Moral coherence processes: constructing culpability and consequences

Cory J Clark[1], Eric Evan Chen[2] and Peter H Ditto[2]

We review recent research in moral psychology that demonstrates a fundamental human motivation for a morally coherent world, that is, a world in which the moral qualities of actors and actions match the moral qualities of the outcomes they produce. The striving for moral coherence explains many seemingly contradictory patterns of judgment found in the moral reasoning literature, such as the general tendency for people to reverse engineer descriptive beliefs to fit desired prescriptive conclusions. Many recent phenomena in the moral reasoning literature demonstrate coherence-based reasoning, among them, the construction of morally culpable agents, the construction of victims and harms, and altered beliefs about the effectiveness and consequences of policies with moral implications.

Addresses
[1] University at Buffalo, United States
[2] University of California, Irvine, United States

Corresponding author: Ditto, Peter H (phditto@uci.edu, phditto@gmail.com)

When Hurricane Katrina struck the gulf coast of the United States in 2005 it caused massive damage to homes and property and killed almost 2000 people. Interestingly, many individuals, prominent political and religious leaders among them, chose not to blame the hurricane for the death and destruction. Instead, they saw it as divine retribution for the sins of the American people [1].

Blaming God and victims is a common reaction to natural disasters that reflects the ongoing human struggle to make sense of the moral world, often by understanding harmful events as caused by morally culpable agents and delivered upon only those who deserve it. In this essay we explore this fundamental desire for moral coherence, and argue that many phenomena in the moral reasoning literature reflect a struggle to construct coherent narratives in which the moral qualities of actions and actors match the moral quality of their outcomes.

Explanatory coherence

Social psychologists have long recognized that humans are fundamentally motivated to simplify and organize their social worlds such that beliefs and feelings about oneself and others fit together coherently [2–5]. This desire for cognitive consistency can motivate rational, evidence-based reasoning, but its popularity in the field has flowed primarily from its prediction of motivated or 'backward' forms of reasoning in which beliefs are, in effect, reverse engineered to produce the coherent fact patterns that people desire (e.g., adjusting one's attitudes to fit with past behavior).

Models of explanatory coherence are notable for their explicit incorporation of this notion of multidirectional causation [6,7]. Coherence-based models resemble classic cognitive consistency theories, but take a dynamic view in which beliefs, feelings, goals, and actions all influence one another, and are adjusted iteratively toward a point of maximal internal consistency or 'coherence' [8]. That is, a coherence perspective depicts people as striving to make sense of information available to them in a way that includes both 'rational' bottom-up influences (e.g., adjusting conclusions to fit facts) and less rational top-down ones (e.g., adjusting facts to fit conclusions). Coherence was originally conceived of in terms of the logical consistency between belief elements, but later work has recognized that people seek consistency not just between various beliefs, but also between their beliefs and their emotions, preferences, and motivations [9].

Moral coherence

Researchers interested in moral reasoning have also long recognized that individuals seek to construct coherent, emotionally satisfying views of the world. In the late 1960s, struck by people's inclination to blame victims of misfortune for their own fate, Melvin Lerner traced this tendency to a core desire to live in a world of just deserts [10,11]. According to Lerner, the internal logic of the moral world is one of justice, where people get what they deserve and deserve what they get. Unfortunately, events like natural disasters and other varieties of seemingly senseless victimization do not always comport with this moral logic, and so maintaining belief in a just world sometimes requires people to construct narratives in which victims deserve the misfortunes that befall them [12–14].

A recent view of moral judgment that explicitly captures the flavor of coherence-based reasoning is the dyadic view of morality championed by Gray and colleagues [15,16]. This view argues for a fundamental dyadic template underlying all moral judgments in which one individual (the agent) acts in a way that intentionally harms or helps a second individual (the patient). If either component of this template is not readily available (i.e., there is no obvious agent or patient), people construct them through a process of dyadic completion [15,17,18]. That is, exposure to harmed patients (e.g., victims of a natural disaster like Hurricane Katrina) motivates the search for a culpable agent (e.g., God [19]), and exposure to even ostensibly victimless acts (e.g., masturbation) motivates the search for harmed patients (e.g., the masturbator him or herself [20••]).

In the sections below, we review other recent research in moral psychology that demonstrates people's penchant for constructing coherent moral narratives by reverse engineering beliefs to fit desired moral conclusions.

Constructing culpability

Rationally, people should only receive blame for behavior that they intend and control [21,22]. This normative principle is well represented in formal legal systems (e.g., via consideration of premeditation, mental competence, emotional state) and reflected in the judgments of everyday people [23,24]. A wealth of research also demonstrates, however, that people engage in the reverse inference process: when motivated to blame and punish others, people construct morally culpable agents by adjusting their beliefs about intention, causation and control.

Attributions of intention

Infants as young as 6 months old attribute more agency for bad outcomes than for good ones [25••] and this same asymmetry has been found repeatedly in studies on adults' attributions of intention and related constructs. Research on the 'side-effect effect,' for example, demonstrates that side effects of identical actions are perceived as more intended when those side effects are morally bad (e.g., harmful to the environment) than when they are morally good (e.g., helpful to the environment) [26]. The side-effect effect has been demonstrated across cultures [27] and in children [28], is robust across numerous ways of phrasing the intention-related judgment ('deciding', 'in favor of', 'advocating', and 'opposed to' [29]), and also appears for direct effects of actions even when luck is said to have played a major role in bringing about the consequences [30,31]. The same asymmetry found for behaviors with negative and positive effects occurs for identical behaviors committed by liked and disliked others. People ascribe more negative motives to others with whom they disagree [32] and are more likely to make dispositional attributions when negative behavior is

committed by an outgroup member than an ingroup member [33].

Attributions of causality and control

People who perform morally harmful actions are also perceived as being more causally responsible for their outcomes compared to those who perform morally ambiguous or morally positive actions [34–37]. Like the asymmetry shown in intention-related judgments, this pattern is robust and has been replicated for a host of different judgments related to causation and control [38,39]. Two recent studies also reveal a hallmark of coherence-based reasoning: complicated or even contradictory patterns of judgments created as individuals struggle to bring judgments in line with one another. One study had people judge the causal responsibility of two agents who jointly contributed to an outcome. When one agent's actions were morally wrong, not only did people attribute more causal responsibility to him, but the other agent was perceived as less causally responsible for the outcome even though her actions did not vary [40]. Another study [41••] had participants read about a sailor who was forced by his captain to throw either passengers or cargo overboard. Participants judged that a sailor throwing passengers overboard was more free (less forced) than when he threw cargo overboard. At the same time, the captain was judged as forcing the sailor more (rather than less) when jettisoning passengers rather than cargo. Consistent with a coherence-based account, making one agent salient over another shifted the target of participants' motives to blame, leading to contradictory evaluations in which both the forcer and forcee were held maximally responsible for the morally reprehensible action of killing passengers.

Higher order culpability

The desire to assign responsibility for immoral actions can extend beyond one-time judgments about particular individuals to the human capacity for moral responsibility in general. Clark and colleagues [42•] found that exposure to the immoral actions of others led people to increase not only their belief that those specific actions were freely chosen, but also their belief that all of humankind is capable of free action. One study analyzed nation-level data and found that the higher a country's crime and homicide rates, the more citizens of that country believed in free will. The authors argue that exposure to immoral behavior increases the motivation to blame and punish, which in turn leads to increased belief in free will (because free will is generally seen as a prerequisite for moral responsibility and punishment [22]). As was shown with causality-related judgments, such coherence striving can sometimes produce seemingly contradictory patterns of judgments, such as the view that determinism and free will are compatible. When told to assume a completely deterministic universe, for example, people will absolve an individual of moral responsibility for morally neutral