



ELSEVIER

The side-taking hypothesis for moral judgment

Peter DeScioli

A recent theory proposes that moral judgment is an evolved strategy for choosing sides in conflicts. Evidence from moral psychology undermines previous evolutionary theories of morality focused on cooperation. I show how the side-taking hypothesis explains these otherwise puzzling patterns of moral judgment — especially its focus on actions rather than intended consequences. Moral judgment offers an alternative to choosing sides based on status or relationships by conditioning support on disputants' actions rather than their identities. Observers can use moral judgment to take the same side, avoiding costly fights among themselves, while dynamically changing who they support in different conflicts. Observers can benefit from moral side-taking even when condemnation harms other individuals and society.

Address

Stony Brook University, Stony Brook, NY, United States

Corresponding author: DeScioli, Peter (pdescioli@gmail.com)

Current Opinion in Psychology 2016, 7:23–27

This review comes from a themed issue on **Evolutionary Psychology**

Edited by **Steven Gangestad** and **Josh Tybur**

<http://dx.doi.org/10.1016/j.copsyc.2015.07.002>

2352-250/© 2015 Elsevier Ltd. All rights reserved.

Imagine you are alone on a desert island when a hulking man says to hand over your food supplies or he will kill you. You protest that his threats are morally wrong. How much protection could your moral judgment really provide? If you had to choose, would you defend yourself with a hand axe or moral arguments? Now imagine instead that there are fifty people on the island. In this case, morality might actually be the better weapon: a persuasive moral argument could rally dozens of armed defenders to your side.

A recent theory begins to uncover how something so seemingly feeble as moral judgment could possibly offer real advantages in evolutionary competition. DeScioli and Kurzban [1**] argued that this puzzle can be resolved by thinking about how humans choose sides in conflicts. A persuasive moral argument can launch a barrage of stones at an opponent. It can sever an opponent's relationships, cut off trade and supplies, and cast them into exile. However, these powers require an audience who will

listen to moral accusations. Hence, a critical challenge is figuring out why observers care about who is right and wrong, and why they can sometimes be persuaded to oppose wrongdoers, even their own family and friends.

Explaining moral judgment, not cooperation

Traditional evolutionary theories of morality focus on understanding cooperation [2–4]. These theories do an excellent job of explaining why humans and other animals care for offspring, cooperate in groups, trade favors, communicate honestly, and respect property. However, they do not explain moral judgment — why humans compute the wrongness of actions, announce moral judgments, condemn other people's violations, and debate over moral rules. For example, kin selection explains why many animals care for their own offspring; it does not explain why humans morally judge other parents who neglect offspring — an unusual behavior absent in other species.

Research from moral psychology reveals the information-processing structure of moral judgment (reviewed in [5]), which is the primary evidence for distinguishing its evolved functions [6]. Moral judgment takes as input a person's action and computes as output a wrongness value for that action. People then announce their moral evaluations, argue about the wrongness of actions, and show hostility toward wrongdoers.

Moral judgment's focus on actions is among its most distinctive and mysterious features. Similar to Kant's deontological philosophy, people often judge actions such as lying, theft, killing, premarital sex, or blasphemy to be morally wrong even when these actions can achieve better outcomes [5,7,8]. A well-known example is people's judgments that it is morally wrong to kill one person to save five people in the footbridge trolley problem [9]. Deontological judgment is puzzling because evolved decision mechanisms are usually consequentialist in that they evaluate the outcomes of actions rather than the actions themselves. For example, people's precautionary judgments aim to avoid injury (an outcome) rather than judging skydiving or bullfighting (actions) to be inherently injurious.

Why then do humans assign moral values to actions? Moral judgment is not required for cooperation: Many non-human organisms cooperate while few, perhaps none, assign moral values to actions. Neither is moral judgment a particularly effective executor of cooperation: Research shows that people frequently behave selfishly even when they judge the same behavior to be immoral

[10]. Another idea is that moral judgment enforces cooperation by punishing cheaters. But this theory too predicts consequentialist mechanisms, specifically punishment aimed at deterring harm. Instead, research shows that people's moralistic punishments are aimed at retribution rather than deterrence [11], and its targets include harmless and beneficial behaviors such as same-sex marriage, birth control, interest-bearing loans, and scientific inquiry.

In fact, moral judgment can be extremely destructive: moral condemnation reduces economic efficiency [12], incites political extremism [13,14], damages close relationships [15], radicalizes terrorists [16], motivates sexual-orientation hate crimes [17], and blocks medical services for pregnant women, people addicted to drugs, and HIV patients [18,19].

A common fallback for cooperation-based theories is to attribute evidence inconsistent with cooperation to mistakes in moral cognition due to erroneous information [20] or simple heuristics [21,22]. These proposals stand on weak theoretical ground because people's deontological "mistakes" are large in magnitude, highly structured, and absent in non-human animals as well as other types of human judgments (e.g., welfare, precautions, economics). They also stand on weak empirical ground. Haidt [5] tested and falsified a mistake hypothesis by finding that people condemn actions that even they themselves report to cause little or no harm. Kurzban *et al.* [23] tested and falsified a mistake hypothesis by finding that participants reported greater willingness to kill a brother to save five brothers, compared to killing a stranger to save five strangers, contradicting the claim that deontic resistance to killing is driven by altruism heuristics. DeScioli *et al.* [24] tested and falsified a mistake hypothesis by finding that people are more likely to choose harmful omissions when there is a threat of punishment, contradicting the claim that people's greater willingness to harm by omission than commission results from cognitive errors (see also [25,26]).

In short, theories of cooperation explain many features of human social life. If humans were Vulcans from *Star Trek*, these theories might also explain morality. However, the data show that humans are hybrid Vulcan-Kantians whose moral judgments focus not only on intended outcomes but also on the particular actions people choose. Our Kantian side formulates moral rules of action, argues about the rules, and creates the diversity of moral systems across cultures. It is the aspect of morality that is most distinctively human and challenging to understand.

Is morality mightier than the sword?

There is little evidence in other animal species of moral judgment, moral announcements, or moral accusations, much less moral debates. Non-human animals do not

argue about the morality of sharing food, paying debts, same-sex relationships, or military invasions, even when they exhibit these behaviors. To understand such an unusual trait, it might help to consider other distinctive features of human social ecology. One possibility is human conflict.

Conflict is much more complicated in humans than most animals because people not only support each other in fights, but also switch sides [27]. In most animals, fighting is largely a matter of assessing the power of a single opponent and knowing when to back down [28,29], but assessment is much more difficult when the opponent can recruit help from other individuals [30]. An opponent who initially appears weak can suddenly multiply in strength with the arrival of a crowd of supporters. This complexity applies not only to fistfights and arrow volleys but also to verbal arguments and gossip campaigns.

In the arena of human conflict, moral judgment is a powerful weapon. Accusations of wrongdoing such as lying, infidelity, or blasphemy can mobilize a mob of aggressors against an opponent. Moral condemnation turns cold shoulders toward colleagues, sends terrorists after political cartoonists, and pulls nations into costly wars.

In this social environment, it is easy to see why natural selection would favor psychological adaptations for using moral judgment against opponents, and for avoiding prohibited actions to escape the wrath of condemners. The difficulty is understanding why an audience cares about right and wrong, and why their punishment is directed at a diverse and changing list of particular actions.

Choosing sides

The audience members observing a human conflict face their own challenges. Humans live in dense social networks and individuals often have cross-cutting ties to both sides of a conflict. When both sides request support from the same person, that person will inevitably damage at least one relationship, creating a difficult choice. Further, to assess the consequences of their side-taking decisions, observers also need to evaluate how other observers will choose sides.

One side-taking strategy is bandwagoning — supporting the higher-status against lower-status fighter. This tactic curries favor with the more powerful disputant and joins with others using the same strategy. However, bandwagoning can set a dangerous precedent: higher-status individuals learn that they can easily exploit lower-status individuals, creating a threat to everyone except the top of the hierarchy.

A second side-taking strategy is alliance-building — supporting one's own supporters [27,31]. By teaming

up, individuals can combine their power against formidable single opponents. However, alliances create new problems. Snyder [32] showed how forming alliances is a social dilemma similar to an arms race. Everyone forms alliances to protect against everyone else's alliances. The result is that when conflicts occur they can quickly spread to include many supporters, creating larger and more costly fights. Moreover, when everyone supports their own allies, the sides tend to be evenly-matched and fights are more likely to escalate. Hence, alliances create a coordination problem among observers such that if they take opposing sides then they will all endure higher fighting costs.

Ethnographic research shows the costly consequences of alliances, especially in small-scale societies without strong policing institutions [33–35]. Trivial squabbles can quickly expand to hundreds of people brawling as outsiders intervene to support family and friends. People in many cultures are also obligated to avenge an ally's murder by killing their enemies, leading to endless cycles of violence [36,37]. Due to the high costs of supporting allies, evolution might have favored alternative strategies for choosing sides that allow observers, at least in some cases, to disregard their alliances and take the same side.

Moral judgment as a side-taking strategy

Moral judgment offers a third strategy for choosing sides [1**] (see Figure 1). Observers morally judge the actions of each disputant and then side against the person who chose the most morally wrong action. If a majority of observers use this strategy, and also share the same moral rules, then they will take the same side, such as opposing a liar, thief, or blasphemer. Choosing the same side reduces observers' fighting costs by avoiding evenly-matched, escalating disputes. Observers benefit by taking the same side because they coordinate their choices to avoid an evenly-matched and costly fight. Moreover, observers achieve coordination without empowering high-status individuals to exploit them in the future. That is, moral side-takers achieve *dynamic coordination*: They take the same side without always siding with the same person because their choices are based on actions rather than identities.

According to the side-taking hypothesis, humans assign moral values to actions so that when conflicts occur, observers can dynamically coordinate to choose the same side. Moral judgment focuses on actions in order to create an alternative to choosing sides based on who the disputants are, including their statuses and relationships. Moral judgment includes an ideal of impartiality [38**] for the same reason: Observers must (at least appear to) compromise their personal relationships if they are to avert divisive and escalating alliances. This is why people are sometimes willing to oppose family and friends for moral reasons. People announce and argue about moral

Figure 1

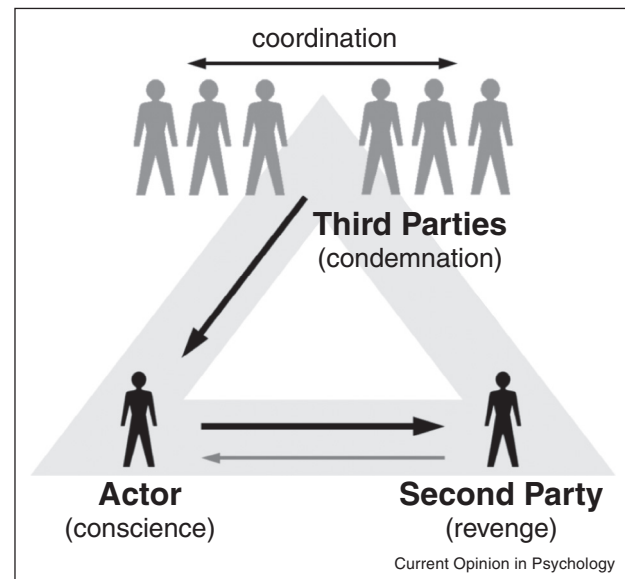


Diagram of a strategic interaction among two disputants and multiple third-party observers who choose sides. Third parties face a problem of coordination because if they choose different sides then they each pay the costs of fighting each other. Moral judgment allows third parties to choose the same side, minimizing their own costs, by opposing the disputant who has taken the most morally wrong action. Distinct psychological mechanisms (shown in parentheses) manage each role in the game.

From: DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, 139, 477–496, APA, reprinted with permission.

judgments [39] in order to negotiate side-taking. Humans show aggression toward wrongdoers because morality is built for battle, mainly for fighting other people's battles.

Humans debate moral rules because these rules determine who wins particular kinds of disputes. Moral debates are essentially meta-fights — fights over how fights will be settled. Experiments show that people advocate for the moral rules that most benefit them [40,41*,42,43,44**]. At the same time, group members need to agree on the moral rules if they are to serve a coordination function. This is why humans are averse to moral disagreements and anxiously seek consensus about moral rules.

A side-taking function allows for flexibility in the content of moral rules. From the observers' perspective, choosing the same side as others is paramount; which side they support is less important. For example, observers can use either a rule against stealing or a rule against refusing to share, even if the two rules lead to opposite wrongdoers in a dispute over resources, as long as they agree about which rule to apply. Like all coordination games [45,46], moral judgment has multiple equilibria because alternative and opposite rules can be equally effective at synchronizing

observers' decisions. This flexibility allows moral rules to change over time and to vary across cultures. It also allows for destructive rules because moral judgment is designed for synchronizing side-taking rather than promoting welfare.

Importantly, moral side-taking does not displace bandwagon or alliance strategies. Just as different moral regimes are equilibria, so too is moral side-taking only one among other equilibria for choosing sides. An individual's best side-taking move depends on how other group members choose sides, as well as the precise costs and benefits for themselves and their family and friends [47,48**]. Due to multiple equilibria, domains of social life vary in whether conflicts are resolved via hierarchy, alliances, or moral rules.

Conclusions

The side-taking hypothesis holds that moral judgment is a strategy for choosing sides in conflicts. Observers use moral judgment to choose the same side while also dynamically changing who they support by focusing on disputants' actions rather than their identities. The side-taking hypothesis helps understand why people blame and punish perceived wrongdoers, why moral judgment focuses on actions rather than intended consequences, and why moral condemnation damages relationships and societies.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. DeScioli P, Kurzban R: **A solution to the mysteries of morality.** •• *Psychol Bull* 2013, **139**:477-496.
Proposes the side-taking theory for moral judgment. The paper develops the side-taking game as a model of the adaptive problem that moral cognition functions to solve. It then applies the model to explain deontic moral judgment, moralistic punishment, and moral impartiality.
2. de Waal FBM: *Good Natured: The Origins of Right and Wrong in Humans and Other Animals.* Harvard University Press; 1996.
3. Ridley M: *The Origins of Virtue.* Viking: Penguin Books; 1996.
4. Wright R: *The Moral Animal.* Pantheon; 1994.
5. Haidt J: *The Righteous Mind.* Vintage Books; 2012.
6. DeScioli P, Kurzban R: **Mysteries of morality.** *Cognition* 2009, **112**:281-299.
7. Baron J, Spranca M: **Protected values.** *Organ Behav Hum Dec* 1997, **70**:1-16.
8. Tetlock PE: **Thinking the unthinkable: sacred values and taboo cognitions.** *Trends Cogn Sci* 2003, **7**:320-324.
9. Mikhail J: **Universal moral grammar: theory, evidence and the future.** *Trends Cogn Sci* 2007, **11**:143-152.
10. Batson CD, Thompson ER: **Why don't moral people act morally? Motivational considerations.** *Curr Dir Psychol Sci* 2001, **10**:54-57.
11. Carlsmith KM, Darley JM, Robinson PH: **Why do we punish? Deterrence and just deserts as motives for punishment.** *J Pers Soc Psychol* 2002, **83**:284-299.
12. Frank RH: *The Darwin Economy: Liberty, Competition, and the Common Good.* Princeton University Press; 2011.
13. Ryan TJ: **Reconsidering moral issues in politics.** *J Polit* 2014, **76**:380-397.
This paper reports survey studies and finds that moral conviction incites political extremism and hostility.
14. Skitka LJ: **The psychology of moral conviction.** *Soc Pers Psychol Compass* 2010, **4**:267-281.
15. Gottman JM, Silver N: *The Seven Principles for Making Marriage Work.* Random House; 1999.
16. Atran S: *Talking to the Enemy: Faith, Brotherhood, and the (Un)Making of Terrorists.* HarperCollins; 2010.
17. FBI: *Hate Crime Statistics.* 2012 <http://www.fbi.gov/about-us/cjis/ucr/hate-crime/2012>.
18. Marlatt GA, Larimer ME, Witkiewitz K (Eds): *Harm Reduction: Pragmatic Strategies for Managing High-Risk Behaviors.* Guilford Press; 2011.
19. Singer P: *The Real Abortion Tragedy.* Project Syndicate; 2012 <http://www.project-syndicate.org/commentary/the-real-abortion-tragedy-by-peter-singer>.
20. Turiel E: **The development of morality.** In *Handbook of Child Psychology: Social, Emotional, and Personality Development*, vol 3. Edited by Damon W, Eisenberg N. Wiley; 1998:863-932.
21. Sunstein C: **Moral heuristics.** *Behav Brain Sci* 2005, **28**:531-543.
22. Gigerenzer G: *Gut Feelings: The Intelligence of the Unconscious.* Viking Press; 2007.
23. Kurzban R, DeScioli P, Fein D: **Hamilton vs Kant: pitting adaptations for altruism against adaptations for moral judgment.** *Evol Hum Behav* 2012, **33**:323-333.
Experiments examine judgments about moral dilemmas to test whether deontic moral rules function as simple heuristics for altruism. The results contradict the altruism-heuristic hypothesis: participants report greater willingness to kill a brother to save five brothers, compared to killing a stranger to save five strangers.
24. DeScioli P, Christner J, Kurzban R: **The omission strategy.** *Psychol Sci* 2011, **22**:442-446.
25. DeScioli P, Asao K, Kurzban R: **Omissions and byproducts across moral domains.** *PLoS ONE* 2012, **7**:e46963.
26. DeScioli P, Bruening R, Kurzban R: **The omission effect in moral cognition: toward a functional explanation.** *Evol Hum Behav* 2011, **32**:204-215.
27. DeScioli P, Kurzban R: **The alliance hypothesis for human friendship.** *PLoS ONE* 2009, **4**:e5802.
28. Arnott G, Elwood RW: **Assessment of fighting ability in animal contests.** *Anim Behav* 2009, **77**:991-1004.
29. Pietraszewski D, Shaw A: **Not by strength alone.** *Hum Nature-Int Bios* 2015, **26**:44-72.
This paper uses scenario studies to examine children's inferences about power and alliances in conflicts.
30. Harcourt AH: **Coalitions and alliances: are primates more complex than non-primates?** In *Coalitions and Alliances in Humans and Other Animals.* Edited by Harcourt AH, de Waal FBM. Oxford University Press; 1992:445-471.
31. DeScioli P, Kurzban R, Koch EN, Liben-Nowell D: **Best friends: alliances, friend ranking, and the MySpace social network.** *Perspect Psychol Sci* 2011, **6**:6-8.
32. Snyder GH: **The security dilemma in alliance politics.** *World Polit* 1984, **36**:461-495.
33. Cooney M: *Warriors and Peacemakers: How Third Parties Shape Violence.* New York University Press; 1998.
34. Cooney M: **The privatization of violence.** *Criminology* 2003, **41**:1377-1406.
35. Phillips S, Cooney M: **Aiding peace, abetting violence: third parties and the management of conflict.** *Am Sociol Rev* 2005, **70**:334-354.

36. McCullough ME, Kurzban R, Tabak BA: **Cognitive systems for revenge and forgiveness.** *Behav Brain Sci* 2013, **36**:1-15.
37. Pinker S: *The Better Angels of Our Nature: Why Violence Has Declined.* Viking; 2011.
38. Shaw A: **Beyond 'to share or not to share': the impartiality account of fairness.** *Curr Dir Psychol Sci* 2013, **22**:413-417.
Reviews experiments showing the role of impartiality in fairness judgments. To appear fair, children prefer to waste resources rather than giving unequal amounts to two recipients. Children also prefer fair partners over reciprocating partners. These observations support the idea that fairness functions to signal impartiality rather than promoting welfare and cooperation.
39. Mercier H: **What good is moral reasoning?** *Mind Soc* 2011, **10**:131-148.
40. Aarøe L, Petersen MB: **Hunger games: fluctuations in blood glucose levels influence support for social welfare.** *Psychol Sci* 2013, **24**:2550-2556.
41. DeScioli P, Massenkoff M, Shaw A, Petersen MB, Kurzban R:
• **Equity or equality? Moral judgments follow the money.** *Proc R Soc B Biol Sci* 2014, **281**:20142112.
Reports experiments using economic games and finds that people advocate moral rules that most advantage them over others. Participants decide whether to divide earnings equally or unequally based on different amounts of effort contributed. Participants' judgments about the morality of each rule change in minutes as they learn their role to favor the rule that gives them more money.
42. Kurzban R, Dukes A, Weeden J: **Sex, drugs and moral goals: reproductive strategies and views about recreational drugs.** *Proc R Soc B Biol Sci* 2010, **277**:3501-3508.
43. Petersen MB, Aarøe L, Jensen NH, Curry O: **Social welfare and the psychology of food sharing: short-term hunger increases support for social welfare.** *Polit Psychol* 2014, **35**:757-773.
44. Tybur JM, Lieberman D, Kurzban R, DeScioli P, Disgust: **Evolved function and structure.** *Psychol Rev* 2013, **120**:65-84.
Develops a novel evolutionary and computational theory for the emotion of disgust, including pathogen, sexual, and moral disgust. The account for moral disgust proposes that an individual's pathogen disgust and sexual disgust toward a behavior indexes their own fitness value for that behavior and hence enters into people's cost-benefit calculations when deciding whether to moralize and condemn particular actions.
45. Schelling TC: *The Strategy of Conflict.* Harvard University Press; 1960.
46. Thomas KA, DeScioli P, Haque OS, Pinker S: **The psychology of coordination and common knowledge.** *J Pers Soc Psychol* 2014, **107**:657-676.
47. Jensen NH, Petersen MB: **To defer or to stand up: how offender formidability affects moral outrage.** *Evol Psychol* 2011, **9**:118-136.
48. Petersen MB: **Moralization as protection against exploitation: do individuals without allies moralize more?** *Evol Hum Behav* 2013, **34**:78-85.
Analyzes data from cross-national surveys and finds that people who have less friends and allies are more likely to moralize other people's behavior. The results support the hypothesis that moral judgment and alliances are alternative strategies that people deploy selectively depending on their own relative strengths and cost-benefit calculations.