanding of moral essence, and employ ecologically valid real-world scenarios whenever possible (Schein, 2020) to construct more comprehensive integrated theoretical models. Just as monists have recently incorporated emotional factors into harm-centered moral judgment models to reconcile the opposition between dyadic morality theory and the social intuitionist model (Gray et al., in press), pluralists should also continue to clarify the essential differences between different moral foundations, simplify existing theories according to Occam' s Razor principle, and explore whether moral foundations exist that are qualitatively different from previous research, thereby constructing more parsimonious and precise theoretical models.

Finally, although moral pluralism emphasizes cultural sensitivity and moral monism emphasizes cultural universality, the vast majority of empirical research is dominated by "WEIRD" (Western, Educated, Industrialized, Rich, and Democratic) samples (Hester & Gray, 2020; Matsuo & Brown, in press). When facing "non-WEIRD" samples, moral foundations theory and dyadic morality theory often lack explanatory power, such as the existence of previously undiscovered moral dimensions (e.g., Atari et al., 2020; Willard & McNamara, 2019) and different correspondence patterns between political orientation and moral foundations (e.g., Davis et al., 2016; Frimer, 2020). Corresponding moral measurement tools also suffer from low construct validity and inadequate cross-cultural reliability and validity (e.g., Iurino & Saucier, 2020; Zakharin & Bates, 2021). Future research should urgently consider the possibility of conducting moral research in different cultures, particularly exploring the nature of morality in indigenous Chinese culture. Chinese philosophy contains numerous discussions on the nature of morality, and compared to the Western pursuit of truth, Chi-

nese people more ultimately pursue morality as the pursuit of goodness (Jin Guantao, 2021; Yu Feng, 2021). Confucius advocated the "Three Bonds and Five Constants," essentially viewing the monistic "ren" (benevolence) as the ultimate moral realm (Lao Siguang, 2019). Mencius' s so-called "four beginnings" may appear pluralistic, yet all originate from the monistic moral psychological and emotional foundation of "the heart that cannot bear to see others suffer" (Li Zehou, 2021). Similarly, Song Dynasty Neo-Confucian Zhu Xi' s advocacy of "preserving heavenly principle" and Ming Dynasty scholar Wang Yangming' s proposal of "extending innate knowledge" also represent monistic moral views abstracted from the "four beginnings" (Li Jingde, 2020). Thus, Chinese philosophy from Confucius to Wang Yangming exhibits characteristics of moral monism. However, Chinese people particularly emphasize holistic and dialectical thinking, believing that all relative and opposing things are not absolutely contradictory but can transform into each other (Peng Hua, 2017). In other words, the debate between monism and pluralism does not mean the two are diametrically opposed: an absolute monistic moral essence (such as "benevolent love") can manifest as different concrete moral phenomena in different situations and cultures, such as fairness and care. Regardless of the specific manifestation, all aim to follow the moral value of "benevolent love." Therefore, from the perspective of traditional Chinese thinking, both "oneness" and "manyness" in morality can be tolerated simultaneously, and the two can transform and coexist. Finally, from the perspective of indigenous psychology, China' s "relationalism" culture differs from Western individualistic culture, causing Chinese morality to be based on interpersonal responsibilities rather than individual rights (Huang Guangguo, 2006; Yang Zhongfang, 2009). These factors may lead Chinese people' s understanding of moral essence to differ substantially from Westerners' . Research has found that Chinese people consider violations of "civility" to be immoral (Buchtel et al., 2015) rather than being based on harm or other factors as Westerners believe, and there is no clear conceptual distinction between "uncivilized" and "immoral" (Berniūnas et al., 2021). These considerations may all provide Chinese solutions to resolving the problem of moral monism versus pluralism.

*Note: Figure translations are in progress. See original paper for figures.*